Research in Science Education (2005) 35: 347–361
DOI: 10.1007/s11165-004-5599-z

The Potential Impacts of Upcoming High-Stakes Testing on the Teaching of Science in Elementary Classrooms

Rose M. Pringle and Sarah Carrier Martin
University of Florida

Abstract

In 1983, the National Commission on Excellence in Education in the United States issued a report called *A Nation at Risk: The Imperative for Educational Reform*. This report and other policy initiatives such as the No Child Left Behind Legislation recommended that the individual states institute assessments to hold schools accountable. This research explored the potential impact of impending standardised testing on teaching science in elementary schools in one school district in Florida. We explored the teachers' concerns about the upcoming high-stakes tests in science, possible impact on their curriculum and what changes, if any, will be made in the approach to science teaching and learning in their classrooms. As the teachers look toward the implementation of high-stakes testing in science, they have recognised the need to teach science. This recognition is not borne out of the importance of science learning for elementary school children, but rather out of fear of failure and the effects of tangible rewards or punishments that accompany high-stakes testing. In anticipation, the teachers are preparing to align their teaching to the science standards while aggressively searching for test preparatory materials. Schools are also involved in professional development and structural changes to facilitate teaching of science.

Key Words: elementary science, high-stakes testing, standardised testing, testing and science curriculum

In 1983, the National Commission on Excellence in Education issued a report called *A Nation at Risk: The Imperative for Educational Reform*. This report, along with other reports and policy initiatives such as the recent No Child Left Behind (NCLB) legislation, and subsequent analyses, called for curricular revisions, improved professional development, and the implementation of more rigorous standards and accountability mechanisms. The NCLB mandates that states institute high standards to homogenise and improve curricula and that annual assessments be conducted to hold schools accountable for meeting those standards (Amrein & Berliner, 2002). As a result, a rigorous testing process tied to schools' accountability evolved. In the United States where this study took place, the testing was confined to mathematics and reading; however, high-stakes testing in science also began in 2003. The purpose of our study was to provide baseline data of teachers' perspectives of the impending standardised testing in science. In particular, we explored elementary school teachers' concerns about the upcoming high-stakes tests in science, the possible impact on their curriculum and what changes, if any, will be made in their approach to science teaching and learning in their classrooms. As local school district

policies can impact teachers' perceptions (Shepard, 2002), we conducted this study in one school district in northern Florida.

In this school district, the previous testing in science, before the statewide comprehensive achievement test consisted of a small science section on the Iowa Test of Basic Skills (ITBS), administered at grades three and five. The ITBS, initiated in 1998 was discontinued in 2001. Prior to that, the California Achievement Test (CAT) tested basic science content knowledge. These test data for science did not affect school accountability or student performance records. When the ITBS was discontinued and schools began to be graded by the state based on student achievement, some schools in this district as in other schools around the United States began to focus instructions on the tested subjects; mathematics and reading (Abrams & Madaus, 2003). At the same time, the inclusion of science in the elementary school curriculum was basically left to the discretion of individual school policies and, in some cases, the teachers made the decision for their classes. However, recent educational policies in this state, influenced by the national NCLB act, include a return to achievement testing in science. Standardised science assessment scores will be used to rate students, schools, and districts. School accountability tied to student scores signalled school districts to re-emphasise science teaching in elementary classrooms. In this paper, we discuss the impact of high-stakes testing and report the results of research that sought to understand the impact of the impending standardised testing policies on elementary teachers' perceptions of teaching and learning science before the full implementation of the statewide science assessments.

