



Education Policy Analysis Archives/Archivos
Analíticos de Políticas Educativas

ISSN: 1068-2341

epaa@alperin.ca

Arizona State University
Estados Unidos

Yeh, Stuart S.

Limiting the Unintended Consequences of High-Stakes Testing

Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, vol. 13, 2005, pp. 1-21

Arizona State University

Arizona, Estados Unidos

Available in: <http://www.redalyc.org/articulo.oa?id=275020513043>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System

Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal

Non-profit academic project, developed under the open access initiative

EDUCATION POLICY ANALYSIS ARCHIVES

A peer-reviewed scholarly journal

Editor: Sherman Dorn

College of Education

University of South Florida

Volume 13 Number 43

October 28, 2005

ISSN 1068-2341

Limiting the Unintended Consequences of High-Stakes Testing¹

Stuart S. Yeh

University of Minnesota

Citation: Yeh, S. S. (2005). Limiting the unintended consequences of high-stakes testing. *Education Policy Analysis Archives*, 13(43). Retrieved [date] from <http://epaa.asu.edu/epaa/v13n43/>.

Abstract

Interviews with 61 teachers and administrators in 4 Minnesota school districts suggest that, in their judgment, Minnesota's state-mandated tests were well-aligned with curricular priorities and teachers' instructional goals, emphasizing critical thinking as well as competencies needed to pass the Basic Standards exit exam, and avoiding the type of recall item that would require drill and memorization. This result, in combination with a survey showing that 85 percent of Minnesota teachers support the exit exam, suggests that Minnesota has been unusually successful in designing a high stakes testing system that has garnered teacher support. The success of Minnesota's model suggests that unintended narrowing of the curriculum due to high stakes testing may be avoided if pressure on teachers to narrow the curriculum is reduced through well-designed, well-aligned exams.

Keywords: testing; assessment; accountability; curriculum; policy analysis; large-scale achievement tests; high stakes.

¹ I thank Karen Wrobbel for assistance in conducting the interviews and Karen Seashore and anonymous reviewers for comments on a previous draft of the manuscript.



Readers are free to copy, display, and distribute this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives**, it is distributed for non-commercial purposes only, and no alteration or transformation is made in the work. More details of this Creative Commons license are available at <http://creativecommons.org/licenses/by-nc-nd/2.5/>. All other uses must be approved by the author(s) or **EPAA**. **EPAA** is published jointly by the Colleges of Education at Arizona State University and the University of South Florida. Articles are indexed by H.W. Wilson & Co. Accepted under the editorship of Sherman Dorn. Send commentary to Casey Cobb (casey.cobb@uconn.edu) and errata notes to Sherman Dorn (epaa-editor@shermamdorn.com).

Introduction

The No Child Left Behind (NCLB) Act of 2001 requires annual testing and establishes sanctions for schools that do not maintain “adequate yearly progress” in improving student achievement. Initial estimates are that 80–90 percent of all schools receiving Title I funds will be found “in need of improvement” by the federal government and subject to NCLB sanctions (Bracey, 2002). A concern is that this type of high-stakes testing may be fundamentally flawed, that it may narrow the intended curriculum—defined as the knowledge, skills and habits of thought teachers believe are important for students to learn—causing teachers to “teach to the test” (McNeil, 2000).

However, narrowing of the curriculum may not always be undesirable (Smith & O'Day, 1991). Alignment of state standards, tests, curriculum and professional development may be necessary to ensure that students are taught and tested on the content and skills they are expected to know and be able to do. Well-designed tests could help teachers to identify and address areas of weakness, improving the quality of instruction as well as providing accountability information. It may be reasonable to narrow the curriculum by focusing teachers' efforts and emphasizing topics that are valued (Smith & O'Day, 1991). This reasoning has engendered broad support and a decade of bipartisan legislative initiatives culminating in NCLB (Jennings, 2000). Whatever its flaws, NCLB seems unlikely to be repealed (Elmore, 2004). Currently, however, there is an opportunity to modify the legislation and its implementation to reduce unintended negative consequences. The focus of this study is to understand what can be done to avoid unintended, negative consequences due to narrowing of the curriculum in a high-stakes testing environment.

In Minnesota, high stakes testing was implemented in 1998 and diplomas were withheld from students who failed the state Basic Standards Test (BST) beginning in 2000 (sample tests and items available at http://education.state.mn.us/mde/Accountability_Programs/Assessment_and_Testing/Assessments/index.html). Surprisingly, however, an overwhelming 85 percent of Minnesota teachers support the State's exit exam (Draper, 2000). It is important to understand what may contribute to these positive views. It is unlikely that teachers would support the exit exam if they believe that it forces them to eliminate valuable chunks of the curriculum. If teachers in Minnesota do not feel pressure to narrow the curriculum in a way that reduces educational quality, it is important to understand their reasoning. This reasoning is likely to explain their choices about what to teach in their classrooms—choices that ultimately determine the extent to which the curriculum may or may not be narrowed. If we understand their reasoning, perhaps we can understand the conditions under which the curriculum may be narrowed—and how excessive narrowing can be avoided. The study reported below suggests that the unique design of Minnesota's testing program may be well-aligned with teachers' expert judgment about what should be emphasized. This design may contribute to teacher satisfaction with high stakes testing, and may provide lessons about how to implement high stakes testing in a way that minimizes unwanted narrowing of the curriculum.

Four Concerns

There are four main concerns related to narrowing of the curriculum and the fear that improved test scores merely reflect teaching to the test, rather than broad gains in learning. One concern is that non-tested subjects, such as art, may be de-emphasized (McNeil & Valenzuela, 2000;

McNeil, 2000; Smith, 1991; Smith & Rottenberg, 1991). This might be addressed by requiring schools to provide elective or required instruction in those areas.

A second concern is that even within tested subjects, particular topics or skills not included on the test may be de-emphasized (McNeil & Valenzuela, 2000; McNeil, 2000; Smith, 1991; Smith & Rottenberg, 1991). This might be addressed by ensuring that the most important topics and skills are covered by the state-mandated test.

A third concern is that testing narrows curriculum and instruction to memorization of bits of factual knowledge (McNeil & Valenzuela, 2000; McNeil, 2000; Smith, 1991; Smith & Rottenberg, 1991). This may occur if state curriculum standards and tests aligned to those standards attempt to cover too much factual knowledge. An unanticipated consequence is that teachers may feel it is necessary to drill students on the universe of factual items that may appear on the test, in the hope that students will correctly answer those items that actually appear. Addressing this issue may require changes in state curriculum standards as well as the state test.

A fourth concern is that testing may cause teachers to spend excessive time on test-taking tricks and strategies that have little value outside of the testing situation (McNeil & Valenzuela, 2000; McNeil, 2000; Smith, 1991; Smith & Rottenberg, 1991). For example, students may be taught to make sure that bubble sheets are correctly completed. This issue might be addressed through tests designed to minimize unwanted trickiness.

