

Teacher Attitudes About Standardized Testing and Test Preparation: A Case Study

By

Christopher R. Klager

Submitted to the graduate degree program in Curriculum & Teaching and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science in Education.

Chairperson Dr. Jim Ellis

Dr. Mary Lynn Hamilton

Dr. Phil McKnight

Date Defended: 6/19/13

The Thesis Committee for Christopher R. Klager
certifies that this is the approved version of the following thesis:

Teacher Attitudes About Standardized Testing and Test Preparation: A Case Study

Chairperson Dr. Jim Ellis

Date approved: 6/19/13

Abstract

This study looks at how science teachers' attitudes about standardized testing affect the types of test preparation practice they choose. Interviews, both before and after state testing, classroom observations during test preparation, and documents from test preparation contributed evidence used to create case studies about two teachers. Teacher attitudes related to pressure they felt to obtain high scores, test preparation ethics, and inexperience with state testing contributed to the decisions the teachers made about test preparation. Education for teachers about what test preparation practices are appropriate and the format of the test could help other teachers make informed decisions about test preparation practices.

Chapter 1: Problem Statement

Introduction

Since President Bush signed the No Child Left Behind Act of 2001 (NCLB) into law, a tremendous amount of pressure has been placed on schools to make Adequate Yearly Progress (AYP) in their standardized test scores. Teachers, in particular, have felt this pressure mount as the percentages of students that must pass the tests each year has increased or as their schools face the repercussions of missing AYP goals.

Recently the Obama administration has provided some flexibility in the form of waivers and more options for creating school improvement plans. In the summer of 2012 the state of Kansas received a waiver from the US Department of Education. The waiver requires the state to implement a number of changes regarding college and career readiness, recognition and accountability, and support for effective instruction and leadership.

To meet the objective of making sure students are college and career ready the state agreed to adopt common core standards by 2014, and to implement new high quality assessments. The state has also adopted new English language proficiency standards. To meet the standard for accountability the state agreed to look at improvement in scores, reduction in achievement gaps, and reduction in non-proficient learners rather than schools having to meet a certain cut score as was required before. Rather than looking simply at the percentage of students who score proficient on the state assessment, the state will award points based on how

much students improve and what levels of achievement students reach on the assessments. The state calls this the Assessment Performance Index (API), and they believe that it is a more effective measure of how schools are doing than what was in place before (KSDE, 2012).

To support effective instruction and leadership the state has agreed to implement a new evaluation system for teachers and administrators. These changes will be implemented by the 2014-2015 school year (KSDE, 2012). The focus on test scores still remains, though, because they are still an important part of the recognition and accountability requirements of the waiver, and because they are used as an important indicator of school success by school officials and by the public.

As a high school teacher I have been able to witness the pressures associated with testing play out in an environment where making AYP was not assured. Each year, as the state assessment occurs, I hear the complaints of teachers in my building. The attitudes of those teachers are not all the same, though. A few teachers seem completely unbothered by the whole process, but the majority of teachers seem to have quite negative attitudes. There are a number of factors that play into how individuals feel about testing. My school also has enacted some programs that are controversial among teachers because they place new demands on us, chiefly a seventeen minute block of time, added onto 4th hour each day, dubbed "Skills Time." "Skills Time" is meant for the review of state assessed reading and math indicators, but it has come up several times in faculty meetings that a significant number of teachers are not participating.

Despite the importance that our administrators place on these standardized tests, there are no set procedures for how we are to prepare our individual classes for them. Each teacher, perhaps in consultation with his or her department, chooses how and when they will prepare their students for these standardized tests. I know that my students do not like to take standardized tests in any form. Any time I mention the state physical science exam that my students have to take, they moan and groan about having to take “another test.” They have been taking them for virtually their entire tenure in school, and by the time they enter my classroom they are sick of them. Hoffman and Nottis (2008) found that student motivation and feelings about standardized testing varied widely depending on the situation. They found that extrinsic rewards were the primary motivating factor for most students, but that some teachers were able to develop feelings of self-efficacy in their students, which provided motivation when it came time to test.