Current High-Stakes Testing and Schools' Curricula

In recent decades, test scores have come to dominate the discourse about schools, their accomplishments and subsequent reforms (Amrein & Berliner, 2002; Hilliard, 2000). When assessment in the form of standardised testing is used for accountability purposes, it is referred to as high-stakes testing (Resnick, 2000), a buzzword with far reaching implications for teaching, learning and assessment. In fixing high stakes to assessments, policymakers borrowed principles from the business sector and attached incentives to learning and sanctions to poor performances on tests. Under pressure from the public, state legislatures used these test scores as a source of information to monitor educational policy. Thus, according to Smith and Fey (2000), demands for accountability emerged from the policy makers and politicians who were concerned primarily with the distribution of power and resources in order to further their own agendas. High-performing schools would be rewarded and the staff would be considered effective. Under-performing schools would be penalised, and to avoid further penalties, would improve themselves (Amrein & Berliner, 2002; Popham, 1999). The expectations were that students would be motivated to learn, school personnel would be forced to do their jobs, and the condition of education would improve without too great a cost per state. Now, the scores are used to control

and reform educational practices as well as make decisions about school effectiveness, teacher competence, student graduation and allocation of national and state funds.

Little research can be found in the extant literature about actual impacts of high-stakes testing programs. Critics of standardised testing discuss the lack of evidence in the literature that supports high-stakes testing as an instrument for school reform leading to higher academic achievement (Hilliard, 2000; Kohn, 2000). They argue that any accountability system that focuses on standardised test scores rewards only compartmentalised, superficial learning and therefore thwart meaningful educational reform. In exploring the impact of standardised accountability tests, Mitchell (1997) reports that school administrators are identifying these tests as barriers to school restructuring and in essence driving their educational programs in directions inconsistent with reform while simultaneously diverting resources from reform efforts. More so, these tests have been shown to provide biased measures of performance in the areas of gender, race and socioeconomic status (SES) and according to Hurn (1997), SES along with race and ethnicity are associated with educational achievement and attainment.

Some scholars feel that high-stakes testing creates a system that is both unfair and destructive to learning for all students. Kohn (2000) contends these tests simply fail to assess the skills that matter most and their true purpose is to amass enormous profits for corporations involved in test production. While the tests measure only a narrow range of behaviors and very low level thinking (Pringle, 2001), the pressure to make the grade is resulting in some very disturbing and unethical educational practices. Typical responses to standardised testing have resulted in teachers focusing only on information that will be tested while material that involves high order thinking and problem solving often falls by the wayside, thus compromising the education of the children. Teachers have reported that they spend the majority of the school day preparing students in the subjects included in the state-testing program (Jones, Jones, Hardin, Chapman, Yarbrough, & Davis, 1999). In one report, preparation for the tests resulted in a skimming of minutes off each period of the day to create a new test preparation period (McClaskey, 2001).

The focus on test-prep activities is usurping a substantive curriculum, as teachers are involved in test preparation. Much time is spent on the preparation of test taking skills to ensure high scores without the engagement in authentic learning (McClaskey, 2001; Mitchell, 1997; Traub, 2002) and such skills as problem solving and creativity. In the end, educators, politicians and the public who push standardised high-stakes testing as a means to uphold high national standards are accomplishing quite the opposite. Unfortunately, these practices are not isolated and are harmful to student leaning, their self-image and well-being (Haladyna, Nolen, & Haas, 1991). This, Kohn (2000) asserts, can discourage the development of collaborative and cooperative skills among learners and reduce teachable moments in the classroom to bothersome and interfering spots. More so, while principals expressed an awareness of the constraints offered by the tests, they, driven by the threat of withholding funds for low test scores, discretely urge their faculty to address specific content, attend to given skills and apply particular instructional strategies in

preparation for the tests (Mitchell, 1997). In many reports, teachers are pressured into teaching only those subjects being tested: math and reading, with all other subjects given minor considerations (Abrams & Madaus, 2003; Shepard, 2002). Science instruction often competes with social studies instruction, and teachers must prioritise and condense each curriculum to include either or both subjects (Jones et al., 1999).

