Effects of High-Stakes Testing on Third through Fifth Grade Students: 
Student Voices and Concerns for Educational Leaders

Carol E. Watson  
Matthew Johanson  
Megan Loder  
Julianne Dankiw

Abstract
As the field of education gears up to transition from NCLB to CCSS, the common thread that remains is high-stakes assessment. While many studies report effects on teachers, instruction, and school systems, this investigation draws on personal journals to detail the perspective of elementary students during the testing process. Findings support two primary themes: many students experience a debilitating level of stress manifested both emotionally and physically, and classroom teachers have a significant amount of impact on how students approach and navigate the testing process. Through insights directly from the students’ perspective, this study offers suggestions for educators in supporting positive student attitudes and ideas for subsequent investigations.

High-stakes testing in public schools and the accountability that follows continue to be controversial issues of debate nationwide. In Pennsylvania, the PSSA tests (Pennsylvania System of School Assessments) begin in third grade and are administered near the end of each subsequent grade or course throughout high school (PDE). Even as many states transition from individual state curriculums and No Child Left Behind (NCLB) requirements to the more universal Common Core State Standards (CCSS) and accompanying assessments, the issue of comprehensive assessments with significant consequences for multiple stake-holders remains.

Many effects of the practice of high-stakes assessment are well-documented. Teachers, administrators, parents, and students have all voiced varying responses to the pressure and stress involved in this process (Amrein & Berliner, 2003; Barksdale & Thomas, 2000; Landry, 2006). Few studies in this area, though, focus directly on the students’ perspective in their own voice. The purpose of this study is to explore the nature and extent of the effects of high-stakes testing on third through fifth grade students as expressed directly through their own personal writings.

Literature Review
Rise of High-stakes Assessment

In 2002, No Child Left Behind (NCLB) was signed into law by President George W. Bush. Similar to several legislative predecessors, NCLB’s goal was to implement a system of cumulative testing attached to national funding intended to improve the education system in the United States. With this system in place, teachers, administrators, and schools are held
individually accountable for students’ success and failure based on numeric scores. Although holding educators accountable for their educational impact on students is very appropriate, many factors exist that call into question the fairness, accuracy, and usefulness of this structure of assessment and the resulting accountability.

Stress and anxiety associated with the use of high-stakes testing is a concern for all stake-holders involved based on the consequences attached to student scores. School systems have various guidelines and practices for publishing the results of high stakes assessments, but all are released in some manner and significant consequences applied. In one study, teachers expressed that their teaching morale has been lowered and that they felt “anxious, pressured, guilty, and even embarrassed at times,” due to the accountability of high-stakes testing (Jones, 1999). Some teachers have reached a level of stress which leads them to a “silent surrender” and are moving out of the field of education due to negative experiences with high-stakes testing (Landry, 2006).

**Instruction and Environment**

Numerous findings indicate a significant impact on classroom instruction, student motivation, and curriculum. The type of instructional strategies utilized in classrooms to communicate content have moved from engaging activities involving higher-level thinking toward a focus on more repetitive, rote practices specifically intended to raise student scores on multiple choice tests (Barksdale-Ladd& Thomas, 2000; Sheperd, 2003). Amrein and Berliner (2003) found that high-stakes tests actually decrease student motivation to learn and lead to higher student retention and dropout rates. Content that is actually taught is more narrowly targeted toward what teachers know will be tested, often to the detriment (or even elimination) of a broader, more comprehensive knowledge base (Wright, 2002). Jones (1999) wrote “…science, social studies, and the arts are subjects that are pushed aside and taught only if there is extra time left in the schedule,” (p. 200).

Throughout the school year, the typical classroom is filled with many visual displays such as student work, informational and motivational posters, and colorful decorations. Students become accustomed to the layout of the desks and tables in the classroom, and have a consistent seat in which they feel a kind of ‘comfort zone.’ Many states require that such material is removed for the duration of the testing in order to prevent distraction (Landry, 2006). A classroom is “transformed overnight from a normal, colorful, and comfortable…classroom to an environment reminiscent of a school room scene from Little House on the Prairie,” (p. 33-34). This change from a comfortable, familiar environment can add to the anxiety of testing and may, in fact, be a distraction in itself.