The first concern might be addressed through policy and the last concern might be addressed through proper test design. The focus of this paper is the second and third concerns: the construct under-representation that occurs when tests only assess aspects of the intended curriculum that are easy to measure (Baker & Linn, 2004). The question is whether it is indeed possible to design and implement large-scale assessments, suitable for state-mandated testing, that are properly aligned with the curriculum in order to reinforce intended learning outcomes (Linn, 2004; Rothman, 2004). A review of state-mandated tests and state curriculum standards found that while there is superficial alignment of test content with state curriculum standards in many states, a more careful analysis suggests that lack of alignment may indeed be a problem (Rothman, 2004; Rothman, Slattery, Vranek, & Resnick, 2002).

Effects of State-Mandated Testing on the Curriculum

Two reviews of research found that in some cases—but not others—state-mandated tests narrow the curriculum (Cimbricz, 2002; Mehrens, 1998). To a certain extent, variation in findings reflect individual, school and district-level differences in experiences, training and attitudes with regard to high-stakes testing (Cimbricz, 2002; Grant, 2000). Some individuals have positive attitudes, feeling that testing can be used to diagnose and address weaknesses, and are well-trained to use test results in this way (Massell, 2001). Others have negative attitudes and are not trained or have little experience in using test results for formative purposes. A path model suggests that both teacher attitudes and professional development mediate the effects of testing programs on instructional practices (Pomplun, 1997). However, it seems unlikely that all of the variation in outcomes is due to individual differences. Each state has control over the design and implementation of its state test and it seems likely that those decisions systematically influence the degree to which the curriculum is narrowed. Unfortunately, existing research does not provide a strong basis for specifying exactly what factors or conditions are important.

Stake Levels

A common supposition is that differences in stake levels explain variation in outcomes. It is hypothesized that low stake levels are associated with little or no narrowing, while high stakes are associated with greater pressure on teachers and more narrowing of the curriculum. There is some evidence in support of this hypothesis. During a period of low stakes, when the results of state-mandated tests were rarely tied to decisions about student promotion or graduation, or threats to reorganize schools, Porter, Floden, Freeman, Schmidt, and Schwille (1986) concluded that: "Another myth exposed as being only a half truth is that teachers teach topics that are tested. Little evidence exists to support the supposition that national norm-referenced, standardized tests administered once a year have any important influence on teachers' content decisions" (p. 11). Kuhs et al. (1985) concluded that Michigan teachers' topic selection was not significantly influenced by the state minimum competencies test or the district-used standardized tests.

In contrast, several studies of Texas' high-stakes testing program concluded that preparing students for state-mandated tests narrows the curriculum (Gordon & Reese, 1997; Haney, 2000; Hoffman, Pennington, Assaf, & Paris, 1999; McNeil & Valenzuela, 2000). Smith and Rottenberg's (1991) study of two elementary schools in Arizona also concluded that high-stakes testing causes narrowing of the curriculum as teachers align instruction with test content.

However, high stake levels do not always result in narrowing of the curriculum. Grant (2001) compared the teaching of two high school teachers in New York, where passing the state Regents exams entitled students to a prestigious Regents diploma, and passing the easier Regents Competency Test was necessary to obtain a regular diploma. In this high-stakes situation, Grant found little direct influence of testing on content or pedagogical decision-making and concluded instead that the teachers' personal beliefs governed their teaching (also see Grant, 2000). Similarly, a study of Kentucky's reform efforts concluded that teachers' responses to these efforts in four exemplary schools were shaped by the participants' shared vision of curriculum and commitment to children, rather than by high-stakes testing (Wolf, Borko, Elliott, & McIver, 2000). Notably, in both of these cases, it could be argued that while stakes were high, pressure on teachers was low. In New York, students who failed the more demanding Regents' exams were not denied a diploma as long as they passed the minimum competency Regents' Competency Test. And in the four exemplary Kentucky schools, it appears that strong administrative leadership deflected pressure from teachers and strong teaching served to raise test scores and allay parent concerns. In other words, effective teaching may have reduced the number of students who failed the state-mandated test and reduced pressure to narrow the curriculum.

Performance-Based and Portfolio Assessments

If individual differences and differences in stake levels do not adequately explain variation in the impact of state-mandated tests, what about characteristics of the tests themselves? Some researchers are concerned that multiple-choice tests are inherently poorly-aligned with sound teaching practices, causing teachers to "dumb-down" their instruction (McNeil & Valenzuela, 2000; McNeil, 2000; Smith, 1991; Smith & Rottenberg, 1991). In view of this, it has been suggested that performance-based or portfolio assessments are more likely to shape instruction in desirable ways, to encourage teaching that promotes active learning, compared to multiple-choice assessments, and to avoid negative narrowing of the curriculum (Baron & Wolf, 1996; Rothman, 1995). In principle, well-designed performance-based and portfolio assessments could be aligned with sound teaching

practices and would provide information about student progress regarding important basic and critical thinking skills that teachers would use to improve instruction.

However, evidence regarding this hypothesis is mixed (Mehrens, 1998). A qualitative study of the effects of testing in Maine and Maryland found that performance-based assessments have limited effects on changing basic instructional strategies when stakes are low or moderate (Firestone, Mayrowetz, & Fairman, 1998). In Kentucky, a survey of teachers found that 90 percent agreed that portfolios made it difficult to cover the regular curriculum, and teacher opinion about the effect of portfolios on instruction was evenly split (Koretz, Barron, Mitchell, & Stecher, 1996). Fewer than 45 percent of principals and teachers reported that KIRIS (Kentucky's portfolio-based assessment program) provided a better view of school effectiveness compared to more conventional, commercial standardized tests. The open-response component was rated as having positive effects on instruction more often than performance events and portfolios, suggesting that it may not be necessary to use costly, time-consuming, and difficult-to-score forms of assessment in order to achieve instructional improvement. A RAND study of Vermont's portfolio system concluded that the use of portfolio assessment had substantial positive effects on fourth grade math teachers' teaching practices, but evidence of validity was unpersuasive and unreliability in test scores precluded most of the intended uses of the scores (Koretz, Stecher, Klein, & McCaffrey, 1994). In Arizona, teacher opinion regarding the consequences of the state's performance-oriented assessment program was divided (Smith et al., 1997). To the extent that it occurs, narrowing of the curriculum may happen independently of the format of the test (Haney & Madaus, 1989; Mehrens, 1998). The research reviewed above suggests that performance- and portfolio-based assessments are unlikely to be panaceas for negative consequences of testing.

Research Questions

The present study investigated the responses of teachers and administrators in four districts in Minnesota, and addressed the following research questions: 1) What are the views of teachers and administrators in the four Minnesota districts regarding state-mandated testing? and 2) In what ways does Minnesota's state-mandated testing program influence teacher and administrator choices about what is taught in their classrooms and schools?

Minnesota's State-Mandated Assessments

Minnesota requires all districts within the state to administer two types of tests. The Minnesota Comprehensive Assessments (MCA) are high standards tests. At the time of the research study, they were administered at grades 3 (reading and math), 5 (reading, math, and writing), 10 (reading and writing) and 11 (math). The MCA tests will also be implemented in grades 4, 6, 7, and 8 beginning in 2005-2006. The reading and math tests are composed of multiple-choice and short answer questions. The writing test requires a written essay. The MCAs are designed to be aligned with Minnesota Academic Standards, but no information is available about the actual degree of alignment.