Student achievement is only one of the casualties of high-stakes testing. Hilliard (2000) describes the effects on teachers as "creative strangulation." With the accountability measures mandated for teachers, many feel that they are no longer individuals with knowledge and skill but mindless robots who work in a factory. Teachers also feel ashamed and embarrassed if their students score low and relieved rather than proud when they score high, eliminating their sense of efficacy and diluting their sense of competence (Charlesworth, Fleege, & Weitman, 1994; McNeil, 2000). Testing pressures teachers to provide a stressful watered-down curriculum that is not exciting to teach. These pressures force teachers to provide activities that they know are not appropriate for their students. Some school-mandated materials are totally "scripted," further reducing teachers' abilities to respond to individual learners' needs. Teachers in one state, according to Jones et al. (1999), noted that they were losing their abilities to be creative planners and thinkers because they were only able to teach what "others" had prescribed. Jones et al. (1999) concluded that high-stakes testing and the high-pressure environment created by states' testing programs encouraged rebellion against the very reform goals that led to the testing programs. This, according to Sheldon and Biddle (2000), is a result of the mismatch between what proponents of the standards-based movement say they want to achieve, that is, students with greater knowledge and more sophisticated ways of using the knowledge versus the high stakes attached to the current tests. The authors recognise the lack of balance in the background literature regarding the value of high-stakes testing. However, there is very little literature supporting high-stakes testing as a true measure of student science knowledge and an abundance of literature citing problems with testing practices and programs.

Three major issues arise when the effects of standardised testing on instruction are considered. Charlesworth et al. (1994) identify these as:

> The question of to what extent standardised tests measure what children are actually taught. Second, is the extent to which instruction and testing match the guidelines set forth by national professional organisations. Third, is the issue of how the use of standardised tests may exert direct influence on the curriculum. (p. 197)

The use of standardised tests to measure not only student learning in science but also assess schools and teachers offers a role for the tests that can affect teacher beliefs and therefore teacher behaviors. Bandura (1986) posits that beliefs are the best indicators of the decisions people make throughout their lives. Haney, Lumpe, Czerniak, & Egan (2002) confirmed that there is a relationship between what teachers believe and what they do in the classroom. Research supports the idea that teachers are crucial change agents to educational reform and that teachers' beliefs are precursors

to change (Ajazen & Fishbein, 1980; Pajares, 1992). In the next section, we report what we learnt from the teachers in response to the impending standardised testing in science. We describe the teachers' responses, the impact on their curriculum and the schools' attempts to prepare the children for the tests.

Purpose

Discussions regarding high-stakes testing have been influenced by findings from reading and mathematics. As science becomes added to the list of tested subjects in this state, we recognise the need for research to explore the effects of these tests on science teaching and learning. This initial research attempts to delineate the potential impacts of such tests on elementary science curriculum as indicated by the teachers.

The purpose of this investigation was to examine teachers' perceptions and concerns about upcoming high-stakes testing in elementary science. These data, collected before administration of the new annual tests, will form the base for further empirical studies of the influence of high-stakes testing on science teaching and learning in elementary classrooms. The results from this study will be used to make comparisons and track the state of science teaching in elementary schools over time as influenced by the high-stakes tests. In this research, no effort was made to observe, and document science teaching at these schools in relation to the teachers' perceptions of teaching science.

Method

Sample

A convenience sample (Gall, Borg, & Gall, 1996) of ten schools in a suburban school district in north-central Florida was used. The schools were selected because we had access to them, were able to procure permission from both the school board and the principals, and felt that constraining the study to one district would limit the variability expected across school districts due to highly localised policies. Within the participating schools, questionnaires were administered to 100 teachers, with thirty-eight valid responses returned; a return rate of 38%. The participating schools were representative of the public schools in the district with students ranging from low to high socio-economic status. The respondents were first to fifth grade teachers, ranging in years of experiences from 1 to over 10 years. The average teaching experience for all the participants was 7.6 years. Twenty-three teachers had undergraduate degrees in education while 12 had masters, 2 had completed a specialist program and 1 with a doctoral degree.

Data Collection and Analysis

The questionnaire used in this study was designed by the researchers and consisted of sixteen questions organised into three sections. The first section asked teachers to report demographic information such as the current grade taught, years of experience and the highest degree earned. In the second section, eight questions sought to elicit the frequency and types of science teaching in their elementary schools and the teachers' perceptions of their levels of preparedness to teach science. The final five questions were open-ended and required the teachers to elaborate on areas of concern, possible impact of the upcoming high-stakes testing in science on their curriculum and envisioned changes that will be made to accommodate regular science teaching in their classrooms.

