

**A Critique Analysis of NCLB, Increase Testing, and Past Maryland Mathematics and
Science HSA Exams: What Are Maryland Practitioner's Perspectives?**

Patrice Juliet Pinder

January 2008

**Paper presented at ASTE 2008 International Conference, St. Louis, Missouri, January 10 –
13, 2008**

Abstract

States are establishing high stakes assessments to serve as a measurement tool of students' academic abilities. This study essentially explored and sought to understand some of Maryland's mathematics and science practitioners' perspectives on increase testing, and the No Child left Behind Act. Basically, the study sought to understand and to examine the effects of increase testing on teachers and students. Four Maryland Practitioners were interviewed in May 2007 for their perspectives on the NCLB Act and increase testing. Additionally, a critique of past Maryland HSA exams in mathematics and science were also conducted to see what were some of the positives and negatives of the exams from a practitioner's perspective. Results basically indicated that Maryland practitioners do not feel that NCLB is serving the needs of either urban teachers or students. Qualitative results further revealed: seventy-five percent of those practitioners interviewed indicated that constant testing was ineffective, one hundred percent of practitioners interviewed indicated that NCLB's goals were unrealistic and do not take into account the 'real' needs of those inadequately prepared students who can not meet the challenges of constant testing, seventy-five percent of those interviewed indicated that there was an increase pressure placed on them to 'teach to the test,' and twenty-five percent of those interviewed also indicated that despite the negative views they and other practitioners may have of NCLB and increase testing, it is still felt that the federal government and policy makers will continuously advocate for testing and accountability into the future. Quantitative data findings (as revealed from critiquing past Maryland HSA math and science exams and also from reviewing secondary data as reported in Maryland's report card) suggested that there were more positive aspects of the Maryland HSA exams in math and science than negatives. Positives of the Maryland HSA exams were: (a) test design construction – with there being varied test items / questions which seemed to cater to a heterogeneous population of students, (b) validity of the exams, as the exam content seemed in line with the state's curriculum focus, and (c) reliability scores were reported to be greater than 0.80 for all forms of the math and science HSA exams. Interview data results seemed in line with quantitative data reporting by Phi Delta Kappa Gallup Poll conducted in 2007, as 55% of those polled by Phi Delta Kappa indicated that NCLB was ineffective and 69% of those polled also indicated that there was an increase demand placed on teachers to teach to the test. In light of a few limitations of this study and also taking into account the gaps or limitations existing in the literature at the present time, it is recommended that more qualitative studies be conducted that can explore and seek to understand the perspectives of the various states' practitioners, particularly those in math and science assessment areas.

{Research report & conference paper, 34 references}

INTRODUCTION

The quality of mathematics and science achievement of students in America has been a major area of concern for the last fifty years. Over that time, the federal government launched and participated in a variety of attempts to promote mathematics and science education. Arguably, the most significant of these occurred after the Soviet Union launched Sputnik in 1957 (Mathematics and Science Initiative Concept Paper, 2006). As a response to the Soviet's demonstration of superior technological capability, and within 12 years of Sputnik, America upgraded its mathematics and science education program, launched satellites, saw its own astronauts orbit the earth, and launched the first man on the moon (Mathematics and Science Initiative Concept Paper, 2006).

However, it was not until the mid-1980s that the Department of Education recognized the need to reach more students and teachers (Mathematics and Science Initiative Concept Paper, 2006). The National Science Foundation also worked to broaden the focus of mathematics and science education from a focus on an elite group of students to a focus on all students (Project 2061). From the 1980s onwards, NSF made tremendous efforts to increase curriculum development in mathematics and science for K-12 students and pre-service and in-service teacher education.

According to Diane Ravitch, testing or assessment and accountability are linked in educational assessment; however, it has not always been this way. In the 19th century students were tested to see how they mastered what they were taught. If they did not pass the test they were left back, no one was held accountable, not the student nor the teacher. Back then, most students who managed to make it out of high school and applied to college were accepted. Only the most prestigious colleges, such as Harvard, Princeton, and Yale required the passing of an

examination. This led to the formation of the College Entrance Examination Board in 1900. With the establishment of the field of educational psychology, a reform in educational testing was seen. Scientific methods were applied to perfect testing as a measurement of student performance. In order to make school a more pleasant experience for students, social promotion became popular in education. This practice did not hold students or teachers accountable for students' poor performance. Further, the field of education began pushing for accountability around 1966 with the release of the Coleman Report. This report was based on a study of inequities in opportunities to learn among children of different races. Adding funds to increase resources to low functioning schools brought about a demand for accountability. With this accountability came a demand for school improvement (Ravitch, 2002).

Moreover, Stiggins (2002) suggests that we must improve classroom assessment if we are going to use assessment to increase student achievement and school improvement. Stiggins says that the past 50 years has seen an evolution in school assessment that believes that school improvement requires three things: the articulation of higher achievement standards, transforming those standards into rigorous assessments, and accountability of educators.

History of high school assessments

In the last century or so, society became concerned with high school graduation. Few students attended high school, and even fewer graduated. Concerns in the past forty years about the social problems associated with those who do not complete high school, sparked the debates about educational accountability. The term 'drop out' was coined by Edward Bellamy in the 1960s. As concerns grew, there were many discussions about the differences between the General Education diploma (GED) and the high school diploma and how each fits into a changing society. The web of influences that shaped entrance and exits out of school, as well as

engagement with learning became of interest to those in the business world. The main discussion focused on the reciprocal relationship between labor-market participation and school attendance, and this push-pull relationship was clearly affected by high-stakes testing.

Further, graduation examinations first appeared in the form of the Certificate of Initial Mastery which was proposed by Marc Tucker in the late 1990s. This initial step led to education reform legislation in such states as Washington State, Texas, and Massachusetts in the early 1990s (Wikipedia, 2006). The Certificate of Initial Mastery was awarded to students who passed a series of performance –based assessments by age 16. After being awarded the certificate, students were qualified to choose to either go to work, enter a college preparatory program, or study for a technical and professional certificate (Wikipedia, 2006).