At the time of the research study, the Minnesota Basic Standards Test (MBST or BST) was a minimum competency exit exam administered in grade 8 (reading and math), 10 (writing), and 12

(for seniors who needed to re-take the MBST).² Students were required to pass the test in order to graduate from high school. The test was designed to ensure that all graduates possess basic skills in reading, writing and math, and was composed of multiple choice questions and a written essay. Cumulative passing rates on the MBST (after five attempts) for all Minnesota students were high: 99.1% in math and 99.5% in reading/language arts (Gayler, Chudowsky, Hamilton, Kober, & Yeager, 2004). Because the MBST exit exam was not intended to be aligned to Minnesota Academic Standards, no alignment review had been conducted (Gayler et al., 2004).

The dual system of using high-standards tests as well as a basic skills exam is a reflection of the desire to test both levels of skills but to make high school graduation contingent on passing a basic skills exam. The MCAs and MBST are representative of the current generation of state-mandated tests regarding their format (largely multiple-choice with some open response and essay items), content (reading, math, and writing), and intended use (auditing school performance). At the time of the study, there were no accountability consequences or rewards for teachers, schools or districts linked to student performance on the MCAs or exit exam. The state did not provide special programs of support to schools and teachers in the form of professional development aligned to the assessments and standards. Sample tests and items are available at http://education.state.mn.us/mde/Accountability_Programs/Assessment_and_Testing/Assessments/index.html.

Methods

Data reported here come from a qualitative study of teacher and administrator responses to state-mandated testing in four Minnesota districts during the 2002–2003 school year. Minnesota was selected for two reasons. First, the state testing program is “high stakes”—students must pass the Basic Standards Test, administered beginning in 8th grade, in order to graduate from high school. Therefore, it was possible to investigate the effects of a high-stakes testing program on curriculum and instruction. Second, Minnesota has an unusual system where the Minnesota Comprehensive Assessments (MCAs) are challenging “high standards” tests, but the 8th grade exit exam is a minimum competency test. Unlike systems that rely primarily on minimum competency tests, Minnesota’s dual approach may be less susceptible to narrowing of the curriculum.

Four school districts were selected that were representative, in terms of student demographics, size and distance from major urban centers, of four major categories: 1 large urban (40,000+ students, 74% minority, 68% free/reduced lunch), 1 small urban (<2000 students, 62% minority, 37% free/reduced lunch), 1 suburban (11,000 students, 28% minority, 8% free/reduced lunch), and 1 semi-rural district (<2000 students, 2% minority, 4% free/reduced lunch). Each district represents the challenges that face districts in the corresponding category. These particular challenges may influence the manner and degree to which the curriculum is narrowed. The large urban district represents the challenges of coordinating the education of a large, extremely diverse student population, many of whom are at risk of not passing the state-mandated test. The small urban district represents the challenges of meeting the needs of a diverse student population without the resources that are available in a large district—such as a dedicated director of testing. The suburban district represents the challenges of maintaining a challenging curriculum and meeting the educational needs of students, most of whom are likely to pass the state-mandated test and should not be subjected to remedial instruction designed for students who are at-risk of not passing the

² The MBST is being replaced by the MCA-II high standards graduation test, starting with the class of 2010.

tests. The semi-rural district represents the challenges of meeting student needs with very limited resources.

In each district, interviews were obtained with the Superintendent or Assistant Superintendent, Director of Testing, and the principal and four teachers at 1 high school, 1 middle school and 1 elementary school (except in one district that only had two schools). Schools were selected that represented the socioeconomic characteristics (ethnicity and percent receiving free/reduced price lunch) of each district. One elementary school declined to participate after the principal learned that teachers would not be compensated for the interviews (a second elementary school was substituted). The researcher has no reason to believe that this affected the results.

Table 1
Distribution of Sample

Interviewees	District Type				Total
	Large Urban	Small Urban	Suburban	Rural	
Superintendent	1	1	1	1	4
Director of Testing	1	a	1	a	2
Principal/Assistant Principal	3	2 ^b	3	3	11
Elementary Teacher					
Regular	3	3	4	4	14
Special Education	1	1	1	0	3
Middle School Teacher		a			
Math	1		1	3 ^c	5
Language Arts	1		1	3 ^c	5
Science	0		1	3 ^c	4
Special Education	2		0	1	3
High School Teacher					
Math	1	2 ^c	1	1	5
Language Arts	1	2 ^c	1	2	6
Science/social studies	2	0	1	1	4
Special Education	0	1	1	0	2

^a Not employed in this context.

^b Single middle/high school.

^c Teachers teach multiple subjects.

Teachers were stratified by subject (math and reading/language arts) and tested grade, and randomly selected. In addition, science teachers were randomly sampled in order to assess the impact of testing in a non-tested subject. All participants had bachelor's degrees and approximately one-third also held masters' degrees (six administrators held doctorates). Years of experience in the classroom ranged from 1–29 years. Approximately 60% of the participants were female, and 40% were male. A total of 61 individuals were interviewed. Table 1 presents the distribution of the research sample.

Data Collection

Development of the semi-structured interview protocol (Appendix A) was informed by three rounds of previous interviews with approximately 100 teachers and principals in three states regarding the effects of state-mandated testing. Each participant was interviewed individually.

Interviews were audio recorded and lasted about 50 minutes. The interview protocol included general questions about the impact of testing as well as specific questions about the influence of testing on teaching and learning: “Overall, is your opinion about the impact of the state testing program positive or negative? Does the state testing program influence the type of skills and knowledge that you focus on? Are these skills and knowledge important? Is the test designed so you feel a need to drill students on factual material?”

Data Analysis

In keeping with established methods of qualitative research (Bogdan & Biklen, 1992; Glaser & Strauss, 1967; Miles & Huberman, 1994; Strauss, 1987), data analysis followed the three-part modified constant comparative method. Each interviewer asked follow-up questions to make certain that the meanings respondents intended were understood. As interviews were conducted, emergent findings were checked through follow-up questions with subsequent interviewees. Interviews were transcribed and coded to identify primary themes. The themes were identified jointly by two raters and cross-checked by a third researcher. Discrepancies were resolved through discussion. Interview excerpts were selected that illustrated views held by the majority of participants.

Limitation of the Study

A limitation of this study is that it relies on teacher and administrator judgments regarding the effect of Minnesota’s tests on curriculum and instruction. These judgments may be biased. However, two recent studies suggest that, in some cases, self-report data may be a reliable indicator of actual instructional practices. One study compared teacher self-reports and independent observations of teachers’ enacted curriculum and found strong agreement between the two measures (Porter & Smithson, 2001). A second, national study involving 25,000 teachers found close agreement between teachers’ self-reported instructional practices and students’ perceptions of those practices (Shim, Felner, Shim, & Noonan, 2001).

Results

What is the Effect of Testing on Curriculum and Instruction?

What are the views of teachers and administrators in the four Minnesota districts regarding state-mandated testing? By an overwhelming two-to-one margin, most interviewees answered that the overall impact of state-mandated testing is positive. This margin was consistent across district type (urban, suburban, and rural), and grade level. All administrators answered that the overall impact of testing is positive. The results are consistent with a random survey of teachers across the state, which found that 85 percent support the Basic Standards exit exam (Draper, 2000).