When we designed the questionnaire (Appendix 1), our main concern as science educators was to examine teachers' perceptions toward upcoming high-stakes testing of science in elementary schools. Item by item analysis was conducted on Questions 1–11 to gather background information on the teachers and to reveal the frequency of occurrences of science teaching. In analysing the responses to the open-ended questions, inductive analysis was used. We found clear categories of the teachers' concerns and these became the focus of analysis from which themes emerged. Following van Manen (1990), we recognised that themes do not arise purely from the data, but are influenced by our theories and values; the interaction between theory and data is reciprocal, and open for reinterpretation. Furthermore, according to Patton (2002), the inductive search for patterns is guided by the research questions and the motives of the researchers. The analysis began as we independently read and coded each of the teacher's responses to the questions. Then, both researchers reviewed the independent analyses, identifying common units for further analysis. The emphasis at this stage was to organise and condense the data into manageable units while ascribing major descriptive codes. Further analysis resulted in our identifying pattern codes that were later developed into themes. For the qualitative researcher, pattern codes reduce large amounts of data into a smaller number of analytic units revealing a plot of the terrain (Miles & Huberman, 1994). Further rereading of the responses and testing of the themes against the supporting and non-supporting evidence from the data resulted in the emergence of two major themes: firstly, the teachers' concerns about the upcoming high-stakes testing in science and secondly, the potential impact on the science curriculum. These two themes along with analysis of Questions 1–11 on the questionnaire are used as the basis for discussing the impact of impending achievement testing in science.

Results

Item by item analysis of questions four to eleven on the questionnaire, revealed the actual times teachers reported that they spent on science teaching as well as the schools' mandates of the time requirements. Further analysis of additional items in

this section described their textbook use, hands-on activities and their inclusion of inquiry-based science activities. Fifteen teachers reported that they taught science every day while the others reported an average of twice per week. We asked teachers how often they used the textbook as the basis for their teaching. Thirty-two teachers reported the use of the textbook as the basis for their science teaching in most of their lessons. One, however, reported that the textbook was never used at all. After reviewing the responses on the questions dealing with frequencies of hands-on science activities/demonstrations and inquiry-based activities, we realised that we should have clarified our distinctions between the two terms because of the similar responses. Specifically, "inquiry-based" science relates to the teaching of science focusing on student-driven investigations in which teachers, as facilitators, direct varying parts of the activities (Layman, 1996). "Hands-on activities/demonstrations" involve any physical manipulation of materials resulting in pre-determined results. For both questions, most teachers reported using either hands-on or inquiry-based activities. Eighteen teachers responded using these activities "sometimes" while 15 indicated "most" lessons included these activities.

Teachers' Concerns about the Upcoming High-Stakes Testing in Science

The teachers who responded to the questionnaire had a lot to say about their concerns about the upcoming science portion of the test. Nine of the participants said that they were very concerned; twenty-three were somewhat concerned, while only five teachers had minimal concerns. The open-ended format of the question allowed teachers to elaborate on their responses and provide the researchers with more insights into their concerns.

The teachers' concerns fell into five major categories: concerns about the effects of poor reading skills on student performance, time constraints to include science lessons in the school day, too much emphasis on the test, teacher preparedness, and the unknown about the test expectations such as the format and student preparedness. Of the 35 teachers who responded to this question, eight expressed concerns about students' reading abilities, seven were concerned with time to include science lessons, three felt the upcoming test placed too much emphasis on teaching to the test, five were concerned about their abilities to help students prepare for the test, and twelve were concerned about the unknown format and content of the test.

The latter category, the fear of the unknown, was an area where many of the teachers expressed concerns. Because the teachers are unaware of the test format, they expressed feelings of helplessness in preparing their students for the unknown task. In addition, many teachers expressed frustration with not knowing the levels and the areas of science content knowledge on the test, and therefore they felt ill-prepared to provide the necessary experiences in science for their students. One teacher described his or her concerns as, "The unknown!" followed by, "What is expected?" Another questioned the format of the test, "... can the children understand what they are asked to do? Also, is it more performance-based or fact-based?"