It thus becomes extremely important for me to 'sell' the tests to my students, so that they will take both the preparation for, and the actual tests, seriously. Fedore (2006) provides the examples of holding school-wide assemblies and hanging posters around the school to encourage students and demonstrate the importance of the tests. Privileges are also sometimes used as a motivator for students (Croft, Waltman, Middleton, & Stevenson, 2005). I would like, though, to help my students believe in themselves and want to show off their learning on the test without relying on extrinsic motivators. A major part of developing these feelings has to do with how I present the tests to my students and what we do to prepare for them. The choices I make about test preparation are important not only

because I want the methods I choose to be effective, but because I need my students to take them seriously.

Purpose

This proposed study seeks to find out how teachers view high-stakes standardized tests, and how those views might play into how they approach preparing their students for those tests. This is an important topic to explore because it could provide information about how teachers are unconsciously allowing their attitudes to affect their students' performance on state standardized tests.

Discontent with standardized testing is widespread among teachers (Abrams, Pedulla, & Madaus, 2003; McMillan, Myran, & Workman, 1999; Mertler, 2011; Sunderman, Tracey, Kim, & Orfield, 2004). When a teacher is the one mandating tests, as they do for most of the year as they go about their normal teaching routine, they may treat the process quite differently than when they have to prepare their students for tests mandated by someone else. Many schools are struggling with making AYP. In Kansas, 15.7% of schools and 26.7% of school districts failed to make AYP in 2011 (KSDE, 2011). Data for 2012 is not available. Since the standardized tests mandated by NCLB are of extreme importance to the schools and school districts that take them, it is important for teachers to be aware of how they might have an impact on student test scores and on meeting AYP goals.

Conceptual Rationale

A review of the literature on the topic of high-stakes assessment preparation practices revealed that there was some variance in the type and amount of

preparation done by individual teachers. Teachers were often faced with conflicting ideas about the purpose of their instruction in the classroom. Teachers have to deal with their own feelings about what their professional and moral responsibility is, while also facing the implications of potentially low scores on high-stakes assessments. I think it seems likely that the balancing act teachers must play between increasing student knowledge and increasing standardized test scores affects the methods teachers choose for standardized test preparation.

Throughout the literature there seemed to be no consensus on what constituted “best” practice in test preparation. This may be due, in part, to concerns about test validity and the ethics of test preparation. Certainly there are practices that will raise standardized test scores quite effectively, but an increase in score may not be all a teacher is after. Standardized tests are nothing new, and a lucrative industry has emerged related to effectively preparing test takers. Someone getting ready to take a common standardized assessment such as the SAT or GRE is likely not so concerned with increasing their knowledge as they are with increasing their score, though. For a teacher, the conflicting demands of increasing knowledge and increasing test scores may bring rise to some interesting decisions about what they do in their classroom.

Research Questions

The purpose of this study is to examine the attitudes that science teachers hold about the Kansas State Science Assessment, and how those attitudes might affect the test preparation practices they employ in the days and weeks leading up to the assessment. Teacher attitudes about standardized testing, and how those

attitudes may relate to the test preparation practices those teachers use, are of particular interest. As a result, it will be important to collect data on the following:
1) What are teachers' attitudes in regard to standardized tests? 2) Did the attitudes of the teacher play a role in how they prepared their students for the standardized test?

Definitions

Attitudes – In this study, attitudes specifically refers to not just whether teachers have a positive or negative outlook on mandated standardized testing, but how important teachers feel the tests are. I focus on how useful teachers think the tests are and whether teachers feel like the test is a good use of their time.

High-Stakes Assessment – High-stakes assessments are tests where the results are used to make important decisions in a school or school district (Au, 2007). This potentially includes school finance, teacher salary and retention, and student grade retention.

Chapter 2: Literature Review

History of Standards and Standardized Testing

The origin of standardized testing in schools in the United States dates back to the mid-nineteenth century. In 1845, Horace Mann sought to adopt a common exam to give to all Boston schoolchildren so that teaching and learning could be compared between schools. In the decades that followed, tests were developed to sort and classify students based on intelligence, to test instructional effectiveness, and to determine the talents of individual students (Gallagher, 2003).