Interviewees who viewed testing positively exhibited four patterns when responding to this question. First, teachers felt that testing improved the quality of the curriculum, although they initially feared that testing would have a negative impact, causing teachers to focus only on basic skills. According to interviewees, test results force teachers to acknowledge that many students are below grade level in math and reading, prompting greater collaboration among teachers and

administrators to improve the quality of the curriculum and to align the curriculum vertically from grade to grade. For example, a principal in a suburban middle school explained:

The negative part about it would be if teachers and staff were simply addressing the basics for the kids—teaching the basics, not taking them to the next level because that test has become so important, and that would be their only focus in the classroom. I think there was a lot of initial talk about that when the testing first started. I don't see that happening here in our middle school now that I have that opportunity to be in classrooms... The positive part is that we are looking at education differently and that we're not working in isolation. Things are on the table, so to speak, and so we can work cooperatively as a team and make differences in the curriculum that really enrich the curriculum and enrich the teaching for the kids and the staff. So I see that as the positive side of it. There's more dialogue, there's connection between K–12 curriculum, there's connection between the levels of high school, middle school, [and] elementary.

Interviewees in the majority group asserted that teachers prepare students for the test by integrating the skills needed to pass the test into the school curriculum throughout the year, rather than through isolated test preparation that narrowly focused on the types of items expected on the test. A principal in a semi-rural elementary school explained:

I think the test forces us to look at what's important... it's not in a way that the teachers teach to a test, but I think they each in their own way make sure that they bring to the forefront in their instruction the skills that enable kids to take the test without a high level of anxiety. I think they do a great job of that, and they don't do it a week before or two weeks before; I think they do it in a timely fashion so that we can honestly say that we're not teaching to the test. We're teaching in a way that children can learn.

Second, interviewees in this group felt that testing has made them more accountable and improved the quality of instruction. Teachers believed they were more focused, goal-oriented and reflective about what they needed to teach. Teachers reported more dialogue and communication among themselves. Professional development and staffing are said to be more focused on improving student achievement. Teachers report that after-school intervention programs have been initiated and focus on the needs of low-achieving children. For example, a high school science teacher in an urban school suggested that testing has impacted instruction through a number of different ways:

I think it has been wonderful. I think it has raised a number of eyebrows, and it has caused people to step back and ask themselves, 'What is it that I'm doing in the classroom that they're asking me on a test, and why are they asking.' In other words, it's causing a lot of self-reflection that teachers often don't have time to do. [They do] a lot of planning but often not a lot of time reflecting... It's impacting professional development such that now our professional development is geared more towards the impact on the classroom, not just on me going out to a meeting of the National Science Teachers Association meeting in Hawaii. Now I have to relate it to how it's going to impact the class. So it's impacted the professional development as well. Obviously it has impacted funding at the district level. We now have specialists in the classrooms or in the buildings. And I think it's also raised awareness that you have some kids that are special ed and kids that are not, and those kids that are outside of special ed still need help, and we're seeing that help is made available to them.

Third, teachers felt that testing improved student attitudes, engagement, and effort by holding students accountable for learning. Teachers noticed that more children are seeking help and participating in after-school programs. An 8th grade math teacher in a suburban middle school explained that her view of testing is positive because it causes students to adopt a new attitude toward learning:

Overall, positive...[because] it is finally something that the kids are accountable for. Up through middle school, you are not accountable for anything in Minnesota really. You can fail, fail, fail, fail, and still go into high school. In high school, you're accountable. So this is finally something where the kids say, "Oh my gosh, this counts." So I like that; I like watching the attitude change because for the first time, it counts.

Fourth, teachers felt that testing resulted in greater efforts to ensure that all children succeed and improved student achievement. For example, a 4th grade teacher in a semi-rural elementary school described a program to identify children with low reading scores and to refer them to a special after-school program:

Here in [district name] our children coming out of third grade and into fourth grade that have scored low in the reading area, they are identified and are given the opportunity to be in an after school program called "Soar to Success." I think it's an 18 week program, runs from 3:00 to 4:00. I've seen a lot of kids have great success with the program...It's one of the programs that have come out of the low scores from the state testing. [So you view that as a positive?] Yes, very, very much so, very much so a positive. If nothing else, their self-esteem while they're in that program, they're getting their homework done, they're learning different ways to attack reading.

In What Ways Does Testing Narrow the Curriculum?

In the previous section, interviewees reported that testing led to focusing of the curriculum. Given the concern among educators about the effects of narrowing of the curriculum, the second question addressed in this study is the following: In what ways does Minnesota's state-mandated testing program influence teacher and administrator choices about what is taught in their classrooms and schools? The themes reported below characterized the views of the majority of interviewees who felt, by a two-to-one margin, that state-mandated testing had a positive impact on the curriculum. These themes were consistent across district type (urban, suburban and rural).

A focus on basic skills at the 8th grade level. As might be expected, interviewees stated that teachers focused on basic skills at the 8th grade level, consistent with the emphasis of the Basic Standards Test on minimum competencies. Teachers also focused on particular topics that were emphasized on the state test. What was unexpected was that teachers in the majority group—across suburban, urban, and rural districts—felt that this focus was appropriate. An 8th grade math teacher in a suburban middle school explained:

Absolutely! The first year we took the test, we spent that summer rewriting our curriculum to align with [it]. We broke it down; we took the eight strands and we analyzed which strands [our district's] eighth graders did the poorest on. We took those strands and we put them in the front of the year, so that we had them done by February. Ratios, proportions and percents were the weakest area, so we did that right before the test. [Do you feel that these skills and knowledge that you

teach are important?] Yes, all of those eight different strands are necessary skills for kids. So it's not like I'm teaching surfing or something totally irrelevant. It is relevant to their daily lives.

Although teachers focused on basic skills, they did not feel overly constrained. Both English language arts and math teachers in the majority group viewed the state test as well-aligned with their instructional priorities and believed that they were teaching vital skills that were applicable beyond the test. An 8th grade English teacher in an urban secondary school explained:

I teach them basic skills that I know will be helpful. . . For example, making sure that when you're reading an article, that you look at some of the questions first so that you're not just reading random information...They're good skills to have...I don't necessarily teach to the test so much as I teach some skills that I think will be beneficial for the test [as well as] in other areas.

A math teacher in a semi-rural high school added:

As an adult, and as an educator, yes, I feel [that the skills measured by the Basic Standards Test are] important, because without those skills, most people are not going to be able to make good judgments throughout their lives. With those skills, they have more of an equal opportunity to make good judgments and decisions... [And] I think the grad standards in mathematics as a whole have done a lot to improve math education in the state of Minnesota. I think without it, there'd be many school districts still doing what they did twenty to thirty years ago, because why change if you don't have to?

Teachers in this group believed that the test covers basic math skills that all students should learn. As an 8th grade teacher in an urban elementary school added: "It's all part of our curriculum...Those things are supposed to be basic skills that everyone would have. So I really don't have a problem with the test." An 8th grade math teacher in an urban secondary school put it succinctly: "The topics on the test are, I think, topics that students should know."