Teachers wrote about their concerns that students with poor reading skills would suffer, regardless of their knowledge of science. They worried that the testing format would measure students' reading and writing abilities; therefore, students with below grade level abilities would be at a disadvantage in their science-test performance. From their experiences with other high-stakes tests such as reading and mathematics, the teachers noted that these depend heavily on students reading the question, analysing the information, and providing short and long written answers. The following quotes are representative of the eighteen teachers who expressed such concerns "Will it be another reading and writing test, or will ALL children have a chance to succeed?"

"Since we are science lab. teachers, we question how the writing level of the students will affect their ability to demonstrate their knowledge of science concepts," and, "because about 50% of our students read below grade level, I am concerned that this will affect the students' scores on the FCAT rather than (test) their knowledge of science principals (sic)."

The perceived lack of time is a constant enemy in an elementary school classroom. Many demands are placed on teachers to cover a wide variety of materials in addition to the basic subject areas of reading, writing, mathematics, science, and social studies. Teachers worried about adding science to the busy school day, and according to one teacher her concerns are, "fitting it with state requirements: reading priority – $2\frac{1}{2}$ hours, math priority – 1 hour 15 minutes." Along with this, teachers expressed that, "lack of time would prevent students from getting all necessary information or concepts because science is so broad," and "it would be difficult for one teacher to handle in the way of responsibility, pressure and required coverage of materials."

Thirty teachers felt the upcoming tests in science will bring an additional time pressure for test preparation, and noted that science is an area in which they have little background and therefore are not able to adequately prepare their students for high-stakes science testing. They described a "general lack of science expertise and enthusiasm by elementary teachers," which impacts the level of science teaching and could affect the students' performance on the test. One teacher expressed the following concern:

> Teachers will lose sight of the wonder and motivation that science can be to students. I am torn between being happy that science is finally being attended to by our district and feeling disappointed that we as educators are being motivated toward change by fear of a test. It would be a shame to have a (test-preparation) science class with boring workbooks that teach to the test. Yes ... I am concerned that students will not be prepared. I am hoping that science will find some priority in some of the primary grades so that by the 4th and 5th grade we don't have (to) cram. I think improvement will take a few years.

The various concerns the teachers expressed about the upcoming testing in science illustrate their levels of discomfort. Indications were that all of these would have significant impact on the curriculum developed in their classrooms.

Perceived Impact on Science Curriculum

Teachers were emphatic in their responses that the upcoming high-stakes testing would dictate changes in their curriculum and the approach to teaching science. The focus of these curricular changes will be to ensure that science will be taught as indicated by the state's standards, and their classroom assessment strategies will somehow mimic the formats in the high-stakes tests. Indications were that science was being taught, and some schools' programs were already aligned to the standards of the state. "We already teach the standards and the science expectations and would not need to make changes." Others noted, however, that they would now need to tailor their science teaching to match the state's requirements if their students were going to be prepared for the tests. "Well, we will just change what we are doing in science and now focus on the state's standards," one teacher lamented. Another stated: "Eventually, as with all tests, our curriculum will be structured to cover test materials in the way they present them in the tests, ignoring how children learn science."

A common response among many of the teachers was, "We will now have to structure the curriculum to include all the standards." Following on this statement, some teachers noted that to satisfy new requirements being stipulated by their schools, they were currently involved in workshops preparing them for the new school year "to teach the science expectations and document state standards as they are being taught." Teachers had not been exposed to a sample of the tests and they described feelings of frustration with how to prepare their students for the unknown format and content of the upcoming tests. At this time, the official guiding document for their science teaching will be the state's mandated standards. However, some teachers were concerned that eventually "good" science teaching will be replaced by 'teaching to the test':

> We are already teaching science, but we will now need to structure the curriculum to cover test materials in the way they are on the test. Science will end up just like math, reading and writing where we only teach the kids the information and the skills to pass the tests.