**Students**

Most recently, specific focus has been brought to the negative impact of high-stakes testing on students themselves. A significant amount of research has been conducted on the effects that high-stakes testing has on children’s behavior and morale. Specifically reported are effects such as fear, anxiety, stress, physical illnes, and powerlessness (Triplett & Barksdale, 2005). Studies reporting a negative impact of high-stakes testing have outnumbered studies indicating positive effects nine to one (Buck, 2010).
Students also communicate many forms of stress due to new or past experiences with standardized testing. In a study on children’s perceptions of high-stakes testing, many students reported nervousness as well as other physical ailments associated with anxiety (Triplett et al., 2003). One student stated “…we are just little kids and we shouldn’t oughta probably have to do something like this, but they make us,” (p. 12). Another student from this study expressed anxiety at home prior to a testing day: “I was nervous at night…like when we took it we were supposed to go to bed early, but I couldn’t fall asleep cause I was still thinking about it,” (p. 13).

Cultural Issues

Socio-economic status and ethnic culture are variables that can have enormous effects on each individual student’s performance (Valenzuela, 2000), and consequently, their stress levels. Students attending schools where the majority of the population comes from a background of poverty are often presented with the most intense conditions and are penalized heavily for poor individual and school performance on high-stakes tests (Wiliam, 2010). Multiple investigations also report that many high stakes test questions reflect a significant cultural bias toward white, middle class placing non-mainstream student populations at a higher disadvantage from the start (Booher-Jennings, 2008; Heubert, 2009; Valenzuela, 2000; Wright, 2002). In schools with a high English Language Learner (ELL) population, studies report that the practice of high stakes testing has had long term negative effects for this population. Wright (2002) reports, “The findings reveal that standardized testing has not resulted in higher quality teaching and learning in this school; rather, it has resulted in a narrowed curriculum and harmful effects on both teachers and students” (p. 1). All of these cultural factors contribute to raising the level stress and anxiety as students are pressured to achieve high scores in order to secure benefits for their schools, their teachers, and themselves.

The social impact of high-stakes testing has been the focus of many studies. This form of testing has been said to negate students as “individuals and as cultural beings with distinct experiences, needs, and desires that accompany their differences,” (Valenzuela, 2000). Booher-Jennings (2008) found that students formed a moral hierarchy based on their performance on the state test. Some students who performed well did not want to associate with the students having a difficulty. Even occurrences of bullying have been linked to high-stakes testing. According to Hazel (2010), as teacher stress increased, so did student bullying. An explanation provided for this is that “the testing program led [teachers] to teach in ways that contradicted their ideas of good educational practice,” and the students’ social-emotional development has been ignored (p. 342).

Most investigations into the impact of high stakes testing on students mine data from sources other than the students themselves, relying on the interpretation of observed student behaviors by others. If data is collected directly from the students, they are typically aware that they are involved in a study of some kind and may alter their feedback accordingly demonstrating reactivity or The Hawthorne Effect (Leedy & Ormrod, 2012). In this study design, we reviewed daily journals in which students had previously recorded their thoughts and feelings about high stakes testing as they progressed through the process. The investigation was initiated after the journals were already completed by the students, so at the time they were writing, students were
not actually participating in any type of research. In this way, we were able to mine data directly from the students (through their writing), unaffected by any awareness of participation in data collection. With these issues in mind, the question guiding this investigation is: What is the nature and extent of the effects of high-stakes testing on third through fifth grade students as expressed directly through their own personal writings.

**Methodology**

This study involved the examination of student journals composed during the implementation of the Pennsylvania State Standards Assessment (PSSA) process. The journals were written by third through fifth grade students attending a large elementary school in a large urban school district in Eastern Pennsylvania.

Prior to facilitating the PSSAs, students in the testing grades were given their own personal journal and were invited to utilize the writing process to express their thoughts and feelings as they navigated through the assessment. The idea of having the students scribe in journals before the day’s testing was put into place by the building principal with the thought of giving students an opportunity to relieve some of their stress. A teacher candidate placed in the building for his program field experience shared this practice with his college professor who pointed it out as an example of ‘raw data’ that could be analyzed to discover valuable information about student perceptions of the high-stake assessment experience.

Following completion of the testing process, the teacher candidate asked the principal if the student journals could be borrowed and used to conduct a research study. Permission was granted by the principal to proceed. Work began after approval of the project through the university’s Internal Review Board process. Because all names and identifying information were removed from the student journals and they were completed previous to the initiation and design study, they are designated as ‘archived data’ for which individual consent is not needed. A fundamental aspect of this study, as compared to the many others on the effects of high-stakes testing, is that students had no knowledge that their journals would be used in research. Therefore, the results are minimally, if at all, skewed by any impact of The Hawthorne Effect (Leedy & Ormrod, 2012).