With the dawning of the year 2000 and beyond, we have seen creation of new high school exit exams. Although the exit exams have improved in format since the 1990s, they are still tied to mastery of curriculum content standards in order for one to graduate and obtain a high school diploma. Dorn (2003) writes about three key issues that profoundly affect the relationship between high stakes testing and graduation patterns. First, dropping out is more of an indirect, rather than direct result of high stakes testing. If students can not finish high school by nineteen years of age, they are lured away by the opportunities of the workforce. Secondly, the proliferation of diplomas at secondary or tertiary levels may mask the continuing inequalities in educational opportunity. Due to the high proportions of the population attending college at some point in life, one might claim that high school graduation is merely an intermediate step in formal schooling, on the way to something bigger. The diploma therefore becomes a gateway for college entrance, as well as a direct pre-requisite for the competitive labor – market (p.3).

Thirdly, schooling as a route to economic and social success has become part of the American belief in the existence of equality (Dorn, 2003).

Gaps / Limitations in the Literature

Very few studies perused have utilized qualitative studies that employ phenomenological approaches in interviewing teachers to get their first hand perspectives on NCLB, and increase testing. This (brief) study can help to fill in the gap that exist within the field in regards to this issue of NCLB, and increase testing (see gaps in works by Rose & Gallup, 2007; Stringfield & Yakimowski – Srebnick, 2005). This study addresses outstanding deficiencies of previous literature done on assessment as it employs a qualitative methodology which involves employing a phenomenological method of inquiry to get Maryland’s mathematics and science practitioners’ viewpoints on NCLB, increase testing, and the effects of such on teaching and education as a whole. Additionally, a critical analysis of past HSA mathematics and science exit exams was also conducted in order to identify some of the negatives, and in some cases, the positives of past mathematics and science HSA exams.

Purpose

The purpose of the study was to explore and seek to understand Maryland’s math and science teachers’ perspectives on the No Child Left Behind act and increase testing with an aim at filling in the gap existing in the education research literature. Basically, what are their views on the two issues and how have their teaching been impacted by NCLB and increase testing?

Objectives

The essential objectives of this study were to conduct the following tasks:

- To explore and understand Maryland mathematics and science practitioners' perspectives on NCLB, increase testing, and the HSA exam. Is there a correlation between increase testing and accountability and the increase in teaching to the test?
- To critique past Maryland's HSA mathematics and science tests design and tests construction to determine positives and negatives with regard to the past tests design and tests construction.

NCLB Act, Accountability, and AYP since 2002

The Elementary and Secondary Education Act (ESEA) passed in 1965 as part of President Lyndon Johnson's War on Poverty gave the federal government investment opportunities in elementary and secondary education. The ESEA was reauthorized in 1994, and signed into law by President George W. Bush on January 8, 2002 as the No Child Left Behind Act (Florida Education Association, 2004). The expansion of ESEA into the NCLB Act (2001) has brought with it the increased federal role in education.

The NCLB Act details each state's responsibilities as follows:

- To develop challenging academic standards, especially in mathematics and science.
- To assess students' progress by constant testing measures by way of the various states' assessments designed to ensure that students are meeting the challenging state academic achievement and content standards.
- To give an account of students' progress over the year in order to meet the accountability mandate of the NCLB Act. In other words, each state offering assessments must produce annual state and school district report cards that inform parents and communities about state and school progress (U S Department of Education, 2003).

- To hold schools, school districts, and the individual states accountable for making Annual Yearly Progress (AYP). The NCLB Act seeks to hold schools accountable for making academic progress with their students and for overcoming achievement gaps, such as those seen between poor and rich students, urban and suburban schools, and between minorities and non-minorities (Florida Department of Education, 2006). The ultimate goal of the AYP Plan under the NCLB Act is to accomplish a 100% proficiency level on each state's mathematics and reading assessments by 2013 – 2014 (nothing was mentioned of science).

The AYP formula is as follows:

- Determine how each state will define 'proficient' student performance in mathematics and reading.
- Decide what indicators of student performance they will include in their definitions of AYP.
- Set incremental AYP targets that would establish minimal levels of increased student performance from 2002 – 2003 through 2013 – 2014.
- Set an initial threshold that is at a minimum the higher of the percent of students proficient in the statewide lowest achieving subgroup, or the local school at the 20th percentile in the state (Florida Education Association, 2004).

Additionally, according to the Maryland State Department of Education (2006), in order for AYP to be assessed, each subgroup of students has to take the exams and perform at least at the proficient level in order for AYP to be met. In other words, African Americans, Asians, Hispanics, Whites, American Indians, and Special Education students have to be assessed as separate subgroups in order for AYP to be successfully determined for each school and

school district tested (the achievement levels are: basic, proficient, and advance, but students have to perform at the proficient level for AYP to be met). Based on the federal No Child Left Behind Act of 2001, schools must meet 39 criteria for AYP to have been made at that school. Districts must meet the same criteria, except that school grades are not taken into consideration. If any one of the 39 criteria is not met, the school has not made adequate yearly progress under the federal accountability plan (Florida Department of Education, 2006). If a school does not meet all 39 criteria, but is high performing or received an ‘A’ or ‘B’ school grade under the A+ plan, then the school will be designated ‘provisional’ AYP (Florida Department of Education, 2006). Here is an example of how AYP is determined (see: Florida Department of Education, 2006):

Table 1: Established AYP Standards for Every US State

AYP Status	Yes
Writing Criteria Met	Yes
Graduation Criteria Met	Yes
School Grades not D or F	Yes

<u>Subgroups tested</u>	<u>Reading (95% tested)</u>	<u>Math (95% tested)</u>	<u>Reading met</u>	<u>Math met</u>
Whites	Yes	Yes	Yes	Yes
Blacks	Yes	Yes	Yes	Yes
Hispanics	Yes	Yes	Yes	Yes
American Indians	Yes	Yes	Yes	Yes
Economically disadvantage	Yes	Yes	Yes	Yes
Limited English Proficiency	Yes	Yes	Yes	Yes
Student with disabilities	Yes	Yes	Yes	Yes

Source: Florida Department of Education, 2006.