These views were shared without exception by all of the administrators who were interviewed, across rural, suburban and urban districts. Superintendents agreed that the test has focused the curriculum on important skills and knowledge at every grade level. A superintendent in a semi-rural district said:

Definitely in math and reading... it certainly has influenced what we're looking at, what's being taught in the curriculum, and then where we have gaps, we're purchasing additional curriculum resources and materials. [Do you think that the skills and knowledge that you teach to help kids prepare for the state test, are those the important things that they need to know?] I think I would agree, and most of our teachers would agree, and principals in the area agree.

Principals believed that as a result of focusing the curriculum on important basic skills, fewer students are failing to learn those skills. As a principal in a suburban middle school explained:

Some of those fluffy extraneous things [have been eliminated from the] curriculum—people just doing what they want to do because it's fun. So I think that's been a plus, and I think that we're focusing on what needs to be done for kids to have their skills for future lifelong use. Focusing on the basics is one of them, and having students be able to be successful and function is a basic. Kids aren't as easily dropping through the cracks as they used to be. So I think that's important that we have kind of solidified some things that we need to address

with students for their success...And it definitely has focused more [attention] on the skills that the kids need.

Interviewees in the majority group agreed that testing focused instruction on reading and math—the tested subjects—and felt that this focus over nontested subjects was appropriate. As a Principal in an urban middle school explained:

Teachers do focus on reading and on math, and our students, approximately half of whom are special needs students because they're either special ed or English language learners new to the English language, they have to learn it. . . You have to learn how to read. You can't learn anything if you don't.

Science teachers agreed with the focus on reading and math, welcomed the opportunity to integrate reading and math into science instruction, and asserted that integrated instruction in these areas was desirable. A high school science teacher in an urban school reported:

It has [influenced the type of skills and knowledge that I focus on] because now I have to teach math and reading in my science classroom. I cannot just assume that by giving a textbook at the tenth grade reading level that these kids know how to read at the tenth grade reading level. [How do you feel about having to teach reading and math?] I love it! If I had my way, reading, writing, arithmetic would never have come out of anybody's curriculum. It would be early childhood family education through grade 16. It would be in every single classroom, every single day. It would not be compartmentalized. The thought is totally silly for me that it's even fragmented; it makes no sense.

Teachers in the majority group did not feel that they needed to focus on test-taking tricks and strategies to prepare students for the state test. Instead, they asserted that they prepared students by teaching skills that are broadly useful. For example, a Title I Reading Coordinator in an urban middle school said: "It's just general reading skills...what we focus on is just good reading instruction and helping the kids with test-taking skills, because some kids really need help with that. Good test-taking skills, ironically enough, are also good reading strategies."

All of the principals in this study asserted that they are opposed to teaching to the test—meaning isolated drills on the types of items expected on the test—and that their teachers understand this. The principals said that they would sanction any teacher caught teaching to the test. As a principal in an urban middle school put it: "We don't look at the test and focus on what's on there, and if I ever find a staff member who does, they're in deep trouble." This suggests that the principals in this study provided countervailing pressure against teaching to the test.

Teachers in the majority group asserted that, in any case, drill is not useful in preparing students for the state tests. For example, a special education teacher in an urban high school said:

I find with the math test in particular, the Basic Standards math test, drilling is very little help, because I see the math test is essentially a reading test. My kids have great difficulty comprehending what the question is. Drill to me is helpful when you're drilling on the same thing—for instance, addition facts over and over, or multiplication facts over and over. Well, that test doesn't focus on drill at all, because each question is totally different in the expectation or the skill that it requires. Drill to me means repetition and that isn't one of the things in the math test.

Interviewees in this group felt that testing did not interfere with their ability to provide a balanced educational experience for students. Teachers and administrators felt that they could balance the need to prepare students for the test while maintaining student interest and

encouraging students to be self-motivated learners. A curriculum coordinator in an urban middle school explained:

We have to focus on the strands in both reading and math. We don't teach to it in this building directly although I know that there's a balance, that you have to because it is high stakes. We'd be doing a disservice not to. But I think we try to balance the appropriateness of [instruction]. We're not going to teach the whole year from the strands primarily or solely. [Do you feel that the skills and knowledge in the strands are important?] Yeah, I do. I feel they're important... I think there's a way to teach [the strands] to make it interesting. And I think there's a way to teach to interest a student to become more of a self-learner. I think some programs take it to the extent that, "We're going to design everything around the test and we're going to teach to it," and I think you lose students easily in that type of environment. Whereas if you create an environment where you view the strands as what we're going to focus on, and then create a unit or a project where they're going to use those skills, I think that's more of a creative way to approach it.

A focus on critical thinking in elementary grades. Of the teachers who felt (by a two-to-one margin) that testing had a positive impact on the curriculum, the comments of the elementary level teachers suggest that the more challenging MCA test administered in the elementary grades prompted a greater focus on critical thinking skills. For example, a 4th grade teacher in a semi-rural elementary school explained that the test was well-aligned with her goals of teaching reasoning skills:

Yes, very much so. [Do you feel that the test focuses on important skills and knowledge?] Yes. I think it teaches children to look differently at a question, it teaches children how to answer questions with reasoning. It's not like a multiple guess or multiple choice; they actually have to explain how they arrived at things, and I think that's really good. I think more of it should be done in classrooms.

Elementary level teachers in the majority group believed that the influence of the state test on instruction was positive because it involves challenging multi-step problems, and teachers are required to prepare students to reason through that type of problem. As a 3rd grade teacher in an urban elementary school explained:

I would say the state test is a very challenging test. I think that one thing that my colleagues and I always comment on in third grade is that these are multi-step problems, when I think about the math. That's a very challenging thing for third graders. You begin to do one step problems and you feel pretty good. Two steps, okay; beyond that it's really challenging for them. I'm glad that they have exposure to that; it gives us a high mark to shoot for.

The emphasis of the state test on complex, higher order thinking in the elementary grades presented a challenge for low-achieving students. A 3rd grade teacher in a semi-rural elementary school described the challenge:

My perception is the test is geared more for the higher level thinker. It's kind of stressful for the lower student. One, it depends on if they can read well, because whether it's math or reading, they have to read alone unless they have a special [education] plan. So reading can enter into it. And then if they've got the organizational skills to prepare themselves—how to set up problems or methods. And then the actual thought process involved, I think it's difficult. So I think it's very difficult for the low end student. I'm thinking of the ones not in special ed.

However, I have to say with our practice testing, they don't seem to have anxiety at the time of testing, which is good. They seem to just accept it and they're proud of themselves... I think it's designed more for the higher level thinkers, not for the average third grade student. It gets complex pretty quickly on the test.

Teachers felt that the challenging nature of the state test has raised standards and improved outcomes. A 2nd grade teacher in a semi-rural elementary school described the effect on teaching and learning:

We noticed some years ago that our math scores were generally low, so it's like, "Okay, we've got to do something about this." But we have really upped our standards for all our math. *Whoa*. We have really raised our standards, and so our math scores are coming up.