**Context and Participants**

The school is located in an inner city area surrounded by residential and commercial properties. 89.4% of the student population in this building is identified as ‘low income’. According to the district 2011-12 demographic statistics, the student population of 840 is comprised of 14.3% White, 12.5% Black, 69.6% Hispanic, and 2.6% other (Allentown School District, 2012). It is notable that this is not a school that has been historically successful at meeting expected/required academic benchmarks as measured by standardized tests. As of 2011, the building had not made adequate yearly progress (AYP) for three consecutive years, receiving the status of School Improvement II from the state Department of Education.

The journals, designed by the principal, displayed a picture of the school building on the cover along with the title: 2011 PSSA Journal. The students were provided a line to write their name.
On the inner cover is an encouraging letter from the principal to students taking the test. Each of the subsequent pages is titled at the top with the testing day and subject (ex: PSSA Day 1 – Mathematics). Below the title is a supportive quote chosen for that particular day. Suggestions for what the students may write about are provided on each page. Each day’s instructions begin by inviting the students to describe how they were feeling that day. On several pages, the students were encouraged to write about how they were feeling that day compared to the previous day. The activity of journaling was communicated as completely optional for students; only if it would help them manage the testing process more successfully. It was not a mandatory assignment, but intended as a positive, support activity.

A total of 186 student journals compiled the sample that was analyzed by the research team. The student writers were enrolled in grades three, four, and five. All participating students took the required multi-day high-stakes state assessment (PSSA). Several of the journals (between five and ten) were submitted by students mostly blank reinforcing the previously explained guideline that this activity was entirely optional and not a required assignment.

**Data Analysis**

The journals were collected and transported to the university campus where the analysis was conducted. Data analysis commenced during summer 2011 and was conducted by a research team of three undergraduate students and a university professor. The research team met many afternoons for large blocks of time over the next several months. Initial meetings were spent teaching the undergraduate members further basics of qualitative research methodology through direct instruction and interaction with previously published studies. Subsequent meetings were spent carrying out the process of analysis and applying what was learned about qualitative methodology closely guided by the university professor.

Qualitative analysis was determined to be the best way to extract data from the significant amount of text written by the elementary students (Merriam, 2001). Specifically, constant-comparative methodology was employed to address the stated guiding question (Strauss & Corbin, 1990). The team of researchers sat around the same table reading through stacks of journals; sometimes reading silently and sometimes aloud to share ideas with each other. As common themes began to emerge, they were noted and identified using color-coded post-it notes, each color representing a particular theme. The categories remained dynamic throughout the process as new student thoughts and feelings were included and themes were rearranged based on new information and changing perspectives.

**Findings**

**Student Struggles**

Negative feelings were expressed through multiple descriptors throughout student writing. Students who reported anxiety and/or nervousness typically took the next step of attributing their feelings wholly to PSSA testing. One student wrote, “I have butterflies because I am nervous about the test…” Some of the students who wrote about their anxiety were concerned about the unknown material in the test. Students in third grade were experiencing high stakes testing at this level for the first time, so there was fear and anxiety of the unknown. “And first I felt
nervios because I Don’t no how long it is and how Big it is,” said one student. According to another student, “Everyone is frightend and Nervous.”

Use of the word ‘hate’ was a consistent theme throughout the journals. Generally, the word hate was clearly directed toward the PSSA test. A simple, repeating sentence expressed by many students was, “I hate PSSAs.” Some students went as far as writing this statement, or a similar one, on every daily journal entry. Occasionally, students used the word hate to voice their dislike of specific content areas. While reviewing the journals with the word hate, math emerged as the subject with the strongest association. One comment declared, “Nervous, bumed, I hate math I can’t stand it.” All students who wrote the word hate expressed a negative attitude toward PSSA testing specifically. A quote that summarizes the prevailing perception communicated through the student journals that used the word hate is, “…I hate the PSSAs so bad I never want to take it again.”