So according to the Florida Department of Education (2006), in order for AYP to be met the following is needed:

- The school has to meet the 95% participation criteria for each subgroup of students (such as African Americans, Whites, Asians, etc).
- Writing proficiency must have increased by at least 1% from the previous year or the school has to consistently possess a 90% or better writing proficiency rate.
- Graduation rate must show at least a 1% increase over the previous year or the school must have maintained an overall 85% or better graduation rate.
- School grade must be at least a 'C' or above
- Reading rate for the school must be at least a 44% rate

Moreover, schools or school districts who fail to meet AYP is placed on a 'need to improve list,' and if they continue to not meet AYP after five years then that school becomes a 'failing' school.

Increase testing and school failures

Increase testing is not only leading to increase students' failures but also to increase schools and school districts' failures. Schools and school districts that fail to meet annual yearly progress (AYP) as mandated by the NCLB Act of 2001 is placed on a 'need to improve' listing for five years. After five years, if a school still continues to not meet AYP, it then becomes a 'failing' school. In the state of Maryland, a large number of schools located within Baltimore City (a largely urban, poverty stricken area) failed to meet AYP for 2006. For Florida, the results are similar to that of Maryland. Further, in 2003, only 23 percent of schools were making AYP in mathematics and reading (science was not mentioned).

The NCLB Act mandates that all states must have all children proficient (100%) in mathematics and reading by 2013 – 14 but for Maryland this seems impossible. Furthermore, for

AYP to be met, every tested student subgroups (i.e. African Americans, Economically disadvantaged, Native Indians, Hispanics, Asians and Whites) have to reach proficiency level based on an overall 95% participation rate. So, if any subgroup does not make the 95% participation for testing, then AYP is automatically not met by that particular school. Additionally, each student subgroup tested has to pass at the proficient level for AYP to be met by a school in any given academic year.

Nationally, the problem of failing schools has been a serious one. California in particular, estimated in 2004 that 67 percent of its schools would have failed to meet NCLB AYP requirements in 2005 (Paulson, 2004). One California teacher stated in support of this 2004 prediction (above):

‘What kind of accountability system would you have when over half the schools are failing (Paulson, 2004, p.3)?’

Critics of the NCLB Act and AYP are contending, despite contradictions by some proponents of the Act, that there is a significant correlation between increased testing and AYP and increase numbers of failing schools seen nationwide since 2002.

The large numbers of failing schools seen since the establishment of the AYP standards have led to the Department of Education tweaking some of its rules in order to grant more flexibility to a greater number of states who wanted to make changes to their accountability plans, and this move have led to some failing schools being taken off ‘failing’ lists.

What does literature in the field of education say about testing, effective teaching, and students’ ability to take on challenging exams?

According to the National Science Educational Standards, Teaching Standard A, science and mathematics curriculums and lesson content should originate and be designed from the

interests, knowledge, understanding, ability, and experiences of the students (Krantz, 2003). Moreover, teachers know the importance of cognitive theory in structuring appropriate experiences in science and mathematics. We try to introduce youngsters to those skills or concepts that will grow and carry over to other subject areas and also over into their real world adult experiences. Skills such as observing, classifying, predicting, critical analysis, and problem solving are all important cognitive processes that educators need to help build up in the early phases of students' academic careers (i.e. elementary, middle & high schools). Further, cognitive researchers, Rowell and Dawson (1980) have argued that according to Shayer and associates, students must be well advanced into the formal operational stage of thinking before they can understand complex mathematics and scientific concepts. And they further postulated that only about 30% – 40% of students aged 15-16 years have reached a certain cognitive developmental level to be able to grasp complex concepts in high school science or mathematics (Rowell & Dawson, 1980). However, Piaget's school of cognitive psychology refutes Shayers and his associates; in that, Piaget believed that young students (high school age or younger) could learn complex 'operational concepts', in any discipline, if it was broken down or simplified for them. In fact, it was believed that these youngsters would have a greater capacity for full meaning of the concept (i.e. meaningful vs. rote learning) and have a greater capacity for retention later in life if they were adequately taught concepts in a modified, simplified way by train and knowledgeable educators (Rowell & Dawson, 1980).

Rowell and Dawson's (1980) research also implied the following premises of Shayer (cognitive psychologist):

- Meaningful learning of a concept can result only when the individual has available the cognitive skills / cognitive capacity needed to understand the concept.

- The level of a student's cognitive skill can be measured by a diagnostic test.

Given the points aforementioned, this research contends:

- That it is impossible or unreasonable to give challenging assessments to students who either lack the mental / cognitive capacity to pass the exams, or lack the cognitive skills that would equip them to be able to undertake challenging assessments in mathematics and science.
- Before the federal government can establish AYP requirements for every student subgroup, all students should first be tested by way of diagnostic tests that would initially determine if all students are mentally / cognitively adept to pass high stakes tests.

Some students may be unable to pass challenging assessments, especially in math and science because their cognitive ability may be impaired due to disabilities or mental health issues. Therefore, student subgroups with disabilities may be unable to make their AYP requirements under the NCLB Act if they are not of the mental capacity to do so.

Frisbie (1998) stated that exams should be pitched at a level for all students. Therefore, exams should cater to a heterogeneous population of students and should therefore consist of items that are easy, moderately difficult, and difficult. If tests are pitched at a level that caters to only a homogenous population, for example, advance or very bright students, then this may affect the reliability of test results. If, the Maryland HSA consists of items that may be viewed as too difficult for the average 15 to 16 year old, then it is expected that their test scores will be lower than if the exam consisted of items that were of a wider range (easy, moderately difficult and difficult items).

