J Sci Teacher Educ (2012) 23:233–257
DOI 10.1007/s10972-012-9266-3

A National Survey of Middle and High School Science Teachers' Responses to Standardized Testing: Is Science Being Devalued in Schools?

Mehmet Aydeniz · Sherry A. Southerland

Published online: 24 February 2012

Abstract This study explored American high school and middle school science teachers' attitudes toward the use of standardized testing for accountability purposes, their justification for the attitudes they hold and the impact of standardized testing on their instructional and assessment practices. A total of 161 science teachers participated in the study. Analyses were based on teachers' responses to a questionnaire including nine-item likert-scale questions and two-item open-ended questions. The analyses revealed that science teachers have mixed reactions to the administration of standardized tests and its use for accountability purposes. The findings also reveal that standardized testing has a significant influence on science teachers' instructional and assessment practices in ways that are counter to the learning goals promoted by science education reformists. Our discussion focuses on the implicit and explicit influences of the NCLB Act on science curriculum, teaching and assessment, and how the NCLB driven policies undermine the goals of science education reform.

Keywords Science · Standardized testing · Science teachers

Introduction

Assessment can play a central role in efforts to bring about improvements in the educational system, curriculum, quality of instruction and student learning (Aydeniz 2007; Brickhouse 2006; National Academy of Sciences [NAS] 2006; National

M. Aydeniz (✉)
Department of Theory and Practice in Teacher Education, College of Education, Health and Human Sciences, The University of Tennessee, Knoxville, USA
e-mail: maydeniz@utk.edu

S. A. Southerland
FSU-Teach, School of Teacher Education, Florida State University, Tallahassee, FL, USA

Research Council [NRC] 2001, 2005). Assessment can support students' learning and improve the quality of instruction when used formatively (Abell and Volkmann 2006; Aydeniz 2007; Bell and Cowie 2001; Black and William 1998; Brookhart 2006; Duschl and Gitomer 1997; Klassen 2006; McMillan 2001; Shepard 2000). When used summatively, it can help monitor the effectiveness of a particular curriculum (NAS 2006; NRC 2005), help evaluate the quality and effectiveness of instruction (Bell and Cowie 2001; NAS 2006; NRC 2001; Shepard 2000), and enhance the efficiency of the school system (Brickhouse 2006; Linn 2000, 2003; NAS 2006; Popkewitz 2000).

Although assessment can serve multiple purposes, school systems in the United States are increasingly emphasizing the summative function of assessment. This emphasis began with the signing of the No Child Left Behind [NCLB] legislation Act into law (U.S. Department of Education 2002) (Brickhouse 2006; Darling-Hammond 2004; Davis et al. 2007; Stiggins 2004). Because the NCLB Act holds school systems accountable for improved test scores, teachers are pressured to focus on using traditional instructional and assessment practices that have been effective in improving students' achievement scores on statewide-standardized tests (Brickhouse 2006; Darling-Hammond and Adamson 2010). This emphasis on increasing students' test scores undermines reform efforts that encourage science teachers to use assessments that are aligned with the instructional goals promoted by the science education reform documents such as the National Science Education Standards [NSES] (NRC 1996) (Aydeniz 2007; Madden 2008).

The NCLB Act calls for greater accountability to improve students' achievement scores on statewide-administered summative tests for both teachers and school systems (Brickhouse 2006). In an effort to achieve the goal of improving students' test scores, the NCLB legislation mandates each school system to write a set of Adequate Yearly Progress (AYP) objectives in each core subject area (mathematics, language arts, and science) aligned with the state standards. To evaluate how states are meeting the standards, NCLB requires each state to measure student learning annually and requires schools' efforts to focus on supporting all students to meet the Adequate Yearly Progress (AYP) objectives. Not only are all students in a state held to the same rigorous academic standards (e.g., in mathematics and language arts) within this accountability system but also schools that fail to meet their Adequate Yearly Progress (AYP) objectives for student learning are threatened with sanctions (e.g., lessening of funds), including the take over of the schools by the states (Abrams et al. 2008; Darling-Hammond 2004; Madden 2008).

The proponents of NCLB aim to achieve accountability through the administration of standardized tests across core subject areas (e.g., mathematics, language arts, and science). With the increasingly common use of standardized testing as an accountability measure for school systems and teachers, there is a growing concern among educators about its influence on curriculum, teaching, instructional time, and student learning (Abrams et al. 2003; Brickhouse 2006; Darling-Hammond 2004; DeBoer 2002; Linn 2003; Madden 2008; Shaver et al. 2007). Thus, educators, parents, and politicians are engaged in on-going discussions about the social and academic consequences of standardized testing in the U.S. schools (Brickhouse 2006; Popham 2006).

Proponents of standardized testing argue that test scores are valid and reliable indicators of student learning in a particular academic subject and that testing is an effective system in ensuring the attainment of minimum academic competencies by all students (Greene and Winters 2003). Contrary to this view, opponents argue that, despite its potential benefits for bringing about improvements at the system level, basing high-stakes decisions (i.e., restructuring of schools, rehiring of teachers, and decisions related to student graduation) on the results of one single test does not necessarily ensure the quality of science education delivered in nation's science classrooms (Aydeniz 2007; Brickhouse 2006; Madden 2008) as well as address equity issues in our schools (Darling-Hammond 2004, Popham 2006). This critique continues as researchers emphasize the huge role of test scores on science teachers' decision-making, the influence they have on the breadth and depth of curriculum addressed in the nation's science classrooms, the nature of classroom instruction and assessment (Aydeniz 2007; Brickhouse 2006; Darling-Hammond 2004; Madden 2008; Popham 2006). These researchers find this approach to educational improvement problematic as it encourages science teachers to use instructional and assessment practices that are not effective in promoting the learning outcomes advocated by the science education reform documents such as National Science Education Standards [NSES] (Abell and Volkmann 2006; Brickhouse 2006; DeBoer 2002; Madden 2008; Pringle and Carrier 2005). These scholars maintain that using standardized test scores as a measure of teacher effectiveness pressures teachers to reduce the content of their curriculum to students' acquisition of only the knowledge and skills necessary for passing the test (Aydeniz 2007; Madden 2008). Because these assessments at best serve as only weak proxies of the true knowledge and skills required in the workplace and real life (Brickhouse 2006; Darling-Hammond and Adamson 2010), use of these assessments as the sole measure on which to make high-stakes decisions undermines the achievement of the ambitious goals of science education reform outlined in reform documents such as the NSES (NRC 1996). More specifically, because of the intensive pressure of increasing test scores for the vast bulk of their students, teachers do not (as though they have the time and resources necessary to) differentiate instruction to address the learning needs of students who need additional assistance (i.e., underachievers) or those who need further challenge (i.e., gifted students). Rather, they cater instruction to the learning needs of average achieving students—who have the greatest likelihood of making gains on these assessments. These decisions made by teachers fail to address the learning needs of the students who have low motivation for learning science or those who need the most help to understand science (Aydeniz 2007; Darling-Hammond 2004). Such teaching practices help maintain status quo rather than helping all students to learn science in such ways that the national reform documents such as NSES (NRC 1996) recommend (Abrams et al. 2008). Moreover, these scholars argue that not only do accountability pressures motivate teachers to focus their instructional planning on test content (Whitford and Jones 2000), but also it encourages them to devote more instructional time to preparing their students for test-taking techniques (Aydeniz 2007; Grant 2000; Madden 2008; Mehrens and Kaminski 1989; Smith and Rottenberg 1991). This happens at the expense of teaching science for students' acquisition of scientific inquiry skills (Abrams et al.