A plethora of physical ailments were described in many students’ writing. The range of symptoms included various types of physical pain, exhaustion, and nausea. Exhaustion or sleepiness was the most prevalent ailment expressed by students. Prior to even beginning the day’s testing, many students reported that they were sleepy. Others described how certain subjects made them tired; “But on reading, I get sleepy...” Symptoms of nausea and physical pain were reported by a significant number of students. On the first day of testing, one student reported, “I feel like I’m about to trow up.” A typical report of physical pain was expressed in another journal: “Back pain chest pain migran…”

Many journal entries indicated that students were stressing over the significant length of time the tests require. In the beginning of the 3 week testing window, many expressed feelings of confidence and even contentment. As time and the testing process continued, however, their feelings began to change. The longer the tests went on, the more students began to experience the physical symptoms mentioned earlier such as exhaustion, nausea, and pain. It was obviously hard for many students to sit and take a test for 3+ hours and remain focused. A few of the students just wanted to go to sleep because they felt so tired. Following many days of testing, one student wrote, “I don’t want to take any more tests and I want to sleep.” Towards the end of the three-week testing period, students were relieved and exhilarated that the tests were over. In the beginning, students anticipated getting Advanced or Proficient on the PSSA’s. But towards the end, their only goal was to finish the test.

Growing Confidence

Although much of the findings reflect a growing negative feeling toward the testing across the process, there were many students who were satisfied and happy taking the PSSAs. “I feel cool about the test I think it is awesome,” wrote a student who communicated positive statements on every journal entry.

Mixed emotions emerged as a common theme very early in the analysis process. Students initially reflected feelings such as fear, nervousness, and worry. Some of their expressed reasoning behind these feelings were fear of getting questions incorrect and failing. Others mentioned that it was their first day of testing and they did not know what to expect. However,
after the first or second day, many entries began to take on a different tone. For example, one student reported, "I feel fine because I took the tests two times already and I am thinking I will do fine on this third one and I am so proud of myself because i did my best and want to pass the test." As the testing progressed, many students who were originally uncomfortable with the testing process adjusted and displayed more positive and confident statements. One student who reported that they felt as though they were going to throw up on the first day appeared to be much more comfortable with the testing process later reporting “good day” and “awsome day” on later journal entries.

**Drawings and Diagrams**

Many students chose to convey their personal specific days through drawings; sometimes with accompanying text and sometimes without. Expressive pictures illustrated many diverse pertaining to the PSSA’s including confidence, fear, nervousness, and excitement. A positive communicated in one drawing of a girl sitting at with a pencil in her hand, the PSSA booklet in and a big smile on her face. Another communicated confidence by student with a wand in hand and the will slay you" coming out of his mouth looking at the PSSA test. Other indicated more active thoughts. One buildings with a boy pushing the off the edge into the water illustrating could defeat the test. Another student feelings the day-to-process. This flowcharts two boxes for containing a single word with an arrow pointing to both boxes. On day 4, the words in the boxes were nervous to confident. On day 5, they read confident to ready. On day 6, they were ready to easy. This display actually demonstrates both of the previously discussed themes of using diagrams to communicate feelings and the progression from anxiety toward confidence during the testing process.
It is significant that the majority of the feelings communicated through these nonverbal modes were positive and confident. This direction is opposed to a much more negative message overall communicated through most of the verbal communication.

**Distinction Among Classrooms**

As the previously discussed themes emerged, we began to recognize that some teachers had played a significant role in student attitudes toward the testing process. Several teachers seem to have had a significant impact on the demeanor with which students approached and navigated the testing process. Some students in specific classroom groups discussed various feelings towards the testing using common language or similar references. Statements such as “I feel happy today and ready for the test” and “I will do great today I know it” were frequent answers from many students in one particular group of journals. It became more evident that different classrooms of students had a varying collective demeanor toward the PSSA test which we attribute to the teacher’s influence. This is evidenced through group commonalities of word phrases as well as attitudes reflected in many of the student comments.

One of the classes showed a distinct pattern of assertiveness toward academic success. The majority of the students felt nervous about the test, however, those same students also felt confident in themselves, and stated so specifically. Students reported working towards Advanced (the highest ranking) stating, “I feel excited because I feel like I already have Advanced” and “I am nervous. Scared of the PSSA’s. But I wanna be advanced.” Another student predicted similarly on behalf of the whole class group, “I feel a little nervous about the PSSA and I also feel ready to take it and I feel that I can get advanced and everybody else can!” Student entries from another class group used common vocabulary consistently across individuals. The words confident, proud, and believe in myself were repeated in almost every journal in the group.

Several class groups were consistently and noticeably not confident or positive. Their comments were also consistent across individual students within the classroom group. All of these trends seemed to be class group-specific.