According to the social constructivist theory, effective teaching must be improvisational because if the classroom is scripted and directed by the teacher, the students cannot co-construct

their own knowledge (Sawyer, 2004). So students' co-construction or active participation in their own knowledge acquisition is important for true understanding of scientific or mathematical concepts to take place. It is posited that if students are allowed to make meaning of scientific and math concepts (esp. in middle & high schools) rather than being guided into rote memorization of scientific and mathematical concepts then true learning would occur in American classrooms and thus test scores in math and science may improve. Conversely, if students are not allowed to make meaning of math and science concepts then their achievement may not improve over time. Freire (2005) favored the administration of constructivist's ideals in classrooms in order to bring about effective teaching and learning. He argued that rote learning should be gotten rid of in favor of meaningful and critical learning. Freire (2005) wrote in support of the aforementioned viewpoint the following:

.....machine like transference from which results machine like memorization, which I have already criticized. Critical study correlates with teaching that is equally critical, which necessarily demands a critical way of comprehending and of realizing the reading of the word and that of the world, the reading of text and of context. (p.40).

And

.....in general what has been done in schools lately is to lead students to become passive before the text. Children learn early on that their imagination does not work, it is a forbidden sin. (p.57).

Further, Freire (2005) rejected the 'educational banking' of conceptual knowledge. In other words, he rejected the notion of teachers making 'deposits' of information into students, which students are to receive, memorize, and repeat. He rejected this passive recipient approach by learners in learning. Like John Dewey, and other constructivists, he too believed in the ideals of:

- critical assessment of knowledge
- synthesis of knowledge after making meaning of knowledge through critical analysis
- complete rejection of rote learning
- active participation versus passive approaches to learning by students
- educators creating creative lessons that can relate any new concepts to the real world experiences of their students

It is posited by Freire (2005) that if a teacher cannot relate new knowledge with students' real world experiences then this will cause a disconnection in learning by students. In other words, students will not assimilate nor accommodate conceptual knowledge; rather, they will reject such knowledge because critical assessment had not occurred due to disconnection between content understanding and students' reception of the concept / skill (Freire, 2005; Posner, Strike, Hewson & Gertzog, 1982). So, if teachers can introduce new ideas in ways their students can relate to, then their students should be able to accommodate and assimilate those concepts to make for better test scores and improved student achievement.

Finally, social and economic issues may impede students' learning and students' achievement. It is postulated that students' socioeconomic backgrounds may significantly impact learning and achievement, either positively or negatively. Mortenson (2001) points out that his belief is that parents' socioeconomics can have powerful affects on both k-12 learning and success in college. For example, 13 year old students whose parents had not graduated from high school scored 9 percent lower on math and science achievement tests and 12 percent lower on reading achievement tests than students whose parents had at least some college (US Department of Education, 2000a). Further, average scores for 9 year olds whose parents had some college were roughly comparable to those of the 13 year olds in the study. About 25 percent of high

school graduates from the lower SES levels go to college and earn a degree; whereas, nearly 80 percent of the graduates from the highest SES levels do. Also, Black households headed by single mothers were found to suffer from high levels of poverty. Children of these individuals usually underperformed in schools as well (US Bureau of Census, 2002; Kauchak & Eggen, 2005). To further advance the socioeconomics argument, socioeconomic status has been found to be related to school success. Researchers have found that high SES students scored higher on intelligence and achievement tests, got better grades in school, stayed in school, and the quality of their home life positively influenced student learning and performance (Kauchak & Eggen, 2005; Macionis, 2003; Economic Policy Institute, 2002). Konstantopoulos (1997, p.5) and others have found that low SES students usually scored lower on intelligence and achievement tests, got lower grades, stayed in school for a shorter duration of time, and the quality of their home life negatively influenced student learning and performance.

Brofenbrenner's (1974) ecological model posits this social ordering of social variables in a child's life and the impact on his/ her overall development. Brofenbrenner (1974) postulates

MICROSYSTEM-----→EXOSYSTEM-----→ MACROSYSTEM

(family, classroom) (school, community) (economics, culture, social values)

According to Brofenbrenner (1974), each person is significantly affected by interactions among a number of overlapping ecosystems. At the center of the ecological model is the individual. Microsystems are the systems that intimately and immediately shape human development. The primary Microsystems for children include the family, peer groups, classroom and neighborhood. Interactions among the Microsystems, as when parents and teachers coordinate their efforts to educate the child take place through an intermediary system located between the microsystem and exosystem called the mesosystem. Further, surrounding the Microsystems is the exosystem

which includes all the external networks, such as community structures and school structures. Surrounding the exosystem is the larger macrosystem. The larger macrosystem which includes social values, cultural values, and economics influence the other two systems; that of, the microsystem and exosystem.

The relevance of Bronfenbrenner's (1974) ecological model to this study is that visually it depicts that parents' cultural values, social values, and economics (macrosystem elements) can influence family values, classroom interactions, and school performances (microsystem elements) which in turn can have lasting impacts on the school community (exosystem). So, families' socio-cultural issues and socio-economic issues can impact students' classroom interactions (either positively or negatively) and also their academic achievement (Bronfenbrenner, 1974; Ogbu & Simmons, 1998; Ogbu, 2003).

Overall, all of the theories outlined has influence the study as follows:

- Piaget and Shayer's cognitive theory (outlined in the article by Rowell & Dawson, 1980) have caused the author to examine the possible correlation between students' true cognitive abilities and achievement levels. That is, are some students or subgroups more equipped mentally to comprehend and retain knowledge, and thus are better able to achieve high scores on achievement tests than others who are mentally challenged?
- Freire's (2005) constructivist theory have caused the study to examine the possible relationships between teachers' effective teaching, students' critical assessment and accommodation of concepts, and their resulting achievement levels.

The theories have formulated the following critical questions as it relates to this study:

- Do high stakes test truly measure students' abilities and comprehension of mathematics and science concepts?

- Are high stakes test pitched at a level that can produce good results for all students? Do high stakes exams truly cater to a heterogeneous student body with disabilities, special needs, mental challenges, and money or economic problems?
- Can teachers' adoption of constructivists' ideals improve classroom teaching, students' motivation to learn, and thus increase test scores in mathematics and science?