2008; Aydeniz 2007; Madden 2008) and conceptual understanding of key scientific facts in a meaningful way (Aydeniz 2007). Past research has shown that traditional instruction which includes teachers lecturing and teaching to the test fail to engage the students who have limited interest in academic learning and who also happen to be the student group that determines whether a school meets its AYP objectives or not. Thus, standardized test-based accountability measures: 1) do not encourage the type of teaching that holds potential to address the learning needs of low-performing students and 2) fail to promote students' acquisition of learning outcomes advocated by the reform documents and current science education literature.

Zancanella (1992) and Cimbricz (2002) argue that the assumption that state-mandated testing leads to better teaching and triggers significant growth in students' learning is more an expression of hope than reality. However, given current legislation, many states continue to administer standardized testing for accountability purposes (NAS 2006). Therefore, we argue that science educators should work diligently to understand the influence of standardized testing on science teacher practice and student learning. Such understandings are crucial for our work to better support the implementation of science education reform practices in our nation's science classroom. Thus, the purpose of this study is to contribute to the on-going discussion on science teachers' attitudes toward standardized testing, the justifications they have about the attitudes they hold toward standardized testing and the impact of standardized testing on science teachers' instructional and assessment practices.

Accountability and Equity

Achieving equity in educational opportunity is one of the primary goals of the NCLB Act (U.S. Department of Education 2002). In an effort to ensure that each student has equal access to rigorous academic content, the NCLB legislation requires that each state report AYP disaggregate data for the minority students, such as African–American students, English as Second Language Learners (ESL), students with disabilities and those that are coming from socioeconomically disadvantaged populations (Berliner 2005; Darling-Hammond 2004; Kim and Sunderman 2005). It is hoped that mandating this disaggregated reporting of AYP will encourage school systems to pay increased attention to the academic progress of these historically disadvantaged aforementioned populations and pressure school systems to allocate sufficient resources to ensure their achievement on the state-mandated standardized tests (Darling-Hammond 2004; Kim and Sunderman 2005).

However, the ways in which school systems try to achieve accountability is in conflict with the purposes of achieving equity in American public schools (Kim and Sunderman 2005; Abrams et al. 2008). Although equity is being voiced in the strongest language in NCLB legislation, equity is viewed only through the lens of increased test scores (Kim and Sunderman 2005). Science educators find this approach to ensuring equity in education problematic. They argue that holding schools accountable only for increased test scores encourages teachers to use

traditional teaching methods (i.e., lecturing) that are more likely to increase student performance when the tests measure students' declarative knowledge (Abrams et al. 2008; Madden 2008). Science educators working toward the reform of teaching practice suggest that endorsing policies that will encourage science teachers to teach meaningful, inquiry-based instructional activities is more likely to ensure equity—that is holding the same goals for different students mindful that at each student has different needs and strengths so that instruction needs to be crafted mindful of student differences (Lynch et al. 2005). Science educators agree that such equitable science instruction—instruction that varies in light of the needs of the students served—is more likely to allow all students to construct meaningful understandings of essential scientific concepts and processes.

Evidence of ineffectiveness of the NCLB policies in ensuring equitable instruction is reported in current literature. Kim and Sunderman (2005) examined how AYP policies under the NCLB Act affect the schools that primarily serve students that come from high poverty and minority families. Using data from six states (Arizona, California, Georgia, Illinois, New York, and Virginia) Kim and Sunderman (2005) found that the requirements of the NCLB do the greatest harm to the schools that host students coming from high poverty and minority backgrounds. They argue that these schools have high number of teachers who are not highly qualified in their teaching subject area. Kim and Suderman (2005) describe that students with limited resources and that have teachers with limited qualifications are being forced to compete with those that have sufficient resources and well-qualified teachers on the same test. Thus, they argue that testing only documents the failure of students coming from schools that have limited resources and less qualified teachers and the success of those that are coming from schools that have more resources and better prepared teachers (Darling-Hammond 2004; Popham 2006). Moreover, because many states use the test scores as a means to assume greater control over the operations of school systems such as the authority to take over the schools for poor test scores (Aydeniz 2007), the schools serving primarily to the minority and low SES students are often pressured to change the way they teach. For instance, Kim and Sunderman's (2005) analysis indicate that Black and Latino students represent 87% of all students attending schools in New York and California, the two states with highest minority enrollment, that are in the 3rd or 4th year of improvement plans. These are the schools that are subject to corrective action or school restructuring.

These findings suggest that achieving accountability through the requirements of NCLB poses significant challenges to achieving equity and excellence in science education. That is the case not only because current tests serve as only weak proxies of the kind of knowledge and skills needed for students to become scientifically literate (Aydeniz 2007; Brickhouse 2006) and but also because states have been given the flexibility whether to include students' science achievement scores in their calculation of their AYP objectives or not. This flexibility can potentially encourage some school districts and states to provide limited support and resources for improving the quality of science instruction- as it is hard to increase students' test scores especially of those lacking motivation, instead provide more support and resources for improving the quality of instruction in subjects such as mathematics

and reading as it is easier to increase students' test scores in these two subject areas through practice (Aydeniz 2007; Madden 2008). Similarly, because increasing test scores becomes the central purpose of school systems, teachers may not feel pressured to teach science for the purposes of promoting the learning goals advocated by the science education reform documents such as the NSES (NRC 1996). Instead, they may focus on teaching the science content tested on standardized tests, which at best only represents a weak proxy of students' basic scientific knowledge (Brickhouse 2006) and in certain cases allocate instructional time to improve students' reading comprehension and mathematical skills in science classrooms—as most schools count test scores in these two areas in their AYP calculations. Finally, the NCLB policies may encourage science teachers to use instructional activities that are most successful for mainstream students. Given the wide scope of many states' science standards, and the limited time allotted to meet these standards, such policies may discourage teachers from differentiating instruction to meet the learning needs of all students. Thus, the NCLB-related policies may further contribute to the achievement gap in science between mainstream students and historically underrepresented student populations such as African Americans and those coming from low-socioeconomic backgrounds. Therefore, it is important to understand how science teachers view the role of standardized testing in bringing about improvements to science education and the impact they feel on their practice if we are to inform policy related to current and future accountability measures.

Accountability and Science Teacher Practice

The administration of standardized tests has significant implications for science teachers who work in public schools. For instance, entertaining the goals of science education reform such as teaching science through inquiry (NRC 1996), accommodating the learning needs of all students, and ensuring all students' success on standardized testing presents significant challenges for science teachers in the face of NCLB legislation (Abrams et al. 2008; Aydeniz 2007; Madden 2008). In the wake of the first administration of science assessments in the 2007–2008 school year for grades 3–5, grades 6–9, and grades 10–12, science educators are beginning to recognize how the goals of science education reform may be undermined (Aydeniz 2007; Brickhouse 2006; Pringle and Carrier 2005). Because many states do not include students' science scores in their calculations of the AYP reports, increasing students' achievement in science may not be a priority for the school administrators. As a result, administrators may not hold teachers accountable for instructional practices that hold potential to promote students' conceptual understanding in science or their acquisition of scientific inquiry skills (Aydeniz 2007; Madden 2008). It is our argument that this special treatment of science in the NCLB Act has significant implications for the place of science in school curriculum, the teaching of science and assessment of students' learning in science. In such a context, only a limited number of teachers who understand and commit to the goals of science education reform may enact curriculum and instruction in ways that are consistent with the goals of science education reform.