In 1925 the SAT was developed as a college entrance exam, and in 1959 the ACT was developed as an alternative to the SAT. Both of these exams are widely used in college admissions even decades later (Alon and Tienda, 2005). The college admission test preparation industry has developed into a very lucrative business because of what a large part of the education system it has become (Buchmann, Condron, & Roscigno, 2010).

Congress enacted the Elementary and Secondary Education Act (Title 1) in 1965 with the major aim of reducing the achievement gap between students by providing extra funds to schools. To qualify for funding, though, schools had to demonstrate that they qualified via standardized test results. This was the beginning of widespread use of standardized testing to determine the effectiveness of particular types of instruction (Gallagher, 2003). Congress has reauthorized the Elementary and Secondary Education Act every five years since its inception,

meaning that standardized testing has remained an important part of how schools are funded.

In 1969 the National Assessment of Educational Progress (NAEP) was created to compare students in different states. Each year a sample of students in each state are tested to determine how they compare to each other. The NAEP is now designed so that the results are “linked” to international assessments such as the Trends in International Math and Science Study (TIMSS), thus allowing individual state results to be compared with international results (National Center for Education Statistics, 2011). Today the NAEP is especially important because it is how different states are compared in President Obama’s Race to the Top program (US Department of Education, 2009).

In 1983, the famous *A Nation at Risk* report was released, calling for a number of changes in education lest the United States fall behind competing countries in commerce, science, and technology (National Commission on Excellence in Education, 1983). The report cited falling test scores and lack of preparation for the workforce as indications that the education system in the US was failing. The report contained thirty-eight recommendations for the reform of K-12 as well as college education. These recommendations were divided into five categories: Content, Standards and Expectations, Time, Teaching, and Leadership and Fiscal Support. In summary, the recommendations called for schools to cover more rigorous content, to use standardized tests to measure the level of student knowledge, and to provide a tool for administering interventions to students who needed remedial or more advanced work. *A Nation at Risk* also made note of

America's low international ranking in science and math education, with many of the recommendations targeted toward improving that ranking.

Since then, a great focus has been placed on the development of standards by which students can be measured. Under President George H. W. Bush a report entitled *America 2000: An educational strategy* (US Department of Education, 1991) called for "standards", as did a law passed during the Clinton administration entitled *Goals 2000: Educate America Act* (US Congress, 1994) which called on states to develop "rigorous standards." Another key goal of *Goals 2000* was that the United States was to be ranked first in the world in science and math achievement.

In 2001, what had been recommendations or goals were replaced by new standards and penalties for schools that failed to meet them (Sadovnik, 2006). Since the passage of the No Child Left Behind Act of 2001 (US Congress, 2001) schools and teachers have had to cope with the pressure of meeting standards or facing severe consequences. If a school fails to meet AYP for two consecutive years it is required to implement a two-year improvement plan. Students at that school also have the option to attend other schools with Title I funds being used to pay for their transportation to the new school. If the schools continues to not make AYP, they must allocate more Title I funds to programs to help underperforming students. After this, schools are faced with implementing new curriculum, replacing staff, increased state oversight, or being turned into a charter school under private management (Wageman & Bartoszek, 2004).

A major part of No Child Left Behind (NCLB) is a shift toward a performance model of pedagogic discourse (Burnstein, 2000), which in this case means a shift to

the use of standardized tests to measure students' knowledge of content standards (Barrett, 2009). Even though standardized testing is nothing new in schools, Barrett points out that the focus of teachers and students had never been to simply pass a test like it is under NCLB.

The Obama Administration has made some alterations to how federal funds are to be disseminated to the states for education. Standardized tests are still an important part of how progress is to be measured. While NCLB is still in effect, the Obama Administration has granted waivers to 32 states, including Kansas (Foster, 2012). This means that those states have demonstrated that they have plans in place to improve schools without having to demonstrate that all students are proficient by 2014, as NCLB requires.

One of the requirements for obtaining a waiver is that Kansas has agreed to the adoption of new college and career readiness standards which the US Department of Education emphasized in its reauthorization of NCLB in 2010 (US Department of Education, 2010). Forty-five states have adopted the Common Core Standards in reading and math. The National Governors Association and the Council of Chief State School Officers developed these standards to provide consistent and high standards for college and career readiness (Common Core State Standards Initiative, 2012). One of the criticisms of state established standards was that they varied in scope from state to state and that some states lowered their standards once it became more important for them to demonstrate proficiency.