A level of anxiety was identified among the participants in their expressed need to match their teaching and assessment strategies to the format of the test. All teachers in our sample had plans to incorporate test-prep activities in their curriculum as a means of ensuring that the students would be prepared for the tests. Their quest was to find and use materials to facilitate test preparation activities along with their teaching. In many cases, finding these test-prep materials and study guides was a school-wide effort. Some teachers reported that their schools were already in the process of procuring such materials for the new school year. Some teachers indicated that a reliance on these test prep materials would result in a narrowing of the science curriculum and a shifting of the focus away from meaningful science teaching and learning activities. According to one teacher, "We will probably have to use some types of study guides to make sure we are covering the basics and cut out a lot of hands-on activities or experiments."

While many curricular changes would involve the inclusion of strategies mimicking test formats, some teachers identified interdisciplinary teaching as a curricular change that would need to occur under the constraints of the quantity of information and the given teaching time. The teachers understood that even though time to teach science was a constraint, they now had no choice but to pay attention to science teaching and this would involve incorporating science with other subjects; "I will incorporate science into reading and writing," or, "I will integrate math and science, and in this way, the students can learn more of each of the subjects," were two of the typical responses.

Teachers also noted that their schools were involved in other efforts to boost science teaching and to make sure that their students would be prepared for the tests. Reports were made of planned professional development in science, and an increase in the amount of money available to procure science materials and resources, and in two instances, regular classrooms were transformed into science laboratories. There appears to be a definite scurry within the participating elementary schools to revive science teaching. Unfortunately, this is not borne out of a need for meaningful science teaching and learning, but rather a response to testing and accountability. As one teacher writes:

> What children need to know in science should not be influenced by the fact that the state is testing it. However, if it takes a test to make educators pay attention to science, so be it.

Discussion

This study explored the concerns of teachers and the impact on their curriculum of impending high-stakes testing in science, through self-reported data collected on a questionnaire. Data analysis revealed that many teachers, though not frequently, were already involved in science teaching in their classrooms. There was a heavy reliance on the reported use of textbooks when compared to activities to develop scientific skills, attitudes and values while learning content knowledge. Many of the teachers were most comfortable in the use of the science textbooks in their teaching. Rigden (1999) declares that such heavy reliance on the textbook is due to the teachers' discomfort with the subject and their lack of knowledge of both the content of science or the way scientific knowledge is acquired. In this school district, the heavy emphasis on reading and literacy may also contribute to an emphasis on a science textbook. In using the textbooks, the teachers according to Huber and Moore (2001) are eliminating evidence-based science thus promoting erroneous and impoverished concepts regarding the nature of science. Their actions are contrary to the state's expectation of science teaching as indicated in Florida's Curriculum Framework (1996), that it is important for students to experience and interact with the natural world before they learn terms, symbols, and equations that scientists use to explain the natural world and not just to absorb facts.

The teachers in this study indicated a willingness to teach science in their elementary classrooms influenced by the mandates in their state's standard for science and the accompanying Grade Level Expectations. A guiding principle, stated in the standards, clearly supports the teaching and learning of science as inquiry. It states:

> ... Students use creative thinking skills to generate new ideas, make the best decision, recognise and solve problems through reasoning, interpret symbolic data, and develop efficient techniques for lifelong learning. (Florida Department of Education, 1996)

However, driven by the impact of the high stakes attached to these tests, teachers and their schools have expressed that they are forced to focus on efforts to prepare the children for the tests. These high stakes include monetary rewards, job security and possible stigma for faculty, parents and children associated with low-test performances. The perception therefore is that if the results of the tests decide the level of rewards or punishments, then teaching to the test is the right thing to do. It is this focus that has had many educators lamenting over the impact of high-stakes testing (Amrein & Berliner, 2002; Hilliard, 2000; Mitchell, 1997) and the resulting narrowing of the curriculum while amassing enormous profits for corporations involved in test preparation and the accompanying test-prep materials (Kohn, 2000). At issue also, is the risk that when teachers feel such strong pressures to improve students' scores, they abandon what they know as effective teaching in order to involve students in test-prep activities. Educators agree that in the quest to improve scores, teachers could lose their ability to be creative planners and thinkers hence they become unskilled workers (Finneran, 2002; Jones et al., 1999).