As science is just becoming part of some states' AYP reports, science teachers' reaction to standardized testing requires special attention. Given this already difficult backdrop, science educators need to develop a more useful understanding of how science teachers respond to the pressures represented in standardized high-stakes assessments. As a community, we need to monitor science teachers' instructional and assessment practices to understand how the administration of standardized testing for accountability purposes influences science teachers' assessment and teaching practices and as a result their students' learning. Such scrutiny of science teachers' instructional and assessment practices can provide invaluable understandings and knowledge that we can use to support science teachers to implement the goals of science education documents (AAAS 1993; Duschl et al. 2007; NRC 1996, 2000) in science classrooms. This study was designed to shed some light on science teachers' attitudes toward standardized testing for accountability, the justification they have for the attitudes they hold, and the influence of standardized testing on their instructional and assessment practices. More specifically, the purposes of this study were as follows: (1) to explore secondary school science teachers' attitudes toward standardized testing, (2) to explore their justifications for the attitudes that they hold toward standardized testing, and (3) to report on the teachers perceptions of changes in secondary school science teachers' instructional and assessments practices due to the administration of state-mandated standardized tests.

Methods

We used a descriptive, quantitative approach in this study. Data were collected through the use of a survey administered to a wide variety of 161 secondary school science teachers in 14 states across the United States. Data were analyzed through descriptive statistics.

Instrumentation/Data Collection

This study is part of a larger study that looked at various dimensions of science teachers' understandings and practices of assessment through a forced-response, likert-scale survey and a set of open-ended questions. Only nine forced-response and two open-ended questions in the survey focused on science teachers' understandings and practices of assessment in the context of standardized testing. These questions were designed to elicit secondary school science teachers' attitudes toward standardized testing and explore the changes in their instructional and assessments practices due to the implementation of standardized testing. One of the open-ended questions asked the participants, "Please elaborate on the advantages and disadvantages of the administration of the [name of state-standardized test]? and the other one asked them, "Please elaborate on the changes you have made a) in your instructional practices as a result of the administration of [name of state-standardized test] b) in your assessment practices as a result of the administration of [name of state-standardized test]? These questions were designed based on the

analysis of a qualitative set of data collected from secondary science teachers in a previous study (Aydeniz 2007) and review of literature on science teachers' assessment beliefs and practices.

The survey was then placed in an online platform through a secure server and the link to the online survey was sent out to the 50 state science coordinators through an email. The state science coordinators were asked to distribute the link to the online survey to the middle and secondary school science teachers in their states. The state science coordinators were asked to distribute the link to the online survey to the middle and high school science teachers in their states and to encourage them to complete the survey.

Participants

A total of 161 teachers from 14 different states in the United States completed the survey. These states include: Georgia, Tennessee, Florida, Louisiana, Kentucky, Missouri, Massachusetts, Maine, Arizona, New Hampshire, Wisconsin, North Carolina, South Carolina, and Texas. Participants had diverse characteristics and professional backgrounds. Of the 161 participants who completed the survey, 10 (6%) held a Ph.D. degree, 18 (11%) Ed.s degree, 77 (48%) a Masters degree, and 56 (35%) Bachelor's degree. In terms of grade levels, 76 (47%) of the participants taught at a middle school and 85 (53%) taught at a high school. In terms of teaching experience, 13 (8%) teachers had 1–2, 37 (23%) 3–7, 54 (34%) 8–15, 39 (24%) 16–25, 18 (11%) more than 25 years of teaching experience. As seen in these statistics, most teachers had at least 3 years of teaching experience which is considered to be a critical point for teacher attrition (Ingersoll and Perda 2010; Keigher 2010). In terms of their respective schools' achievement level on standardized testing, 72 (45%) came from schools that performed above their state's mean, 62 (39%) came from schools that performed at their state's mean, and 27 (17%) came from schools that performed below their state's mean. In terms of the socioeconomic state of their students, 89 (55%) of the participants taught students who were mostly in the low-socioeconomic bracket, 11 (7%) mostly in the high-socioeconomic bracket, and 61 (38%) mostly in the middle socioeconomic bracket level students. In terms of students' achievement level, 15 (9%) taught mostly low-achieving students, 19 (12%) mostly high-achieving students, and 127 (79%) taught students of mixed abilities. In terms of geographic location, 25 (16%) teachers taught in an urban, 54 (34%) in a suburban, and 82 (%51) in a rural school setting. These characteristics are summarized in Table 1.

Limitations of the Study

Although we controlled for certain variables in the design of our study, the voluntary nature of the participants and the characteristics of the participants places certain limitations on our findings. First, the manner in which standardized tests are designed and used for accountability purposes vary from state to state; therefore, the quality and nature of standardized tests may have had an influence on the interpretations of science teachers who participated in this study. Second, because of

Table 1 Characteristics of participants

Degree	PhD	Eds	MSc	Bsc	
# of teachers	10	18	77	56	
Years of experience	1–2	3–7	8–15	16–25	>25
# of teachers	13	37	54	39	18
Geographic location of schools	Urban	Suburban	Rural		
# of teachers	25	54	82		
Grade level	Middle school	High school			
# of teachers	76	85			
School's achievement level on AYP	Above state average	At state average	Below state average		
# of teachers	72	62	27		
Students' achievement level	High achieving	Mixed abilities	Low-achieving		
# of teachers	19	127	15		
SES of students	High SES	Middle SES	Low SES		
# of teachers	11	61	89		

the voluntary nature of the participants, each state is not equally represented in our sample. However, given the wide range of our sample (i.e., participants from 14 different states), we argue that our findings should provide valid insights into science teachers' attitudes toward standardized testing and reflect the type of changes that standardized testing might have triggered in science teachers' instructional and assessment practices. We caution our readers and encourage them to keep these limitations in mind as they interpret the findings of this study.

Data Analysis

We used two forms of analyses: quantitative and qualitative. Quantitative data analysis include descriptive analysis of science teachers' responses to a nine-item survey that measured science teachers' attitudes toward standardized testing. We used the following procedures for quantitative analysis. First, we added the number of responses that corresponded with the "agree" and the "strongly agree" categories and calculated the percentage of positive responses for each question. Second, we added the number of responses that corresponded with the "disagree" and the "strongly disagree" categories and calculated the percentage of negative responses for each question. Then, we calculated the number of answers that corresponded to the "neutral" statement for each question. Then, we compared the percentage of positive statements with the percentage of negative statements to understand the participant teachers' attitudes toward various dimensions of standardized testing and its influence on their instructional and assessment practices.