Assessments are being developed based on the new common standards although they will not be ready for implementation for several years. The Obama

Administration's aim is to increase the quality of standardized tests, although they still will be an important tool for assessing the effectiveness of schools (US Department of Education, 2010).

Another important development during the Obama Administration is the implementation of the Race to the Top (RTTT) program. This program uses money made available through the 2009 economic stimulus plan to provide incentives for states to implement changes in education. States compete for over 4 billion dollars in federal funds by demonstrating their compliance in the following categories: state success factors, standards and assessment, data systems to support instruction, great teachers and leaders, and turning around the lowest achieving schools (US Department of Education, 2009).

In this program a great deal of emphasis is placed on standardized assessments. Two of the criteria for showing state success factors are that scores improve in reading and math and that scores show a shrinking of achievement gaps. The standards and assessment category requires states to adopt widely used standards, like the common core standards, and to implement a new generation of assessments based on those standards. To show that a state has great teachers and leaders they are required to implement teacher evaluation systems that use student growth, presumably measured by standardized tests, as an indicator of teacher effectiveness (US Department of Education, 2009). Race to the Top moves away from penalizing schools for poor performance on tests to rewarding states for improvement. The importance of standardized tests, though, remains extremely high.

Effects of NCLB

There have been numerous studies conducted on how NCLB has changed schools, but whether those changes have been positive or negative overall is still a matter of contention (Abrams et al., 2003). Maleyko and Gawlik (2011) note that there is a consensus in the field of education that the philosophical intent of NCLB is good. It is difficult to argue with the idea that “all children will count and no child will be left behind. (p.614)”

Dee and Jacob (2011) point out one of the underlying assumptions of the move from low-stakes tests like NAEP to high-stakes tests like those under NCLB is that if the public has information about the quality of schools, the productivity and focus of those schools will improve. Because of the possibility of sanctions and the fact that those sanctions are determined by the results of these high-stakes assessments, however, school behavior has changed in ways that may be detrimental to students.

One of the changes that educators fear as a result of a move towards high-stakes tests is that teachers will “teach to the test.” Teachers have reported changing the content they cover in their classrooms to focus more closely on standards when standards are implemented in their schools (Jones et al., 1999). Jones et al. also note that instructional strategies sometimes change to meet requirements in the standards (1999). Guilfoyle (2006) makes the point that while the US Department of Education claims that tests will help teachers improve their teaching methods by diagnosing student weaknesses, the tests actually only

measure a very narrow list of subjects and teachers are better off focusing solely on what will appear on the tests.

Dee and Jacob (2011) show that there have been gains in math scores since the inception of high-stakes testing, but that there has been only minimal headway on closing achievement gaps. Math scores may be going up, but if teachers are feeling forced to narrow their focus to topics that only appear on the tests, are the increase in test scores worth the cost? Malyko and Gawlik (2011) write that basic skills are being favored over high level thinking skills and innovation, and Firestone et al (2002) note that a great deal of focus is placed on memorization of facts. The recent move towards the common core standards may give teachers reason to focus on high level thinking skills, but it is yet to be seen what kind of assessments are developed to go along with these new standards. Regardless, Dee and Jacob (2011) also note that there has been a shift in focus to reading and math and away from other subjects like social studies, art, and music.

Fuller, Wright, Gesicki, and Kang (2007) note that some state standards even have been lowered to, presumably, make it easier for schools to show that their students are meeting standards. Harris (2007) showed that different states can make students look much better by manipulating the cut scores they use to signify whether a student is achieving at standard or not. There also are cases of individual teachers or school districts manipulating tests to make sure that students meet standards (Jacob & Levitt, 2003). Thirty-five school district employees recently were indicted in Atlanta on charges that they physically changed answers on tests and then covered it up (Strauss, 2013). The National Center for Fair and Open

Testing (2012) claims that cheating or test manipulation have occurred in 37 states and Washington DC in the past four years. Firestone et al. (2002) note that advocates of testing claim that high-stakes testing is necessary to hold teachers and students accountable by using test scores as the grounds for providing rewards or levying sanctions. Others (Darling-Hammond, 2007a; Maleyko & Gawlik, 2011), however, point out that having a single-measure accountability system is unreliable and can lead to manipulation of data. The recent cheating scandal in Atlanta may be an example of how the pressure to perform on such important tests can lead to unethical behavior.