While science may be taught in more classrooms as a result of the testing, there exists a conundrum when the quantity of science increases at the cost of the quality of the science. Within the constraints of the research and nature of the teachers studied, the analysis revealed that the upcoming high-stakes test in science is causing an increase in activities in elementary science while creating a high level of anxiety among the teachers and school administration. Indications are that much of these activities will be restricted to documentation of science standards, efforts to procure resources to facilitate test-preparation, and a heavy reliance on efforts to teach to the test. The high anxiety level among all the teachers in this study, and their need to obtain test-preparation materials to supplement the teaching of science to ensure students would be prepared for the tests is not unique. Several studies report that this is a common phenomenon among teachers due to the high stakes attached to the results (McClaskey, 2001; Mitchell, 1997; Traub, 2002). The movement toward high-stakes testing in science might block science teaching advocated by the very standards the teachers plan to uphold. Those included in this study strongly suggest that teaching to the test will dominate the teaching of science in their elementary schools. A focus on science facts will certainly not facilitate arguing and debating over evidence, and students are likely to miss the opportunity to learn very valuable reasoning and higher order thinking skills. In addition, the pressure to prepare children for the test will work against wider involvement in inquiry-based instruction. In this era of assessment and accountability teachers are overwhelmed by the amount of

content, and so the time it takes to engage students with hands-on, minds-on activities might be seen as a time-consuming luxury given the perceived needs to cover the content for these high-stakes tests.

Teachers over time develop knowledge and beliefs about their teaching which is consistent with their practice (Haney, Lumpe, Czerniak, & Egan, 2002; Pajares, 1992). When changes are implemented as in the case of testing procedures, inconsistencies may arise between their beliefs and the expectations of the impending changes. Teachers will need to be supported to accommodate these changes. The findings in this research have implications for professional development experiences for teachers. Indications are that the teachers are not prepared to effect curriculum changes consistent with standards when there is a perception that tests will measure only science facts. Credence has been given to professional development experiences in increasing teacher preparedness to deal with issues of curriculum and classroom management (Smith, Banilow, McMahon, & Weiss, 2002; van Driel, Beijaard, & Verloop, 2001). Given the enormous pressures that standardised tests place on teachers, rather than focusing on the tests, these professional development activities could provide opportunities to develop effective strategies that enhance learning, and also the sequencing and pacing of these strategies to build on children's prior knowledge.

Limitations

We recognise the need to address the limitations of this study. First, the small and limited sample of teachers is problematic. Second, in an effort to collect baseline data on teachers' perceptions before the first wave of standardised tests was administered in the state, time was a limiting factor preventing the usual piloting and validation of the instruments. However, as science educators, we recognise that there is a paucity of information about the impact of high-stakes testing on meaningful science learning in this school district. This research will provide the science education community with baseline data to enable further empirical analysis of the impact of testing on science teaching and learning.

Conclusion

As the teachers in this research look toward the implementation of high-stakes testing in science, they have recognised the need to teach science. While this is a positive effect of the test, the recognition is not borne out of the importance of science learning for elementary school children, but rather out of fear of failure and the effects of tangible rewards or punishments that accompany high-stakes testing. In anticipation, the teachers are preparing to align their teaching to the science standards while aggressively searching for test preparatory materials in science. Schools are also involved in activities such as professional development and structural changes to facilitate teaching of science. The findings in this research have implications for

further research. Assessment and accountability have played prominent roles in many of the reform efforts over the last decades. While these assessments are being externally mandated, it is the teachers who are responsible for the actual changes in the classroom. As they teach science mandated by the standards, further research should highlight how teachers' interpretations of the standards become translated into teaching and learning activities and the long-term effects of standardised testing on the teaching of science in elementary classrooms. The question is: Will high-stakes testing in science encourage meaningful science for our elementary students?

Correspondence: Rose M. Pringle, University of Florida, 2403 Norman Hall, PO Box 117048, Gainesville, FL 32611-7048, USA
E-mail: rpringle@coe.ufl.edu

References

Abrams, L., & Madaus, G. (2003). The lessons of high-stakes testing. *Educational Leadership, 61*(3), 31–35.

Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives, 10*(18). Retrieved 4 September, 2002, from http://epaa.asu.edu/epaa/v10n18/

Ajazen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1986). *Social foundation of thoughts and actions: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Charlesworth, R., Fleege, P. O., & Weitman, C. J. (1994). Research on the effects of group standardized testing on instruction, pupils, and teachers: New directions for policy. *Early Education and Development, 5*(3), 195–211.

Finneran, K. (2002). Testy about testing. *Issues in Science and Technology Online.* Retrieved 4 September, 2002, from http://www.issues.org/issues/19.2/editorsjournal.htm

Florida Department of Education. (1996). *Sunshine State Standards*. Tallahassee, FL: Florida Department of Education.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research* (6th ed.). New York: Longman.

Haladyna, T. M., Nolen, S. B., & Haas, N. S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher, 20*(1), 2–7.

Haney, J., Lumpe, A., Czerniak, C., & Egan, V. (2002). From beliefs to action: The beliefs and actions of teachers implementing change. *Journal of Science Teacher Education, 13*(3), 171–187.

Hilliard, A. G. (2000). Excellence in education versus high-stakes standardized testing. *Journal of Teacher Education, 51*(4), 293–304.

Huber, R. A., & Moore, C. J. (2001). A model for extending hands-on science to be inquiry based. *School Science and Mathematics, 101*(1), 32–42.

Hurn, C. (1997). *The limits and possibilities of schooling* (3rd ed.). Boston, MA: Allyn and Bacon.

Jones, M. G., Jones, B. D., Hardin, B., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan, 81*(3), 199–203.

Kohn, A. (2000). Burnt at high-stakes. *Journal of Teacher Education, 51*(4), 315–327.

Layman, J. W. (1996). *Inquiry and learning: Realizing science standards in the classroom*. New York: The College Board.

McClaskey, J. (2001). Who's afraid of the big, bad TAAS? Rethinking our response to standardized testing. *English Journal, 91*(1), 88–97.

McNeil, L. M. (2000). *Contradictions of school reform: Educational cost of standardized testing*. New York: Routledge.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.

Mitchell, K. J. (1997). What happens when school reform and accountability testing meet? *Theory into Practice, 36*(4), 262–266.

National Commission of Excellence in Education. (1983). *A Nation at risk: The imperative for educational reform*. Washington, DC: US Government Printing Office.

Pajares, F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307–332.

Patton, M. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Popham, W. J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership, 56*(6), 8–17.

Pringle, R. (2001). Standardized testing and accountability: A vision for the future. *Florida Educational Leadership, 1*(2), 16–21.

Resnick, L. B. (2000). From aptitude to effort: A new foundation for our schools. In L. Abbeduto (Ed.), *Taking sides: Clashing views on controversial issues in educational psychology* (pp. 206–210). Guilford, CO: McGraw-Hill.

Rigden, J. S. (1999). Training K-6 teachers to teach science. *Education Digest, 64*(8), 59–61.

Sheldon, K. M., & Biddle, B. J. (2000). Standards and accountability and Perils and pitfalls. In L. Abbeduto (Ed.), *Taking sides: Clashing views on controversial issues in educational psychology* (pp. 211–220). Guilford, CO: McGraw-Hill.

Shepard, L. (2002). The hazards of high-stakes testing. *Issues in Science and Technology Online*. Retrieved 4 September, 2002, from `http://www.issues.org/issues/19.2/shepard.htm`

Smith, M. L., & Fey, P. (2000). Validity and accountability in high-stakes testing. *Journal of Teacher Education, 5*(15), 334–344.

Smith, P., Banilow, E., McMahon, K., & Weiss, I. (2002). *The national survey of science and mathematics education: Trends from 1977 to 2000*. Chapel Hill, NC: Horizon Research.

Traub, J. (2002). The test mess. *New York Times Magazine.* Section 6 (4.7.02).

van Driel, J., Beijaard, D., & Verloop, N. (2001). Professional development and reform in science education: The role of teachers' practical knowledge. *Journal of Research in Science Teaching, 38*(2), 137–158.

van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy*. Albany, NY: State University of New York.