The focus of the test on critical thinking skills reportedly has had a positive influence on the curriculum in nontested subjects such as science. In both science and math, the curriculum now emphasizes 'real life' situations and higher order thinking. A 5th–6th grade special education teacher in a semi-rural middle school described the changes in the curriculum:

Our curriculum has changed in science and math so that we're teaching the children life skills, real life situations, higher level thinking, math strategies, and science items... It seems to make more sense to them that they're learning something that they can use in the future, rather than to just work on a page of multiplication facts.

Need for segmented tests. Interviewees who were critical of testing suggested a change in the design and implementation of state tests that could address multiple concerns: administering state tests in short segments spread throughout the school year. Segmented testing would allow the number of topics covered to be increased (because the total amount of testing could be increased), reducing the pressure to narrow the curriculum, without forcing students to sit for long periods of time for any single segment. Multiple testing checkpoints throughout the school year could also improve the delivery of diagnostic feedback to teachers, creating opportunities to help students who need assistance.

Discussion

The results of this study suggest that teachers and administrators in the four Minnesota districts felt, by a two-to-one margin, that the overall impact of state-mandated testing in Minnesota is positive, consistent with survey results showing that 85 percent of teachers state-wide support the state's exit exam (Draper, 2000). Within this majority group, teachers and administrators felt that the impact of the state tests on curriculum—focusing attention on basic skills at the 8th grade level and critical thinking skills at the elementary level—was appropriate.

Their reasoning explains their choices about what to teach in their classrooms and, thus, the nature of curriculum narrowing in Minnesota due to high stakes testing. Teachers and administrators reasoned that students need to know basic skills in math and reading as well as critical thinking skills in order to succeed in the world outside of school. In their judgment, the state-mandated tests are well-aligned with key instructional priorities and well-designed to avoid construct under-representation. In their view, Minnesota's state-mandated tests assess students on skills that students should know. Teachers in the majority group felt that it is appropriate to emphasize these skills and to de-emphasize less important outcomes. This suggests that a properly aligned, well-designed testing system can avoid excessive narrowing of the curriculum.

The teachers' claims that drill-oriented instruction is not helpful in preparing students for the state-mandated tests suggest that the tests may be consistent with inquiry-based teaching that emphasizes the learner as an active constructor of knowledge. Whether the Minnesota teachers endorsed inquiry methods was not directly addressed in this study, but the evidence from this study does not support the hypothesis that state-mandated testing in Minnesota is inconsistent with an inquiry-oriented approach.

Three additional features of Minnesota's system contributed to proper alignment and sound design. First, Minnesota's unique testing system encouraged a focus on basic skills as well as critical thinking, maintaining breadth and depth in the curriculum. The high standards MCA exams offset pressure to narrow the curriculum to the basic skills tested on the MBST exit exam. In contrast, states that only use minimum competency tests inadvertently risk encouraging teachers to focus only on the basic skills needed to pass those tests. Second, Minnesota's decision to select a minimum competency exit standard resulted in a high cumulative passing rate—over 99 percent—and low pressure on teachers to teach to the test. Thus, teachers felt that they could prepare students for the exit exam without drilling students on the types of items expected on the exam. Third, administrators reportedly discouraged teaching to the test, providing countervailing pressure against any tendency to drill students on narrow test-focused skills.

The implication of this study is that well-aligned, well-designed tests that emphasize basic skills and critical thinking, plus minimum competency exit standards that reduce pressure on teachers to teach to the test, and administrators that discourage teaching to the test, can avoid much of the curriculum narrowing that reduces the overall quality of education received by students in a high stakes testing environment. To the extent that excessive narrowing of the curriculum occurs in other states, the cause may be state-mandated tests that are not well-designed or well-aligned to the intended curriculum.

The results of this study highlight the importance of aligning the design of state-mandated tests with the intended curriculum, and are consistent with prior research regarding the importance of alignment (see Fuhrman, 2001). In some instances, this requires changing how tests are designed. In other cases, it requires changing state curriculum standards so that they are not excessively oriented toward factual knowledge and do not inadvertently lead to tests that emphasize recall of such knowledge. The results of this study also suggest the importance of *not* linking the compensation or job security of administrators to test scores through bonus systems and performance reviews because this may inadvertently undermine the motivation of administrators to provide the type of countervailing pressure noted in this study, possibly creating an environment where teachers feel intense pressure to raise test scores. It may also be desirable to explore the feasibility of segmenting state-mandated tests into several shorter tests and spreading the administration throughout the school year. Segmented testing would allow the number of topics covered to be increased, reducing the pressure to narrow the curriculum.

The main finding of this study—that it is possible to avoid excessive curriculum narrowing when stakes are high—may seem surprising if high stakes are equated with high pressure. However, this finding is consistent with prior research that makes clear that *stakes*—formal consequences for students and/or schools that are linked to test results—are not synonymous with *pressure*—communications and informal consequences intended to induce staff to increase test scores (Corbett & Wilson, 1991). In a study of 300 school districts in two states, negative impacts on the curriculum occurred when stakes were high and pressure to raise test scores was high (Corbett & Wilson, 1991, p. 126). Positive impacts on teaching and learning occurred when stakes were high but pressure to raise test scores was low (Corbett & Wilson, 1991). High stakes and low pressure can occur when students must pass a test to graduate from high school but almost all students pass easily. This type of test is not likely to generate concern among staff or parents or pressure to change the curriculum.

Under conditions of high stakes and low pressure, schools responded in a positive way (Corbett & Wilson, 1991). Educators in these schools accepted the test as a valid indicator of student learning and used it as the basis for making improvements in instruction while maintaining a balanced curriculum. Schools had adequate time to respond to student weaknesses indicated by test results. Under this condition, teachers implemented strategies for improving learning in a broad, balanced way that led to a rise in test scores without having to teach directly to the test. Teachers were most likely to adopt, adapt or invent more effective instructional practices as the best means of improving student learning (Corbett & Wilson, 1991, p. 116).

Thus, although there are clearly many factors that influence whether state-mandated tests have positive or negative effects, Corbett and Wilson's study suggests that, in high-stakes situations, the degree of pressure that teachers feel to improve student achievement is paramount. The current study suggests that teachers feel less pressure when state-mandated tests are well-designed and properly aligned with curricular and instructional priorities, the exit exam focuses on minimum competencies rather than high standards, and administrators provide countervailing pressure against the urge to teach to the test.

This framework can be extended to incorporate other factors, such as the dropout rate of minorities and ESL students—these factors may be understood as influencing the pressure that is placed on teachers to improve student achievement. A high dropout rate places more pressure on teachers. Similarly, teachers with a higher level of pedagogical content knowledge may feel less pressure to teach to the test, providing a countervailing force against pressures to narrow the curriculum. The Corbett-Wilson theoretical framework suggests why the impact of testing may be positive in one school district and negative in another. It suggests, for example, that despite strong consequences for students, testing *may* have a positive impact on curriculum and instruction *if countervailing factors reduce the pressure on teachers*. Conversely, a lack of countervailing factors may cause teachers to resort to drill and practice.