METHODS AND DATA SOURCES

Method

The current study utilizes a mix methods design in order to capture the widest range of effects of testing and accountability efforts, 'the what, and so what of reform efforts' (Stringfield & Yakimowski - Srebnick, 2005). Qualitative and quantitative data analyses have been utilized (i.e. quantitative findings as reported in Maryland's Report Card data reports). Qualitative data have been collected by way of semi-structured interviews with Maryland's mathematics and science practitioners. Additionally, quantitative data have been collected through critical analysis of past exit exams in mathematics and science and from data obtained from Maryland State's Report Cards.

Qualitative Design

Four practitioners, 2 Maryland math high school teachers and 2 Maryland science high school teachers were interviewed on May 1, 2007, by a single interviewer. A semi-structured interview format was used to allow the researcher flexibility in adapting the particular question(s) to the interviewee's understanding. So, if an interviewee wanted clarification, the interviewer could have rephrased the question(s) or ask a different, but relevant question, as it related to the issue at hand. So, the interviewer was not restricted in her ability to change the questions or the interview format. Three areas of interest investigated during the interviews were:

benefits (or lack thereof) of increase testing, the retention of testing and accountability within 5 - 10 years, and testing and accountability affects on teaching and teachers in general. There was a list of five suggested questions, such as: *Do you think the goals mandated by NCLB are realistic goals? Do you think the intended goals of NCLB will be realized within the next 5-10 years? How have your teaching been affected by testing and accountability? Do you think testing and accountability will increase or decrease in importance over the next 5-10 years? And, do you think constant testing is beneficial to students, especially urban students?*

The researcher introduced the topic(s) of conversation and through questions steered the course of the interview, although the interviewer was free to change a question if the interviewee needed further clarification. The interviews were conducted during the spring of 2007, and the interviews were conducted in the morning and evening hours. The interview data were coded and some examples of the codes that were used to sort the interview data were: *ineffective (useless) Act, inadequately prepared, continued advocacy, and increase pressures*. These codes (aforementioned) were used to identify common themes between the states' math and science practitioners.

Quantitative Design

Quantitative data analysis involved critiquing past Maryland mathematics and science exit assessments for positives and negative aspects of the past mathematics and science HSA exams as it related to such aspects of the exam as: state's standards, rubric / scoring, reliability issues, validity, readability and biasness of test items over the last few years of exam administration.

Limitation / Bias of Study

There were limitations to the study with the interview process. Firstly, only Baltimore City School practitioners were interviewed and not other practitioners from other Maryland counties. Due to time constraints and limited resources, only a few practitioners were selected for the interview through convenience sampling, rather than random sampling, and an assessment of their view points of the NCLB Act, increase testing, and its impact on them and their students were explored for a richer and deeper understanding. Future research may address these limitations by employing a study that may interview practitioners representing more Maryland school districts and not just the Baltimore City School System, and by employing random sampling instead of convenience sampling. Due to this limitation, the small population of Baltimore City school teachers' interview responses can not be seen as generalizable to all teachers' perspectives representing all Maryland school districts, but rather, the views expressed in this study represents only the opinion of a limited teacher population representing a single school district.

Qualitative Results

Maryland's mathematics and science practitioners had an opportunity to express their view points on NCLB, high stakes testing, and its impact on teachers and students. About 70% of those practitioners interviewed mentioned that constant testing was not benefiting students, especially urban students within Baltimore City. Furthermore, 75% of those teachers interviewed said that constant testing was *ineffective* and students were not benefiting from the barrage of tests. The following excerpts demonstrated practitioners' thoughts on the benefits of the NCLB Act and constant testing.

Practitioner 1 (math): 'In my opinion, the NCLB 2001 mandate is not working effectively for

students. Students don't get any benefits from it because they don't take their studies seriously.'

Practitioner 2 (biology): 'The constant barrage of tests is of no benefit to students. Baltimore City students are not benefiting from constant testing. The various state wide assessments are of no use.'

Practitioner 4(biology): 'No. Testing has no benefit to the students. Most of the students in Baltimore City are performing below their grade level. So testing them at their grade level does not make any sense. A solution should be found first to alleviate the situation.'

Practitioners felt that not only was the NCLB Act's mandate of constant testing not beneficial to students but the federal government and others failed to create a realistic and practical mandate that can meet the practical needs of all students, that is a heterogeneous student body of academically adept students and academically challenged ones. One hundred percent (i.e. all interviewees) of the practitioners interviewed talked about how unrealistic the goals of the NCLB Act was and how it failed to address the 'real' needs of those students who are inadequately prepared to take on the challenges of constant testing. The data collected clearly shows that practitioners think that the NCLB's goals are unrealistic and therefore do not take into account the inadequately prepared students who cannot pass high stakes assessments.

Practitioner 1 (math): 'I think the goals mandated by NCLB are not realistic because they don't give importance or emphasis to the root cause why students don't perform. They should do something about how to discipline the students. If students have self discipline then achievement will follow.'

Practitioner 2 (math): 'No, they are not realistic because some students will be left behind and

there are no accommodations for special education students and for students who are way behind academically.’

Practitioner 3 (biology): ‘The goals are not realistic. Students do not have the required skills to achieve the goals. The work ethics (academic preparation, attitude, aptitude) of students are in complete contrast to the goals of NCLB.’

Practitioner 4 (biology): ‘No. Students need preparatory courses for higher classes.’

But, not only is the NCLB Act and constant testing affecting only students, teachers are also feeling ‘the painful affects’ of this Act. There is an increase need to ‘teach to the test content’ rather than focusing on a broad and generalized high school curriculum within science and mathematics. In fact, about 75% of those teachers interviewed talked about an increase pressure on them to ‘teach to the test’ in order to satisfy school administrators and the school district.

Practitioner 1 (math): ‘It has a lot of pressure because it is expected that my students will have to perform or achieve. I need to use different strategies to be able to motivate them to do their work especially the BCR and ECR parts.’

Practitioner 2 (math): ‘Much of my teaching is aligned to assessments. I am held accountable for how well my students perform on these tests. When my students don’t perform well, administrators will constantly be on top of me to say what I am doing wrong.’