One unexpected finding was evident not in the student writing, but in teacher writing in the student journals. One particular teacher responded to each individual student’s daily entry with a personal message almost every day of testing. This consistent practice transformed the student journals into an interactive conversation between student and teacher. She expressed pride in the students’ hard work and attitudes toward the test, confidence that they would do well, and encouragement to stay calm and do their best. Some notes offer helpful tips in response to specific student concerns expressed in a previous entry such as, “Sometimes the questions might be tricky, but if you concentrate, I know you’ll be able to figure it out! Never give up!”, and “If you get sleepy, try to sit up nice and tall in your chair”. To another student who expressed that he was ‘getting mad at 2 answers’ the previous day, she wrote, “It’s OK to get stuck on a few questions—as long as you take your time and go back to try to figure them out”. Other comments seemed to be in response to teacher observations from previous days such as, “Slow down today and really take your time. You can do it!” Early in the testing, one student had written, “I don’t feel good I don’t think I am going to pass the pssa I think I am going to get held back in 4th
grade.” Near the end of the testing, this same student wrote, “Thank you for telling me I can do it now I believ I can do it”. Almost every student in this classroom at some point in their journal thanked their teacher for her notes and stated the positive impact it had on them as they faced the tests.

**Discussion**

Findings from this study yielded significant evidence of both expected (as indicated from previous research) and unexpected outcomes. Emergent themes were also similarly stratified with both negative and positive trends reflected. The overall indication is one of very strong student feelings and perspectives, regardless of which end of the spectrum they fell. This practice of high-stakes testing evokes extremely strong emotions in elementary students that are sometimes recognized, and sometimes not. During the rush of adult/parent/educator stress and responsibility involved in preparing for and facilitating high-stakes assessments in schools, the view from a student’s perspective may become lost, forgotten, or just not valued. Teachers and principals need to remember that students are experiencing significant and sometimes, debilitating stress as well. It is important for this phenomenon to be recognized, acknowledged, and addressed for the sake of both student health and student success. The perspective of students cannot be ignored or even minimized if we are to move forward toward success for all stakeholders.

In response to this issue, one of the most useful patterns that emerged in this study is that various classroom groups had distinctly differing collective demeanors toward the testing process. The most likely reason for this is the influence and attitude modeling of the teacher. The common vocabulary and language utilized by students within classroom groups, the cohesive negative or positive group attitudes toward success, and the dynamic of changing attitudes from beginning to end of the testing process all indicate teacher influence on the class as a whole. This evidence lends credibility to previous studies suggesting that teachers and other educators can make a difference in the level of confidence with which students approach testing (Supon, 2008; Triplett & Barksdale, 2005).

Several implications for educational leaders at both the building and district levels are raised by these findings as well. As many states embark on newly implemented structures of merit teacher pay largely based on student test scores, it is likely that educator stress over high-stakes testing will continue to elevate. By sharing these findings on the potential impact teachers can have on student attitudes during testing and offering specific suggestions for modeling, building principals and district leaders can empower teachers toward positive practice which may, in turn, create a context in which students feel empowered and confident. We can deduce from some of the specific statements students communicated that building/district practice may need to move away from a paradigm of constant reminders related to the tests to one of more calmness and confidence. Finally, the whole idea of creating and implementing the student PSSA journal to validate student concerns and aid in their positive navigation through the testing process was originally the idea of the building principal. This practice is easily duplicated both to support students and to find out more about what they are thinking and feeling. This information could easily lead to more ideas and practices to support both teachers and students through a successful testing experience.
Further research on the specifics of practices facilitated by teachers whose students, as a group, approach testing with confidence and a positive attitude may provide useful strategies for other teachers to duplicate in the future. Also of benefit would be to conduct a study similar to this one correlating levels of success based on test scores to determine if positive student attitudes (individually or as a class group) produce higher academic success.

We would like to close this article with some things to keep in mind for teachers and administrators gleaned specifically from this investigation’s findings:

1) Teachers need to be acutely aware of how high-stakes tests are affecting students, actively recognize and acknowledge their stress, and be ready to offer both individual and group strategies for students to manage it.

2) Verbal and written support expressing confidence and encouragement from the classroom teacher is significantly effective if it is specific, individual, and daily.

3) Students take their cue/attitude from the teacher – if a teacher conveys confidence in their success, they are more likely to adopt that perspective as well. Although teachers are not able to remove the stress from students during this mandated process, they may have much more power than they realize to instill self-confidence and to teach students how to develop it for themselves.

4) School leaders determine the atmosphere in any building or district. Armed with research-based information such as this, they have the potential and the responsibility to implement practices that support healthy and confident attitudes toward mandated high-stakes testing practices.

References


Heubert, J. D. (2009). High-stakes testing: Opportunities and risks for students of color, English language learners, and students with disabilities.