The role of pressure in causing inadvertent narrowing is significant because many states are rushing to implement high standards exit exams that are likely to generate tremendous pressure on teachers to raise test scores (Gayler et al., 2004). The present study suggests that this move may be premature. Minnesota's system of testing offers an alternative way to raise educational standards without creating excessive pressure on teachers to sacrifice the breadth and depth of the curriculum. State and federal policymakers can increase the probability that high stakes testing will have positive, rather than negative, effects on teaching and learning by drawing upon Minnesota's model and ensuring that tests are properly aligned and pressure on teachers is reduced.

References

- Baker, E. L., & Linn, R. L. (2004). Validity issues for accountability systems. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 47–72). New York: Teachers College Press.
- Baron, J. B., & Wolf, D. P. (1996). *Performance-based student assessment: Challenges and possibilities*. Chicago: University of Chicago.
- Bogdan, R., & Biklen, S. (1992). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn & Bacon.

- Bracey, G. W. (2002). The 12th Bracey report on the condition of public education. *Phi Delta Kappan*, 84(2), 135–150.
- Cimbricz, S. (2002). State-mandated testing and teachers' beliefs and practice. *Education Policy Analysis Archives*, 10(2), retrieved October 26, 2005, from <http://epaa.asu.edu/epaa/v10n2/>.
- Corbett, H. D., & Wilson, B. L. (1991). *Testing, reform, and rebellion*. Norwood, NJ: Ablex.
- Draper, N. (2000, February 3). Poll finds most teachers dislike Profile of Learning; An Education Minnesota survey also found general approval for the state's basic-skills tests. *Minneapolis Star Tribune*, 3B.
- Elmore, R. F. (2004). The problem of stakes in performance-based accountability systems. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 274–296). New York: Teachers College Press.
- 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.
- Fuhrman, S. H. (Ed.). (2001). *From the Capitol to the classroom: Standards-based reform in the states. One hundredth yearbook of the National Society for the Study of Education, Part II*. New York: The University of Chicago Press.
- Gayler, K., Chudowsky, N., Hamilton, M., Kober, N., & Yeager, M. (2004). *State high school exit exams: A maturing reform*. Washington, DC: Center on Education Policy.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine De Gruyter.
- Gordon, S. P., & Reese, M. (1997). High stakes testing: Worth the price? *Journal of School Leadership*, 7, 345–368.
- Grant, S. G. (2000). Teachers and tests: Exploring teachers' perceptions of changes in the New York state testing program. *Education Policy Analysis Archives*, 8(14), retrieved October 26, 2005, from <http://epaa.asu.edu/epaa/v8n14/>.
- Grant, S. G. (2001). An uncertain lever: Exploring the influence of state-level testing in New York State on teaching social studies. *Teachers College Record*, 103(3), 398–426.
- Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives*, 8(41).
- Haney, W., & Madaus, G. (1989). Searching for alternatives to standardized tests. *Phi Delta Kappan*, 70(9), 683–687.

- Hoffman, J., Pennington, J., Assaf, L., & Paris, S. (1999). High stakes testing in reading and its effects on teachers, teaching, and students: Today in Texas, Tomorrow? *The Reading Teacher*, 52, 482–492.
- Jennings, J. F. (2000). *Why national standards and tests? Politics and the quest for better schools*. Thousand Oaks, CA: Sage.
- Koretz, D., Barron, S., Mitchell, K., & Stecher, B. (1996). *Perceived effects of the Kentucky Instructional Results Information System (KIRIS)*. Santa Monica, CA: RAND.
- Koretz, D., Stecher, B., Klein, S., & McCaffrey, D. (1994). The Vermont portfolio assessment program: Findings and implications. *Educational Measurement: Issues and Practice*, 13(3), 5–16.
- Kuhs, T., Porter, A., Floden, R., Freeman, D., Schmidt, W., & Schwille, J. (1985). Differences among teachers in their use of curriculum-embedded tests. *The Elementary School Journal*, 86(2), 141–153.
- Linn, R. L. (2004). Accountability models. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 73–95). New York: Teachers College Press.
- Massell, D. (2001). The theory and practice of using data to build capacity: State and local strategies and their effects. In S. H. Fuhrman (Ed.), *From the Capitol to the classroom: Standards-based reform in the States. One hundredth Yearbook of the National Society for the Study of Education* (pp. 148–169). New York: The University of Chicago Press.
- McNeil, L. M., & Valenzuela, A. (2000). *The harmful impact of the TAAS system of testing in Texas: Beneath the accountability rhetoric*. Houston: Department of Education and Center for Education, Rice University.
- McNeil, L. M. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.
- Mehrens, W. A. (1998). Consequences of assessment: What is the evidence? *Education Policy Analysis Archives*, 6(13), retrieved October 26, 2005, from <http://epaa.asu.edu/epaa/v6n13/>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Pomplun, M. (1997). State assessment and instructional change: A path model analysis. *Applied Measurement in Education*, 10(3), 217–234.
- Porter, A. C., Floden, R. E., Freeman, D. J., Schmidt, W. H., & Schwille, J. P. (1986). *Content determinants* (Research Series No. 179). East Lansing, MI: Michigan State University, Institute for Research on Teaching.
- Porter, A. C., & Smithson, J. L. (2001). Are content standards being implemented in the classroom? A methodology and some tentative answers. In S. H. Fuhrman (Ed.), *From*

the Capitol to the classroom: Standards-based reform in the States. One hundredth yearbook of the National Society for the Study of Education (pp. 60–80). Chicago, IL: The University of Chicago Press.

- Rothman, R. (1995). *Measuring up: Standards, assessment, and school reform*. San Francisco: Jossey-Bass.
- Rothman, R. (2004). Benchmarking and alignment of state standards and assessments. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 96–114). New York: Teachers College Press.
- Rothman, R., Slattery, J. B., Vranek, J. L., & Resnick, L. B. (2002). *Benchmarking and alignment of standards and testing* (CSE Technical Report). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing, University of California, Los Angeles.
- Shim, M. K., Felner, R. D., Shim, E., & Noonan, N. (2001, April 10–14). *Multi-dimensional assessment of classroom instructional practice: A validity study of the Classroom Instructional Practice Scale (CIPS)*. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.
- Smith, M., & O'Day, J. A. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233–267). Philadelphia: Falmer.
- Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8–11.
- Smith, M. L., Noble, A., Heinecke, W., Seck, M., Parish, C., Cabay, M., Junker, S., Haag, S., Tayler, K., Safran, Y., Penley, Y., & Bradshaw, A. (1997). *Reforming schools by reforming assessment: Consequences of the Arizona Student Assessment Program (ASAP): Equity and teacher capacity building [CSE Technical Report 425]*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.
- Smith, M. L., & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice*, 10(4), 7–11.
- Strauss, A. L. (1987). *Qualitative analysis for social scientists*. New York: Cambridge University Press.
- Wolf, S. A., Borko, H., Elliott, R. L., & McIver, M. C. (2000). "That dog won't hunt!": Exemplary school change efforts within the Kentucky reform. *American Educational Research Journal*, 37(2), 349–393.

About the Author

Stuart S. Yeh

University of Minnesota

Email: yehxx008@umn.edu

The author is currently Assistant Professor in the Department of Educational Policy and Administration at the University of Minnesota. His research focuses on improved ways of designing assessment and accountability systems, and he is writing a book that recommends changes in federal, state, and district level testing policies.