In terms of qualitative data analyses, we analyzed science teachers' responses to the two open-ended questions that measured science teachers' perceived advantages and disadvantages of administering standardized tests in science, and the particular

Table 2 Science teachers' attitudes toward standardized testing and its impact on instruction and classroom assessment

Questionnaire statements	*n*	SD	D	N	A	SA	Mean
1. I believe the administration of the {#test} will improve student learning	161	37	60	31	28	5	2.40
2. I make changes in my assessment because I believe the {#test} related policies will increase student learning	161	22	65	40	29	5	2.56
3. I make changes in my assessment because I believe it is not fair to my students to take the test without necessary preparation	161	3	15	22	86	35	3.84
4. I make changes in my assessment because I do not want to lose my job	161	31	44	39	39	8	2.71
5. The {#test} dictates how I teach and assess my students' learning	161	21	30	29	54	27	3.22
6. I feel encouraged by my school administration to organize my lessons around the {#test} objectives	161	9	14	30	62	46	3.79
7. Performance differences in student achievement on {#test} reflect differences in the characteristics of students rather than teacher effectiveness	161	2	22	42	67	28	3.59
8. The {#test} forces me to teach in ways that contradict my own beliefs about effective science instruction	161	20	43	39	32	27	3.02
9. The {#test} scores are not accurate measure of what my English as a Second Language Learners (ESL, ESOL, ELL) students know and can do in science	161	–	13	59	51	38	3.70

SD Strongly disagree, *D* diagree, *N* neutral, *A* agree, *SA* strongly agree

changes that they made in their instructional and assessment practices due to the administration of standardized tests. After we printed participant teachers' responses to the two open-ended questions, we went through their responses line by line to identify themes that are central to our research questions. Our analysis led to the emergence of the following themes: views related to student learning, views related to instructional practice, and views related to assessment practices. These themes were color coded throughout the entire file. Then, we counted the number of positive and negative statements separately for each theme (i.e., changes in science teachers' assessment practices). We report the overall statistics in Table 2, and more specific statistics in Tables 3, 4, 5, 6, 7 and 8.

Findings

Two sets of data analyses are reported in the findings section: quantitative and qualitative. Three of the nine questions (i.e., Q1, Q7, and Q9) focused on science teachers' perceptions of the validity of standardized tests administered in their relative states. Three of the nine questions on the survey focused on the impact of standardized testing on science teachers' instructional practices (i.e., Q5, Q6, and Q8) and three questions focused on the impact of standardized testing on science teachers' assessment practices (i.e., Q2, Q3, and Q4). Table 2 provides descriptive

Table 3 Aspects of science teachers' positive statements related to the influence of testing on their teaching

Positive statement	N = 55	Representative statement
1. Administration of the standardized tests achieves the uniformity of the content being taught across schools and districts	48	It only influences the fact that I must teach the basic concepts of my course to the best of my ability. I also understand that I must try to cover as much material as the students can reasonably learn so that they are exposed to as many concepts as possible
2. Administration of the standardized tests holds all teachers accountable for teaching rigorous content	7	Requires extended coverage of material. Encourages me to teach a lot more concepts than I used to, which I think is good. It makes teachers to teach actual content instead of fluff

Table 4 Aspects of science teachers' negative statements related to the influence of testing on their teaching

Negative statement	N = 164	Representative statement
1. Administration of standardized tests leads teachers to water down the content	53	I am unable to go in-depth in most disciplines because I have to rush through to cover all of the material. Skimming through is the wrong approach to teaching science I no longer teach parts of my subject that are fun and interesting
2. Such tests do not inform subsequent instruction	13	It does not allow for re-teaching
3. Administration of such tests forces teachers to teach to the test	98	Unfortunately, I often am forced to" teach to the test" rather than making sure that students are truly learning the "big idea" concepts that would serve them better

Table 5 Aspects of science teachers' positive statements related to student learning

Positive statement	N = 21	Representative statement
1. Allows for the evaluation of student thinking	4	Allows for the evaluation of student thinking
2. Achieves the uniformity of curriculum taught in the classroom	17	Tests a standardized base knowledge so students can transfer from school to school and colleges know what is taught in biology

information on science teachers' responses to these survey questions. These responses served as the bases of our interpretation.

The {#test} refers to the test that was used in the state that the teacher is currently employed. The participants were prompted to state the name of the standardized test that was administered statewide in their relative states for accountability purposes. Once the participants entered the name of the test that name replaced {#test} for the remaining questions.

Table 6 Aspects of science teachers' negative statements related to student learning

Negative statement	N = 238	Representative statement
1. Standardized tests do not measure what students learn in science courses	141	Topics relating to earth science are not on the test
		I teach physics but I am not sure how much of what I teach is on the test
		I think the test only measures what they learned in 8th grade. It is almost like a joke
		One question cannot gauge whether a student knows or understands motion
2. Students do not take the test seriously	89	Students are very smart. They know the test does not count so they do not try their best
3. These types of test discriminate against special needs students such as ESL	8	How can you understand the question if it is not written in your home language?

Table 7 Tally of Science Teachers' Comments Related to Changes in Their Assessment Practices

Type of change	Number of statements	Percentage
1. Spending time on test-taking strategies	N = 145	90
2. Inclusion of more multiple-choice type questions	N = 103	64
3. Effects the content of my test	N = 68	42
4. Affects the format of my test	N = 119	74
5. No effect	N = 16	10

Table 8 Changes in Science Teachers' Assessment Practices

Type of statement	Representative statements
Negative statements	Assessments are geared toward the test instead of students' interest
	Does not allow for variation for students' learning styles
	We have to have [name of the test] Dailies
	At the end of year, I switch from essay type questions to multiple-choice questions so they are ready for the same form of assessments. I spend 2 weeks reviewing the materials to prepare them for the test
	I use trickier questions
Positive statements	I include more open-ended questions to encourage writing
	It influences the type of questions that I ask in my labs. I do more compare contrast, critical thinking type of assessments
	I do formative assessment quizzes every week to determine student strength and weaknesses. The state gives us a template and asks us to write our own assessment questions. Too much work

Our findings show that from science teachers' perspectives, statewide-standardized testing practices used for the purposes of accountability influenced the quality of curriculum implemented, the instructional practices used by science teachers and

the content and methods of assessments used by science teachers. The findings suggest that teachers perceived that the standardized assessments linked to the NCLB legislation encouraged them to use effective instructional and assessment methods that are promising in terms of improving students' achievement scores on the standardized tests that they believed did not measure the type of learning goals advocated by the science education reform documents. We qualify this evaluation in the argument that we present in the discussion that follows.

How Do Science Teachers Perceive the Impact of Standardized Testing on Student Learning?

The analyses revealed that only 20.5% ($n = 33$) of the participants agreed that the administration of standardized testing would improve student learning, 19% ($n = 31$) held a neutral view on the impact of standardized testing on student learning, and 60.5% ($n = 97$) of the participants did not think that the administration of standardized testing would improve student learning in science. Those who believed that the administration of standardized testing would improve student learning suggested that this was the case because such policy held teachers accountable for strictly adhering to the state-mandated curriculum standards in an effective manner as evidenced in their responses to the open-ended questions. Those who believed that the administration of standardized testing for accountability purposes would not improve student learning suggested that this was the case because the pressure of improving students' test scores encouraged them to reduce the content of their enacted curriculum to students' acquisition of factual knowledge as evidenced in their responses to the open-ended questions.

The teachers who supported the administration of the test supported students' acquisition of canonical knowledge; in contrast, those who did not reinforce the administration of the standardized tests supported students' acquisition of inquiry skills. For instance, one teacher who held a positive attitude toward the administration of the test said, "it [the test] holds teachers accountable to teach minimum standards that the students must know to graduate." Another teacher who also held a positive view toward the administration of the test said, "I have also stressed science content vocabulary to my students. I have them use the vocabulary in their spoken and written answers." The teachers who held negative attitudes toward the administration of standardized testing for accountability purposes complained that testing discourages them from emphasizing students' acquisition of higher-order thinking skills. For instance, one such teacher said, "I feel that I am restricted and cannot afford to spend the extra time on higher-level learning and research for students because of time constraints from [name of the test] test dates."

What Do Test Scores Represent?