Practitioner 4(biology): ‘Yes. I am being forced to teach to the tests rather than being allowed to teach the important content areas for the science high school curriculum. Most of the time, I am forced to give students practice questions similar to the actual HSA test.’

According to the interviewees, there is the prediction that despite the unrealistic nature of NCLB, testing, and the accountability movement, testing and accountability will be predicted to persist in importance over the next 5-10 years and even beyond. Twenty-five percent of the practitioners interviewed spoke about the government and policy makers' continued advocacy for testing and accountability.

Practitioner 2 (math): 'I think it would increase because the politicians and policy makers or administrators would continue to advocate for testing and accountability because of their fixation to the two, even though, both do not translate into good student achievement in math or science.'

The above interviews with Maryland's mathematics and science practitioners suggested that increase testing is impacting teaching and students' achievement (esp. for urban poor minorities). In addition, practitioner 1 identified the importance of addressing the root causes of why students under perform (ex. socioeconomics or poverty, and race and ethnicity) rather than ignoring existing problems and creating additional ones (as with NCLB mandated goal of constant testing of all of the nation's youths). The data collected showed that in addition to increase testing affecting students' performance, it also affects teachers as well, creating an intense pressure to 'teach to the test' in order to get more students to pass these high stakes exams and to get administrators of their (i.e. the teachers') backs (see practitioner 2 quote aforementioned).

Summary of Qualitative Results

- 75% of the practitioners interviewed said that constant testing was ineffective.

- 100% of the practitioners interviewed said that the NCLB's goals were unrealistic and do not take into account the 'real' needs of those *inadequately prepared* students who cannot meet the challenges of constant testing.
- 75% of the practitioners interviewed said that there was an *increase pressure* on them to 'teach to the test.'
- 25% of the practitioners interviewed said that the federal government and policy makers will *continuously advocate* for testing and accountability into the future.

Quantitative Results and findings

The objective was to examine and critique the state's past mathematics and science exit assessments on the issue of proper adherence to state's standards, rubrics, scoring, test construction, test design and bias issues. Basically, what were some of the positives and negatives of each of the above examined variables from a practitioner's perspective?

Positives of the Maryland HSA mathematics and science exams: A practitioner perspective

Test Administration

It was a positive point to have administered the exam by grade 11 so that students can retake it before graduation if they fail the exam on the first try.

Test Design

It was a positive point where increase reliability of test scores were concern to have more objective (multiple choice) type questions on the mathematics and science HSA exams, rather than to have had the exams composed of more extended constructed or brief constructed responses (see Frisbie, 1998). Additionally, all of the past exams perused eliminated possible random guessing by utilizing a varied alphabetical selection for multiple choice questions, such as choice selections of either A - D or F, G, H and J for Maryland's HSA. Thus, this may have

attributed to the high Cronbach α rating of greater than 0.80, which was highlighted for the last few years of the Maryland HSA mathematics and science exams (Maryland State Department of Education, 2005). An example of this was question 2 of the fall 2004 mathematics HSA, in which the selection choices were a choice of either an F, G, H or J answer selection. For example:

A park ranger spent \$208 to buy 12 trees. Redwood trees cost \$24 each and spruce trees cost \$16 each. How many of each tree did the park ranger buy?

(F) 10 redwoods and 2 spruce

(G) 9 redwoods and 3 spruce

(H) 3 redwoods and 9 spruce

(J) 2 redwoods and 10 spruce

Also, question 3, where the selection choices were changed to either an A, B, C, or D answer selection as such:

Paul earned scores of 76, 78, and 86 on his first three science tests. What score does he need on his fourth test to earn a mean score of 82?

(a) 78

(b) 80

(c) 84

(d) 88

Readability

Clear and concise questioning was utilized as past exam items were carefully examined; this was a positive attribute of the exams as it ensured clarity of exam questions for students, which may have favored test score reliability and thus may have contributed to the reported high internal reliability rating of greater than 0.80 for the last few mathematics and science examinations perused (Maryland State Department of Education, 2005).

Validity

Content related validity seemed to have been met; in that, the exam items seemed to have correlated with the state's mathematics and science standards and the BCPSS schools' mathematics and science curriculum. This was a positive in the author's opinion (as a practitioner).

Reliability

The state's HSA mathematics and science assessments were reported to have a high cronbach alpha reliability rating of greater than 0.80. Maryland's algebra reliability rating was 0.89 to 0.90 for the three forms of the exam (forms A, B & W). Maryland's biology reliability also exceeded 0.80, at a reported rate of 0.92 to 0.93 for all five forms of the exam (forms A, B, C, D & W). Most literature report good exam reliability to be greater than 0.80; so again, this was a positive attribute in the author's opinion, who also happens to be a practitioner in the field (refer to Frisbie, 1998).

Rubric / Performance levels

Performance levels under AYP rules

Good to have established three performance levels under AYP, which were: Advance = outstanding / highest achievement, Proficiency = average / passable performance, Basic = below standards / not passing.

Rubric

Good to have five levels of rubric scoring, which were outlined as: Level 5 = full and complete understanding of questions / problems, Level 4 = good understanding of problems / questions, Level 3 = basic understanding, Level 2 = some understanding, Level 0 = no understanding of problems.

Test item types

The HSA exit exams in mathematics and science consisted of a wide range of test items. Test items were a combination of easy, moderately difficult, and difficult items. According to Frisbie (1998), a variety of test items is best in order to cater to a heterogenous student body, which may consist of strong as well as weak students. For example, the mathematics and science HSA exams over the past years have consisted of easy, moderately difficult, and difficult items, which is seen as a good exam attribute as it would allow all students, both weak and strong to have an equal chance to answer some of the questions on the exams. Specific examples of this taken from past mathematics and science HSA exams are listed as follows (see below examples): question 2, which is an example of an easy math test item, question 3, which is an example of a moderately difficult math exam item, and question 9, which is an example of a difficult math test item.

Question 2

A park ranger spent \$208 to buy 12 trees. Redwood trees cost \$24 each and spruce trees cost \$16 each. How many of each tree did the park ranger buy?