Appendix A

Interview Questions

1. Overall, is your opinion about the impact of the state testing program positive or negative? Why?
2. Does the state testing program influence the type of skills and knowledge that you focus on?
 - a. Are these skills and knowledge important?
 - b. Is the test designed so you feel a need to drill students on factual material?
3. How does the state test affect low-functioning students?
4. What are the strengths and weaknesses of the computer-adaptive test, compared to the state-mandated test?
5. Recommendations.
 - a. To guide instruction, would you favor re-designing the state test to provide more detailed diagnostic information about student strengths and weaknesses?
 - b. Would you favor consolidation of all of the state and district tests into one test?
 - c. Suppose this was done and suppose that you could receive scores within 7 days. Would you favor, at every grade level, giving the test before the school year, as a pre-test to diagnose students, and at the end of the school year, as a post-test to determine how much students improved during the year?
 - d. Would you favor using the test to determine which students should receive afterschool tutoring?
 - e. Should students who fail the end-of-the-year test be required to attend summer school as a condition of promotion to the next grade?
 - f. Would you favor the use of a computer-adaptive test that adapts the difficulty of the questions to the student's level?
 - g. Would you favor the use of a test that provided the option of performance assessments for low-functioning students, voc ed students, bilingual students, and special ed students, even if it meant that regular ed teachers might have to share the responsibility of administering and scoring the performance assessments? [Perhaps 2 days of work per teacher].
 - h. Would you favor state tests that focus on critical thinking rather than recall of facts?
 - i. Until these changes are made, would you favor delaying the requirement that students pass the state test in order to graduate from high school?
 - j. Are there other suggestions that you would make regarding the design or implementation of the state testing program?

EDUCATION POLICY ANALYSIS ARCHIVES <http://epaa.asu.edu>

Editor: Sherman Dorn, University of South Florida

Production Assistant: Chris Murrell, Arizona State University

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Sherman Dorn, epaa-editor@shermamdorn.com.

Editorial Board

Michael W. Apple

University of Wisconsin

Greg Camilli

Rutgers University

Linda Darling-Hammond

Stanford University

Gustavo E. Fischman

Arizona State University

Gene V Glass

Arizona State University

Aimee Howley

Ohio University

William Hunter

University of Ontario Institute of Technology

Daniel Kallós

Umeå University

Thomas Mauhs-Pugh

Green Mountain College

Heinrich Mintrop

University of California, Berkeley

Anthony G. Rud Jr.

Purdue University

Terrence G. Wiley

Arizona State University

David C. Berliner

Arizona State University

Casey Cobb

University of Connecticut

Mark E. Fetler

California Commission on Teacher Credentialing

Richard Garlikov

Birmingham, Alabama

Thomas F. Green

Syracuse University

Craig B. Howley

Appalachia Educational Laboratory

Patricia Fey Jarvis

Seattle, Washington

Benjamin Levin

University of Manitoba

Les McLean

University of Toronto

Michele Moses

Arizona State University

Michael Scriven

Western Michigan University

John Willinsky

University of British Columbia

EDUCATION POLICY ANALYSIS ARCHIVES
English-language Graduate-Student Editorial Board

Noga Admon
New York University

Jessica Allen
University of Colorado

Cheryl Aman
University of British Columbia

Anne Black
University of Connecticut

Marisa Cannata
Michigan State University

Chad d'Entremont
Teachers College Columbia University

Carol Da Silva
Harvard University

Tara Donahue
Michigan State University

Camille Farrington
University of Illinois Chicago

Chris Frey
Indiana University

Amy Garrett Dikkers
University of Minnesota

Misty Ginicola
Yale University

Jake Gross
Indiana University

Hee Kyung Hong
Loyola University Chicago

Jennifer Lloyd
University of British Columbia

Heather Lord
Yale University

Shereeza Mohammed
Florida Atlantic University

Ben Superfine
University of Michigan

John Weathers
University of Pennsylvania

Kyo Yamashiro
University of California Los Angeles

Archivos Analíticos de Políticas Educativas

Associate Editors

Gustavo E. Fischman & Pablo Gentili

Arizona State University & Universidade do Estado do Rio de Janeiro

Founding Associate Editor for Spanish Language (1998—2003)

Roberto Rodríguez Gómez

Editorial Board

Hugo Aboites

Universidad Autónoma
Metropolitana-Xochimilco

Dalila Andrade de Oliveira

Universidade Federal de Minas
Gerais, Belo Horizonte, Brasil

Alejandro Canales

Universidad Nacional Autónoma
de México

Erwin Epstein

Loyola University, Chicago,
Illinois

Rollin Kent

Universidad Autónoma de
Puebla. Puebla, México

Daniel C. Levy

University at Albany, SUNY,
Albany, New York

María Loreto Egaña

Programa Interdisciplinario de
Investigación en Educación

Grover Pango

Foro Latinoamericano de
Políticas Educativas, Perú

Angel Ignacio Pérez Gómez

Universidad de Málaga

Diana Rhoten

Social Science Research Council,
New York, New York

Susan Street

Centro de Investigaciones y
Estudios Superiores en
Antropología Social Occidente,
Guadalajara, México

Antonio Teodoro

Universidade Lusófona Lisboa,

Adrián Acosta

Universidad de Guadalajara
México

Alejandra Birgin

Ministerio de Educación,
Argentina

Ursula Casanova

Arizona State University,
Tempe, Arizona

Mariano Fernández

Enguita Universidad de
Salamanca. España

Walter Kohan

Universidade Estadual do Rio
de Janeiro, Brasil

Nilma Limo Gomes

Universidade Federal de
Minas Gerais, Belo Horizonte

Mariano Narodowski

Universidad Torcuato Di
Tella, Argentina

Vanilda Paiva

Universidade Estadual Do
Rio De Janeiro, Brasil

Mónica Pini

Universidad Nacional de San
Martín, Argentina

José Gimeno Sacristán

Universidad de Valencia,
España

Nelly P. Stromquist

University of Southern
California, Los Angeles,
California

Carlos A. Torres

UCLA

Claudio Almonacid Avila

Universidad Metropolitana de
Ciencias de la Educación, Chile

Teresa Bracho

Centro de Investigación y
Docencia Económica-CIDE

Sigfredo Chiroque

Instituto de Pedagogía Popular,
Perú

Gaudêncio Frigotto

Universidade Estadual do Rio
de Janeiro, Brasil

Roberto Leher

Universidade Estadual do Rio
de Janeiro, Brasil

Pia Lindquist Wong

California State University,
Sacramento, California

Iolanda de Oliveira

Universidade Federal
Fluminense, Brasil

Miguel Pereira

Catedrático Universidad de
Granada, España

Romualdo Portella do

Oliveira

Universidade de São Paulo

Daniel Schugurensky

Ontario Institute for Studies in
Education, Canada

Daniel Suarez

Laboratorio de Políticas
Publicas-Universidad de
Buenos Aires, Argentina

Jurjo Torres Santomé

Universidad de la Coruña,
España