One of the controversies surrounding the administration of standardized testing for accountability purposes is the meaning that the stakeholders attach to the test scores. While some relate the test scores to teachers' effectiveness others relate test scores to students' level of motivation. Teachers who participated in this study differed in

their views about the meanings of students' test scores. While 15% ($n = 24$) of the participants stated that "students' test scores on standardized tests is an indicative of teacher effectiveness", 59% ($n = 95$) of the participants did not believe there was a positive correlation between students' performance on standardized tests and teacher effectiveness. The rest of the participants ($n = 42$ or 26%) held a neutral view on the correlation between teacher effectiveness and students' test scores. A significant number of the participants, 55% ($n = 89$), did not think that the standardized test scores reflected English as a Second Language (ESL) students' learning in science. The majority of the teachers ($n = 141$; 88%) agreed that the test administered in their relative states did not measure the content that they taught in their courses. Moreover, a significant number of high school science teachers ($n = 72$) indicated in their responses to the open-ended questions that the standardized tests only measured their students' understanding of basic science concepts that are covered in middle school and not the science content that they taught in their courses.

These findings suggest that policymakers and school administrators are making decisions that impact teachers, students, and parents based on a set of data that are problematic at best from the viewpoint of teachers. If the judgements about the quality of instruction, teacher effectiveness and the funding that the schools receive is made based on a set of data that fail to adequately reflect students' level of science knowledge, ensuring the ideal goal of "Science for All" is unlikely to become a reality with the administration of standardized testing through punitive measures (Aydeniz 2007; Brickhouse 2006). Taking into account the fact that teachers do not have access to the content of actual tests in many states and that their perceptions of the content of the test are simply shaped by their interpretation of the released sample items their perceptions of what the actual tests may represent is problematic. This is an important finding in that denying teachers access to the content of the test may mislead teachers and encourage them to focus on students' acquisition of factual knowledge despite efforts to design curriculum standards and assessments that encourage students' acquisition of critical thinking and scientific inquiry skills (Abrams et al. 2008).

How Does Standardized Testing Impact Teachers' Practice?

One of the purposes of this study was to explore the influence of standardized testing on science teachers' instructional and assessment practices. In terms of the impact of standardized testing on science teachers' instructional practices, 51% ($n = 81$) percent of the teachers stated that "standardized testing dictates how I teach and assess my students' learning". One teacher who felt the pressure of teaching to the test said, "I have concentrated on state standards and getting my students ready for the gateway test. This sometimes limits my freedom on doing things that will be more interesting and/or enriching for my students". Majority of the participants ($n = 108$; 86%) stated that they felt encouraged by their school administration to organize their lessons based on the objectives of standardized tests administered in their relative states. Not surprisingly, 37% ($n = 59$) of the participants stated that the changes that they made in their instruction due to

standardized testing contradicted with what they believed teaching and learning of science should look like. When the number of participants who stated a neutral view are equally distributed, the percentage of those who believed the changes that they made in their instruction contradicted their beliefs increased to 47% ($n = 78$). Interestingly, the majority of teachers ($n = 113$; 70%) did not believe that their teaching had an impact on how well their students performed on the standardized tests, in spite of the changes that they made in their instruction.

How Does the Standardized Testing Impact Science Teachers' Assessment Practices?

The majority of teachers ($n = 150$; 93%) who participated in this study stated that they made significant changes in their assessments due to the standardized testing. Only 7% ($n = 11$) of the participants stated that standardized testing did not have any impact on how they went about assessing their students' learning in science. The pressure of increasing students' test scores lead many teachers to more frequently use multiple-choice type assessments than they did in the past. For instance, teachers made statements like, "My assessments are less along authentic lines and more like standardized testing- even when standardized testing does not fit." and "I use fewer project-based assessments and more standard pencil-paper tests due to time restrictions."

However, it is important to note that the administration of the standardized tests encouraged some teachers ($n = 7$) to use assessments that focused on students' acquisition of critical thinking skills. One teacher said, "I include more open-ended questions to encourage writing. Students must explain and give evidence to support their answers." Another one said, "I try to ask more open-ended questions in labs. I also go over the scientific method and its applications in more detail. I give students problems where they have to apply scientific thinking to solve a problem." Additional qualitative data indicated that the teachers in one school district in Arizona were required to meet and design assessments that focused on students' critical thinking skills and were encouraged to use formative assessments for the purposes of improving students' test scores. The teachers who reported the positive influence of standardized testing on their assessment practices are from the same state. As evidenced in this finding, leadership can make a difference in the ways in which the administration of the standardized test used for accountability purposes impact science teachers' assessment practices.

What are Science Teachers' Justification for Making Changes in Their Assessments in the Context of Standardized Testing?

Science teachers provided diverse reasons for the changes that they made in their assessments in the context of standardized testing. Although only 21% ($n = 34$) of the participants stated that they made the changes in their assessments because they believed the administration of the test would improve student learning, 75% ($n = 121$) of the participants stated that they made changes in their assessments because they did not believe it was fair to their students to take the test without

adequate preparation. A sizable number of the participants ($n = 47$; 29%) stated that they made changes in their assessments because they feared losing their jobs. These findings indicate that the administration of the standardized tests encouraged some teachers to use instructional practices that they anticipated would help their students to perform better on the test rather than adopting instructional practices that have been proven to engage students in the meaningful learning of important science concepts (Abrams et al. 2008; Aydeniz 2007; NRC 1996; Madden 2008) and students' acquisition of scientific inquiry skills (Abrams et al. 2008; NRC 2000).

These findings demonstrate that overall, testing for accountability outlined in NCLB does not by and large support the practices described in the science education reform documents. It is important to recognize that the absence of a focus on reform minded pedagogy in the current test-based educational improvement policies may be the reason for why so many science teachers do not make a conscious effort to use instructional practices (i.e., inquiry-based teaching) that are promising in promoting the learning outcomes advocated by the recent science education reform documents. It follows that the testing-based educational improvement policies undermine the reform efforts in science education by allocating instructional time and resources in ways that do not help students to acquire the type of knowledge and skills deemed important by the science education community or the science education reform documents. This is because both teachers and schools are busy at work to increase their students' performance on standardized tests which at best only represents a weak proxy of students' level of understanding of scientific concepts and processes (Aydeniz 2007; Darling-Hammond and Adamson 2010; Fu et al. 2009; Madden 2008).

Advantages and Disadvantages of Using Standardized Testing in Science from Teachers' Perspectives

Teachers' responses about the advantages and disadvantages of using standardized testing in science were placed in two categories; responses related to teaching and responses related to learning. The aspects that are related to teaching involve statements such as "standardized tests force me to teach to the test," or statements such as "standardized tests prevent me from teaching science through inquiry." The responses related to student learning involve statements such as "standardized tests allows for evaluation of students' critical thinking skills" or "standardized tests do not measure what students learn in my science courses". Teachers' comments in each of these two categories were further divided into positive and negative statements. Positive statements reflected the perceived positive influence of standardized testing on teacher practice or student learning (i.e., acquisition of scientific inquiry skills). Negative statements refer to the perceived negative influence of standardized testing on teachers' practice (i.e., teaching history of science concepts instead of critical thinking skills) or student learning (i.e., focus on factual knowledge). These qualitative analyses are reported in Tables 3, 4, 5, and 6.