- (F) 10 redwoods & 2 spruce
- (G) 9 redwoods & 3 spruce
- (H) 3 redwoods & 9 spruce
- (J) 2 redwoods & 10 spruce

Question 3 was an example of a moderately difficult item.

Question 3

Paul earns scores of 76, 78, and 86 on his first three science tests. What score does he need on his fourth test to earn a mean score of 82?

- (a) 78 (b) 80 (c) 84 (d) 88

Question 9 is an example of a difficult item on the 2004 mathematics HSA.

Question 9

Look at the pattern 12, 36, 108, 324.....If this continues, what will be the next term?

- (a) 540 (b) 648 (c) 972 (d) 1,296

As for the science HSA, the exam also consisted of easy, moderately difficult, and difficult test items (see below items).

Question 9 (2005 HSA biology) is an easy test item

The energy required for photosynthesis is provided by

- (a) proteins (b) sunlight (c) chlorophyll (d) carbohydrates

Question 24 (2005 HSA biology) is a moderately difficult test item

Which of these describes the relationship between cholera bacteria and copepods?

- (F) mutualism (G) parasite – host (H) commensalisms (J) producer-consumer

Question 26 (2005 HSA biology) is a difficult test item

A scientist cloned a goat. Which of these is a true statement about the cloned goat?

- (F) It has new genes and traits
(G) It lacks the genes for reproduction
(H) It has genes that are identical to the original goat
(J) It looks the same as the original goat but has different genes

So again, we see a wide variation of test items with the Maryland's HSA mathematics and science exams.

Negativities of Maryland past mathematics and science HSA: A practitioner's perspective

Test Purpose

The Maryland HSA mathematics and science exams are not a graduation requirement right now, but will become a mandatory prerequisite for graduation in 2009; this may be seen as a negative, in that, this may cause students to not want to perform well on the examination prior to 2009, as they may not see the need to achieve high scores on an exam that is not a requirement for graduation.

Bias issues/ cultural and language sensitivity

The Maryland HSA does not cater to a diverse ethnic student population with varied languages, and therefore the Maryland HSA is only given in one form, which is English, this may hinder some student's performances.

Discussion

How do the interview data findings stack up against the existing literature in the field on the issue at hand?

Interview data in part, is in line with the reporting of the 39th Annual Phi Delta Kappa Gall up poll data; in that, interview data showed that seventy-five percent of the Baltimore City practitioners interviewed viewed increase testing and NCLB as being 'ineffective,' and this was in sync with the Phi Delta Kappa reporting. Phi Delta Kappa report indicated that in 2007 about 55% of those people who knew a great deal about NCLB had an unfavorable opinion of the act (Rose & Gallup, 2007, p. 34), and this was in line with my interview data results.

Interview data revealed the emphasis placed on teachers to 'teach to the test' while ignoring other curriculum focus. The Phi Delta kappa gall up poll results seemed to be in line with this finding, as they reported in 2007 that 94% of those people interviewed by them were

concerned about the reduced emphasis placed on teachers to focus on non standardized examination curriculum content (Rose & Gallup, 2007, p. 36), and also, it was reported that 69% of those Phi Delta Kappa pollsters felt that there was a great demand placed on them to ‘teach to the test’ (Rose & Gallup, 2007, p. 37), and also 79% of those interviewed by Phi Delta Kappa felt that it was ‘a bad thing’ to do (Rose & Gallup, 2007, p. 37). Similar, negative sentiments and concerns about ‘teaching to the test’ were also echoed in the research work of Stringfield & Yakimowski-Srebnick (2005).

The negative views or perspectives of NCLB and increase testing held of Baltimore City teachers seems to imply that they simply do not think the increase in testing or the act itself will benefit the urban, poor students. This implicit belief seems to reflect the findings of Kauchak & Eggen, (2005), Konstantopoulous, (1997), Macionis (2003), Mortenson (2001), Paulson, (2004), U.S. Department of Education (2000a), and U.S Bureau of Census (2002) who have all reported on the possible negative impact low socioeconomics have on school performances for urban Black students.

Conclusion and Recommendations for Future Research Study

What is known about teachers’ views / perspectives of NCLB, and increase testing? The research literature is robust with explanations given in numbers; that is, a lot of quantitative studies have been done on NCLB and increase testing and its effects on urban schools in particular. For example, Rose & Gallup (2007) Phi Delta Kappa reporting presented a lot of quantitative data to report generally on teachers and non teachers views of NCLB and testing, and Stringfield & Yakimowski – Srebnick (2005) in their research work utilized a lot of literature reviews and secondary data sources from Maryland Report Cards in order to report on the low academic performances of Baltimore City students. This paper, on the other hand,

addresses the limitation seen in the literature; in that, it utilizes a qualitative approach as opposed to the predominant quantitative approach and employs a phenomenological inquiry tactic in order to get Maryland (Baltimore City) teachers' perspectives on NCLB and increase teaching. By employing a qualitative perspective, we are giving voices to the teachers and we are finding out what are the 'real truths' behind the numbers reported in earlier quantitative studies.

Proponents of statewide testing may argue that a successful graduation exit examination assures standards motivates teachers to teach to the standards and increases school accountability.

However, critics of statewide testing may also argue that there will be challenges. As Statewide testing leads to a narrower statewide curriculum and the tests may require more remediation and preventive measures. There will be those that argue that a single test score does not measure student achievement and preparing for the test requires too much time away from non-testable material (Education Commission of the States, 2006), also arguments can be made for increase funding, as this will always be important, for more resources are needed to improve test scores in urban schools (Bhanpuri & Sexton, 2004).

Specifically, educators in Maryland feel that the increase testing will be seen as less intrusive once more teachers are involved in the development of the test items, and once less pressure is placed on teachers to 'teach to the test.'

In light of a few limitations of this study and also taking into account literature reviewed while conducting the study, it is recommended that more qualitative studies be conducted that will explore and seek to understand the perspectives and viewpoints of the various states' practitioners, particularly those in challenging assessment areas, such as mathematics and science.