The analyses reveal that the majority of teachers' comments about the influence of standardized testing on their teaching practices were negative (i.e., 164 negative

comments vs 55 positive comments). Teachers who acknowledged the negative influence of standardized testing acknowledged that the administration of standardized tests for accountability purposes encouraged them to "water down" the content of the courses they taught, and limit the content of their teaching to the content of the test which they perceived as being factual information rather than emphasizing critical thinking and inquiry skills. Teachers who acknowledged the positive impact of standardized tests on their teaching maintained that administration of such tests for accountability purposes encouraged all teachers to cover state standards more rigorously. They stated that holding teachers accountable for students' test scores ensured the uniformity of the content taught across different schools and districts within their relative states. In terms of science teachers' comments related to student learning, there were more negative comments ($n = 238$) than there were positive comments ($n = 21$). The changes brought about in science teachers' assessment practices due to the administration of standardized tests are reported in Tables 5 and 6.

As data reported in Table 6 indicate that the majority of teachers emphasized the negative impact of standardized testing on student learning. First, teachers did not think that the test results accurately reflected the actual learning that is taking place in the classroom as some students do not take the tests seriously. Second, teachers did not think that a student's level of learning in one particular domain of knowledge could be captured through one or two questions as is often the case with statewide-administered standardized tests. Finally, high school teachers did not think that the tests measured the type of knowledge and skills that are learned in the courses that they taught. Instead, they believed the test questions focused on assessing students' acquisition of basic science knowledge and skills that they thought their students learned in middle school.

Although the majority of the participants' comments emphasized the negative influence of standardized testing on their assessment practices, a small number of participants ($n = 17$) expressed the positive influence that the administration of standardized tests had on their assessment practices. Table 8 shows examples of positive and negative statements that the teachers made about the influence of standardized tests on their assessment practices.

The analysis of science teachers' responses indicates that standardized testing had a significant influence on how teachers go about assessing students' learning in science. About 90% of the participants stated that they now spend significant amount of their instructional time on teaching test-taking strategies. For instance, some teachers allocated the 2 weeks before the administration of the test to teaching test-related content that they did not think they would normally cover in their courses. Although 12% of the participants thought spending time on test-taking strategies helped their students to learn science better, 88% percent of them indicated that spending time on test-taking strategies is detrimental to their students' learning in science.

The changes in science teachers' assessment took place in two forms; changes in the format of their assessments and changes in the content of their assessments. Although 74% of the participants indicated that they made changes in the format of their assessments (i.e., inclusion of more multiple-choice questions, making the look

of their test to mimic the look of standardized test), 42% of the participants stated that they made changes in the format of their assessments (i.e., more open-ended questions) to help their students perform well on the standardized test. Further analyses reveal that standardized testing had a negative impact on the majority of science teachers' classroom assessment practices. Negative impact in this context refers to teachers using assessments that primarily measure students' acquisition of scientific facts rather than their acquisition of scientific inquiry and critical thinking skills, and students' development of conceptual understanding.

Although standardized testing is known to have a negative influence on science teachers' classroom assessments (Jones et al. 2003; Madden 2008; Popham 2006), the results of this study reveal that standardized testing policies had positive influences on some science teachers' assessment practices as well. For instance, 4% of the participants stated that they now use formative assessments more frequently in their assessments than they did before. A small number of science teachers ($n = 13$) indicated that standardized testing had encouraged them to include more critical thinking questions in their assessments. One participant stated that he/she made her laboratories more inquiry-oriented because of the standardized test administered in her state. The majority of science teachers who participated in this study questioned the validity of the test results. Teachers stated that their students knew that science scores did not held them or their schools accountable for the science results so they did not invest the effort that they could have invested otherwise in performing well on the tests. A few number of teachers ($n = 23$) stated that the students did not care about the results of the test even if the test held them accountable for their performance on the test.

Discussion and Conclusion

The purposes of this study were to (1) explore secondary school science teachers' attitudes toward standardized testing, (2) explore their justifications for the attitudes that they hold toward standardized testing, and (3) report on the changes in secondary school science teachers' instructional and assessments practices due to the administration of state-mandated standardized tests. We discuss the findings in the order of the purposes described above.

Science Teachers' Attitudes Toward Standardized Testing

The study participants held diverse views about the role of standardized testing in bringing about improvements to students' learning in science. The majority of participants held negative views on the role of standardized testing in bringing about improvements to student learning in science. However, their views were influenced by their beliefs about what they thought was important for students to learn in science classrooms, the type of test that was administered in their states and their interpretation of the information that was available to them about the content of the test. Therefore, science teachers used different criteria by which they judged the value of standardized testing in improving the quality of student learning in science.

For instance, while the teachers who believed that students should learn substantial amount of scientific knowledge advocated the administration of standardized tests, those that valued students' acquisition of scientific inquiry and critical thinking skills opposed the administration of the standardized tests for accountability purposes.

Previous studies reveal that science teachers' perceptions of the role of standardized testing in educational reform is influenced by the type of students (i.e., standard level students vs advanced placement(AP) level students) that they teach (Finnigan and Gross 2007), their beliefs about curriculum (Bianchini and Kelly 2003), teaching and learning of science (Aydeniz 2007; Madden 2008). The majority of participants in this study taught science to students of diverse ability levels. Their experiences with students from different ability groups may have influenced their attitudes toward standardized testing. However, we cannot make such grand judgements from the findings of this study due to the limited number of participants. Future studies that explore this relationship more systematically may offer more valuable insights into science teachers' perceptions of standardized testing, the factors impacting their decisions and the instructional and assessment practices that their perceptions of standardized testing lead to than we did in this study.

Science Teachers' Justification for the Attitudes They Hold Toward Standardized Testing

Teachers' expressed diverse justifications for the attitudes that they held toward the administration of standardized tests for accountability purposes. Teachers who expressed positive attitudes toward the administration of the standardized tests believed that testing for accountability purposes encouraged teachers to cover the state standards in their teaching in an effective manner and thus ensured the uniformity of the content taught across classrooms in their states. This is consistent with previous findings (Aydeniz 2007; Madden 2008; Shaver et al. 2007). These teachers also believed that administration of standardized tests would ensure that all students are acquiring the type and level of knowledge needed for them to graduate from high school. Teachers who expressed negative attitudes toward the administration of the standardized tests for accountability purposes provided diverse reasons for the attitudes that they held toward such tests. First, they believed that the test limited the type of knowledge and skills that students are learning in science classrooms to the factual information and discouraged them from promoting students' acquisition of critical thinking and inquiry skills. Second, they did not think that the test measured the type of knowledge and skills that they thought their students learned in the courses that they taught. Finally, these teachers believed that administration of the test encouraged science teachers to teach to the perceived test content rather than designing instruction based on students' interests and for promoting students' acquisition of scientific inquiry skills. These findings collectively suggest that teachers held negative attitudes toward the administration of standardized tests for diverse reasons: 1) because of the misalignment between the content of their teaching and that of the test (i.e., the test did not measure what they

covered in their lessons) and 2) because of their perceptions of the type of knowledge and skills that the test measured and what they believed was important for their students to learn in science (i.e., scientific inquiry skills). For instance, while teachers who held a positive view thought that the test ensured students' acquisition of essential knowledge needed for them to graduate from high school, those who held a negative view said the test did not measure students' critical thinking skills. It follows that if the goal of test-based accountability measures is to improve student learning in science, this purpose must be communicated to the teachers in an effective manner with adequate and convincing justifications along with professional development needed for them to become effective teachers. Otherwise, teachers may engage in practices that are counter to the intended purposes due to teachers' misinterpretations of the intentions of such tests.

The Effects of Standardized Testing on Science Teachers' Instructional and Assessment Practices

The findings indicate that the administration of standardized testing has multiple effects on the teaching and assessment practices of science teachers who participated in this study. Although a significant number of science teachers indicated that they made changes in their teaching and assessment practices that they considered negative due to the administration of the standardized testing a small number of the participants reported the positive influences of standardized testing on their assessment and instructional practices. This findings suggest that whether the perceived impact of standardized testing is positive or negative depends on how the test is being administered, how the expectations are being communicated to the stakeholders and how the teachers are being supported as they are trying to make the goals of science education reform prevail in nation's science classrooms while addressing the accountability needs of the system. For instance, in one particular state, teachers felt that they were being monitored by their school district to use formative assessments and incorporate more writing in science instruction. Integration of writing with science though appears to be a positive change, the motivation for this integration lead to some questions that beg answers. For instance, is writing being integrated in science because it supports students' science learning (Yore et al. 2008) or is it simple because science has become victim to the goals of raising students' writing scores on standardized tests—a heavily emphasized aspects of AYP calculations? Further research is needed to understand the very real implications of states' calculations of AYP for the manner in which science is taught (Pringle and Carrier 2005).

The findings of this study show that science teachers have several unique challenges in terms of their responses to the administration of standardized tests as reported in previous studies (Abrams et al. 2008; Aydeniz 2007; Madden 2008; Pringle and Carrier 2005). First, because many school districts do not count students' science scores in their AYP calculations and include their students' reading, writing, and mathematics scores, science teachers are pressured to emphasize reading, writing, and mathematical skills in their teaching (Aydeniz 2007; Madden 2008). In some schools, all teachers across curriculum are expected

to spend 15 min or more a day, five days a week on reading (Aydeniz 2007). Thus, instead of teaching science, science teachers are now expected to spend time on students' reading comprehension skills, not to enhance students' conceptual understanding of science, but to protect schools from financial and administrative sanctions (Aydeniz 2007; Madden 2008).

Second, science teachers who teach in schools in which students' science scores count in the AYP calculations, find themselves teaching to their misconceptions regarding the tests (i.e., such tests only address low level, facutal knowledge) rather than the content advocated in the national standards (Abrams et al. 2008). This is because most states do not make the tests accessible to the teachers due to the costs associated with the production of these assessments. For instance, one teacher said, "I think the test measures 8th grade science, not what I teach in 10th grade. That's why I spend the 2 weeks before the test on teaching general science concepts." As evidenced in this excerpt, the current system of accountability without full understanding of NCLB policies encourages science teachers to teach lower level content. Furthermore, because the focus on improving students' test scores encourages science teachers to use traditional methods of teaching (Abrams et al. 2008; Madden 2008; Pringle and Carrier 2005; Shaver et al. 2007), science instruction fails to address the learning needs of the disadvantaged students (i.e., students' with low motivation to learn science) (Darling-Hammond 2004; Kim and Sunderman 2005; Noblit et al. 2007). For instance, while research suggests that some historically disadvantaged minority groups and underachieving students learn science better through group work and hands-on and minds-on activities (Lee 2005; Noblit et al. 2007), standardized tests encourage teachers to focus on content coverage at a rapid pace (Aydeniz 2007; Madden 2008). Thus, instead of ensuring minority students' participation in science, the accountability systems that rely on test scores serve as an instrument of marginalization, leaving the students who have historically underperformed even further behind (Jones et al. 2003; Shaver et al. 2007).

Implications

These findings led us to drive several conclusions that have implications for multiple stakeholders. We discuss these implications below.

Implications for Policy

The purpose of the NCLB Act as well as other test score-based accountability systems are to improve either the achievement of students who are failing to perform at an acceptable level or the quality of learning for all students. The results of this study show that the majority of science teachers do not think that there is an alignment between the content of the tests administered and the curriculum that they enact in their classrooms. Similarly, the majority of participants emphasized the limitation of the assessments used to predict students' mental model of a science concept in that no one single question can determine what the students know about a

science concept. It follows that policy makers should focus on investing resources into curriculum and professional development of science teachers that has potential to improve the learning of historically underachieving students and of the students who need further challenge if they want to achieve their goals. However, as suggested by the teachers who participated in our study, current accountability system fails to support the kinds of science instruction that holds potential to enhance the quality of learning for all students—especially for those who are currently underperforming. This situation becomes particularly dire as the limited resources of states, districts, and schools are often devoted to the changes that will result in "quick fixes" for something that may require much more long term, and systematic attention (Abrams et al. 2008). If the goal of educational system is to make *"Science for All"* a reality, we need to pay special attention to the ways in which current accountability systems impact science instruction and student learning in science and the ways in which the changes in instruction and classroom assessment impact the learning of historically underperforming and underrepresented student groups.

Implications for Science Teacher Educators

Currently, science teachers are struggling to address the learning needs of all students while trying to accommodate the accountability needs of the system (Abrams et al. 2008; Aydeniz 2007; Madden 2008). Science teachers seem to be receiving two conflicting messages, one that calls on them to address the learning needs of all students by scrutinizing their instruction through testing another that calls on them to teach science through methods such as problem based learning that have been proven to ensure the learning of all students (Aydeniz 2007; Madden 2008). While teachers are being scrutinized to improve students' test scores, there is not a system through which we hold science teachers accountable for promoting the goals of science education reform in the classroom or to use instructional and assessment practices that are deemed effective (Abrams et al. 2008; NRC 1996) for enhancing the quality of learning for students who usually underperform in science in traditional learning environments (Brown 2004; Calabrese-Barton and Tan 2009). Science teacher educators must actively engage in national and local policies that impact the representation and enactment of science curriculum in the accountability measures in place and make sure that the goals of science education reform are not being ignored in favor of current accountability policies (Abrams et al. 2008). Science teacher educators need to consider designing and offering professional development opportunities that will help the practicing science teachers to consider the implications of their test-related practices for their students' learning in light of the new science education standards. Similarly, science educators should place a greater emphasis on pre-service science teachers' understandings of different purposes of assessment. Such understanding is likely to help them make informed choices as they respond to the pressures to increase students' scores on statewide tests.

Implications for Research

Clearly, there is a need for science educators to craft their work mindful of the very real demands placed on schools, teachers, and the students they serve. First, science educators should focus their efforts on conducting research that document and highlight the underrepresentation of science in test-driven recent educational reform measures. These efforts are likely to help ensure the equitable representation of science in the current accountability systems. The fact that most states do not count students' science scores in their AYP calculations could mean that science teachers do not receive enough financial and professional development support to promote students' learning in science. If states exclude science scores in their AYP calculations, the state administrators and school principles will be motivated to divert their resources into curriculum areas that are counted in their AYP calculations to avoid financial sanctions or other punitive measures. However, science educators have limited empirical evidence to support these claims. An increased attention to researching these issues can make a significant difference in the professional lives of science teachers and the quality of science learning experienced by the students in nation's science classrooms.

References

Abell, S. K., & Volkmann, J. M. (2006). *Seamless assessment in science: A guide for elementary and middle school teachers*. Portsmouth, NH: Heinemann.

Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice, 42*(1), 18–29.

Abrams, E., Southerland, S. A., & Silva, P. (2008). *Inquiry in the classroom: Challenges and possibilities*. Greenwich, CT: Information Age Publishing.