References

- Brofenbrenner (1974) ecological model as outlined in Berger, V. (2000) Psychology, grand theories in psychology: Ecological theory. Retrieved Sep. 24, 2006 from <http://www.sesd.sk.ca/psychology/ecological.htm>.
- Bhanpuri, H., & Sexton, S. (2004, May). *A look at the hidden costs of high school exit exams*. Center on Education Policy. Retrieved December 20, 2006 from http://www.cep-dc.org/pubs/hiddencost_may04/HiddenCostFinal.pdf.
- Dorn, S. (2003). High stakes testing and the history of graduation. *Education Policy Analysis Archives*, 11 (1), 1-29, Retrieved November 7, 2007, from <http://epaa.asu.edu/epaa/v11n1>.
- Economic Policy Institute (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC.
- Education Commission of the States (2006). *High school – exit exam Education Policy Issue Site*. Retrieved December 26, 2006 from <http://www.ecs.org/html/issue.asp?issueid=108&subIssueID=159>.
- Education Commission of the States (2006, April). *High school- exit exams. Education Policy Issue Site*. Retrieved May 1, 2006 from <http://www.ecs.org/clearinghouse/68/35/6835.pdf>.
- Florida Department of Education (2006). 2006 Guide to calculating adequate yearly progress (AYP) technical assistance paper 2005 – 2006.
- Florida Department of Education (2006, May). *2006 FCAT mathematics results grades 3 -10 percent comparison by year and district*. Retrieved December 29, 2006 from <http://www.firn.edu/doc/sas/fcat/fcatscor/htm>.
- Florida Education Association (2004). *No Child Left Behind – NCLB / ESEA*. Retrieved December 28, 2006 from <http://www.feaweb.org/page.aspx?a=1588>.pps.1-2.

- Florida Education Association (2004). *In our view: FCAT + AYP = Confusion*. Retrieved December 28, 2006 from <http://www.feaweb.org/page.aspx?a=1834>. pps.1-3.
- Frisbie, D. (1998). Reliability of scores from teacher- made tests. *Educational Measurement: Issues and Practice*, 7, 25 - 35.
- Freire, P. (2005). *Teachers as cultural workers-letters to those who dare teach*. 5500 Central Avenue, Boulder, Colorado: Westview Press.
- Kauchak, D., & Eggen, P. (2005). Learner Diversity differences in today's students. In *Introduction to teaching, becoming a professional*. 2nd ed, (pp. 84-92). Upper Saddle River, New Jersey and Columbus, OH: Pearson, Merrill-Prentice Hall.
- Konstantopoulos, S. (1997, March, p.5). Hispanic –White differences in central tendency and proportions of high and low scoring individuals. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Krantz, P.D. (2003). Flying through the standards with bats. *Journal of Science Activities* classroom projects and curriculum ideas, 39, 29-33.
- Macionis, J. (2003). *Sociology* (7th ed.) Upper Saddle River, NJ: Prentice Hall.
- Mathematics and Science Initiative Concept Paper (2006). Retrieved December 5, from http://www.ed.gov/rschstat/research/progs/mathscience/concept_paper.doc.
- Maryland State Department of Education (2005). Maryland HSA Technical Report. Retrieved June 8, 2006 from <http://www.marylandpublicschools.org/msde/divisions/planningresultstest/2005+hsa+techinal+report.html>.
- Maryland State Department of Education (2006). *What is AYP? How does Maryland implement adequate yearly progress?* Retrieved December 28, 2006

from http://mdk12.org/MSPP/ayp/what_is.html. pps.1-2.

Mortenson, T. (2001, March). The condition of access: What the enrollment numbers say? Paper presented at the annual conference of the coalition of state university aid administrators, St. Augustine, Florida. Posted on the post secondary education opportunity web site: <http://www.postsecondary.org/archives/report/OSUAAAF31801.pdf>.

Ogbu, J.U. (2003). *Black American students in an affluent suburb (a study of academic disengagement)*. Mahwah, NJ: Lawrence Earlbaum Associates.

Ogbu, J.U., & Simmons, H.D. (1998). Voluntary and Involuntary minorities, a cultural ecological theory of school performance with some implications for Education. *Anthropology & Education Quarterly*, 29(2), 155-188.

Paulson, A. (2004). A shortening list of failing schools states gauge progress of Bush's education reforms – and debate. What it really shows. *The Christian Science Monitor*. Retrieved December 28, 2006 from http://www.csmonitor.com/2004/1123/po1s02_uspo.html. pps.1-4.

Posner, G.J., Strike, K.A., Hewson, P.W., and Gertzog, W.A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211 -227.

Ravitch, D. (2002). A brief history of testing and accountability. *Hoover Digest*, (4).

Rose, L.C., & Gallup, A.M. (2007). The 39th Annual Phi Delta Kappa / Gallup Poll of the public's attitudes toward the public schools. *Phi Delta Kappa*, 39, 33 – 48.

Rowell, J.A., & Dawson, C.J. (1980). Mountain or Mole Hill: Can cognitive psychology reduce the dimensions of conceptual problems in classroom practice? *Science Education*, 64, 693 -708.

Sawyer, R.K. (2004). Creative teaching: Collaborative discussion as discipline improvisation.

Educational Researcher, 33(2), 12-20.

Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta*

Kappa, 83(6), 758 – 765.

Stringfield, S.C. & Yakimowski –Strebnick, M.E. (2005). Promise, progress, problems, and paradoxes of three phases of accountability: A longitudinal case study of the Baltimore

City Public Schools. *American Educational Research Journal*, 42(1), 43 - 75.

US Bureau of Census, (2002). Poverty in the United States, 2000. Washington DC: US

Government Printing Office.

US Department of Education (2000a). Condition of education, 2000. Washington, DC: US

Government Printing Office.

U S Department of Education (2003). *Stronger Accountability*. Retrieved September 22, 2006

from <http://www.ed.gov/nclb/accountability/index.html?src=ov>.

Wikipedia (2006). *High school graduation examinations*. Retrieved December 29,

2006 from http://www.en.wikipedia.org/wiki/High_school_graduation_examination.