American Association for the Advancement of Science [AAAS]. (1993). *Benchmarks for science literacy*: *A Project 2061 report*. New York: Oxford University Press.

Aydeniz, M. (2007). *Understanding challenges to the implementation of assessment reform in science classrooms: A case study of science teachers' conceptions and practices of assessment*. Unpublished Dissertation. Florida State University, Tallahassee, FL.

Bell, B., & Cowie, B. (2001). *Formative assessment and science education*. Boston: Kluwer.

Berliner, D. C. (2005). Our impoverished view of educational reform. *Teachers College Record, 108*(6), 949–995.

Bianchini, J. A., & Kelly, G. J. (2003). Challenges of standards-based reform: The example of California's science content standards and textbook adoption process. *Science Education, 87*, 378–389.

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*, 139–148.

Brickhouse, N. W. (2006). Celebrating 90 years of science education: Reflections on the gold standard and ways of promoting good research. *Science Education, 90*(1), 1–7.

Brookhart, S. M. (2006). *Formative assessment strategies for every classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.

Brown, B. (2004). Discursive identity: Assimilation into the culture of science and its implications for minority students. *Journal of Research in Science Teaching, 41*(8), 810–834.

Calabrese-Barton, A., & Tan, E. (2009). Funds of knowledge and discourses and hybrid space. *Journal of Research in Science Teaching, 48*(1), 50–73.

Corbett, H. D., & Wilson, B. L. (1991). *Testing, reform and rebellion*. Norwood, NJ: Ablex Publishing Corporation.

Darling-Hammond, L. (2004). Standards, accountability, and school reform. *Teachers College Record, 106*(6), 1047–1085.

Darling-Hammond, L., & Adamson, F. (2010). *Beyond basic skills: The role of performance assessment in achieving 21st century standards of learning*. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education.

Davis, N., Genc-Kumtepe, E., & Aydeniz, M. (2007). Fostering continuous improvement and learning through peer assessment: Part of an integral model of assessment. *Educational Assessment, 12*(2), 113–135.

DeBoer, G. E. (2002). Student-centered teaching in a standards-based world: Finding a sensible balance. *Science & Education, 11*, 405–417.

Delandshere, G. (2002). Assessment as inquiry. *Teachers College Record, 104*(7), 1461–1484.

Duschl, R. A., & Gitomer, D. H. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. *Educational Assessment, 4*(1), 37–73.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.

Finnigan, K. S., & Gross, B. (2007). Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low-performing schools. *American Educational Research Journal, 44*, 594–629.

Firestone, W., 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.

Firestone, W. A., Monfils, L., & Camalli, G. (2001, April). *Pressure, support, and instructional change in the context of a state testing program*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Fu, A. C., Raizen, S. A., & Shavelson, R. J. (2009). The nation's report card: A vision of large-scale science assessment. *Science, 326*(5960), 1637–1638.

Grant, S. G. (2000). Teachers and tests: Exploring teachers' perceptions of changes in the New York state-mandated testing program. *Education Policy Analysis Archives [On-line serial], 8*(14). Available: http://epaa.asu.edu/epaa/v8n14.html.

Greene, J. P., & Winters, M. A. (2003). *Testing high stakes tests: Can we believe the results of accountability tests? (Report 33)*. New York: Manhattan Institute, Center for Civic Innovation.

Ingersoll, R. M., & Perda, D. (2010). Is the supply of mathematics and science teachers sufficient? *American Educational Research Journal, 47*(3), 563–594.

Jones, M. G., Jones, B. D., & Hargrove, T. Y. (2003). *The unintended consequences of high-stakes testing*. Lanham, MD: Rowan & Littlefield.

Keigher, A. (2010). *Teacher attrition and mobility: Results from the 2008–09 teacher follow-up survey* (NCES 2010-353). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [October 10, 2010] from http://nces.ed.gov/pubsearch.

Kim, J. S., & Sunderman, G. L. (2005). Measuring academic proficiency under the No Child Left Behind Act: Implications for educational equity. *Educational Researcher, 34*(8), 2–13.

Klassen, S. (2006). Contextual assessment in science education: Background, issues, and policy. *Science Education, 90*(5), 820–851.

Lee, O. (2005). Science education with english language learners: Synthesis and research agenda. *Review of Educational Research, 75*(4), 491–530.

Linn, R. (2000). Assessment and accountability. *Educational Researcher, 29*(2), 4–16.

Linn, R. (2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher, 32*(7), 3–13.

Lynch, S., Kuipers, J., Pyke, C., & Szesze, M. (2005). Examining the effects of a highly rated science curriculum unit on diverse populations: Results from a planning grant. *Journal of Research in Science Teaching, 42*(8), 912–946.

Madden, D. B. (2008). *Middle school educators' perspectives of changes in teaching practices caused by implementation of the FCAT Science*. Unpublished Masters Thesis, Florida State University, Tallahassee, FL.

McMillan, J. H. (2001). *Classroom assessment: Principles and practice for effective instruction* (2nd ed.). Boston: Allyn and Bacon.

Mehrens, W. A., & Kaminski, J. (1989). Methods for improving standardized test scores: Fruitful, fruitless or fraudulent? *Educational Measurement: Issues and Practices, 8*(1), 14–22.

National Academy of Sciences [NAS]. (2006). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: National Academy Press.

National Research Council [NRC]. (1996). *National Science Education Standards*. Washington, DC: National Academy Press.

National Research Council [NRC]. (2000). *Inquiry and the national science education standards*. Washington, DC: National Academic Press.

National Research Council [NRC]. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

National Research Council [NRC]. (2005). *Systems for state science sssessment*. Washington, DC: National Academy Press.

Noblit, G., Hwang, S., Seiler, G., & Elmesky, R. (2007). Forum: Toward culturally responsive discourses in science education. *Cultural Studies of Science Education, 2*(1), 105–117.

Popham, W. J. (2006). Branded by a test. *Educational Leadership, 63*, 86–87.

Popkewitz, T. S. (2000). The denial of change in educational change: Systems of ideas in the construction of national policy and evaluation. *Educational Researcher, 29*(1), 17–29.

Pringle, R. M., & Carrier, S. M. (2005). The potential impacts of upcoming high-stakes testing on the teaching of science in elementary classrooms. *Research in Science Education, 35*, 347–361.

Shaver, A., Cuevas, P., Lee, O., & Avalos, M. (2007). Teachers' perceptions of policy influences on science instruction with culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching, 44*(5), 725–746.

Shepard, L. A. (2000). The role of classroom assessment in teaching and learning. *Educational Researcher, 29*(7), 4–14.

Smith, M. L., & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice*, *10*(4), 7–11.

Stiggins, R. J. (2004). New assessment beliefs for a new school mission. *Phi Delta, 86*(1), 22–27.

Whitford, B. L., & Jones, K. (2000). Kentucky lesson: How high stakes school accountability undermines a performance-based curriculum vision. In B. L. Whitford & K. Jones (Eds.), *Accountability, assessment, and teacher commitment: Lessons from Kentucky's reform efforts* (pp. 9–24). Albany: State University of New York Press.

Yore, L., Henriques, L., Crawford, B., Smith, L. K., Zwiep, S., & Tillotson, J. (2008). Selecting and using inquiry approaches to teach science: The influence of context in elementary, middle & secondary schools. In E. Abrams, S. A. Southerland, & P. Silva (Eds.), *Inquiry in the classrooms: Challenges and opportunities* (pp. 39–87). Greenwich, CT: Information Age.