The implementation of high-stakes standardized testing also has caused some teachers to focus a great deal of attention on “bubble kids,” trying to make sure that the students who are close to passing make it there when they test (Murnane & Papay, 2010; Neal & Schanzenbach, 2010). When schools place the focus on trying to get as many students to meet standards as possible, they are tempted to move away from the original intent of high-stakes testing, that all students matter, to use their resources to make themselves look as good as possible.

Jacob (2005) provides evidence of teachers encouraging low-achieving students to skip high-stakes exams in an effort to raise scores for their school. Another important issue that high-stakes testing has brought up is SPED classifications. Jacob also points out that another tactic schools will use to try to increase their test scores is to classify low-achieving students as needing special education services. This allows those students to take a modified version of state tests and potentially increase a school’s scores.

Teacher Attitudes About Standardized Testing

Teachers certainly are not immune to having their own opinions on testing. In fact, numerous studies have been conducted looking at how teachers view high-stakes standardized testing. The opinions of individual teachers cover a wide range, but there are some trends that can be observed. In their study on how teachers view NCLB, Murnane and Papay (2010) found that, in general, teachers seem to support the intent of NCLB in terms of holding schools accountable for the learning of their students. Overall, though, teachers felt that the incentives created by the law have caused a decrease in the quality of education for advanced students because of the focus placed on standards that they already meet. Other studies have concluded that teachers have an overall unfavorable view of NCLB and the effects it has had on education such as the focus on specific test content and increases in pressure teachers feel (Mertler, 2011).

One effect of NCLB is on how teachers actually teach. Teachers have expressed concerns about feeling they have to “teach to the test” which has affected their curricular choices (Darling-Hammond, 2007b). Even if teachers do not go so far as to teach to the test, it is common for teachers to express that they feel like the tests change what they are able to teach. Murnane and Papay (2010) note that teachers tend to like that the increased focus on standardized testing has set clear expectations for what their students should know. In contrast, Jones et al. (1999) point out that the advent of high-stakes testing in North Carolina led teachers to feel like they had lost the ability to be creative planners and thinkers. In Barrett’s

(2009) study teachers expressed that they had amended what they teach so that it conformed better with the state tests, that they felt they had a lack of time to cover all of the content they wanted to cover, and that they did not feel like they could take risks or try new things in their instruction.

Many studies have looked at teacher morale since high-stakes testing became so important. Byrd-Blake et al (2010) say that increased pressure on teachers to obtain high scores has increased stress and lowered teacher morale. Mertler (2011) goes further to say that students also feel the effects of this increase in pressure and their morale also can drop. Teachers realize that a lot of decisions are out of their control, but they don't feel comfortable with the results.

Another major form of stress for teachers is the publication of scores. Barrett (2009) found a major source of stress for some teachers was that their scores would be posted in the school for other staff members to see. This offers the potential to embarrass an individual teacher, but others have noted that teachers also worry about how the school will be perceived in the community. Jones et al (1999) noted that teachers sometimes feel embarrassment and guilt when their school's poor scores are published in local newspapers. Even in schools where scores traditionally are high, teachers expressed a lot of pressure to keep those scores high. Teachers also have reported high levels of anxiety because they felt that a team of administrators might walk into their room at any time to make sure they were implementing policies enacted to raise test scores (Valli & Buese, 2007). The stress and pressure that teachers feel comes from a variety of sources, both in the school and in the community.

In contrast, Rex and Nelson (2004) found that some teachers relied on their own professional judgment and ethical sense when confronted with the challenges of standardized testing, and opted not to let it affect their practice. Rex and Nelson (2004) conclude that testing has an influence in how teachers teach, but the extent can vary quite widely. For example, one of the English teachers they spoke to said that he teaches “real writing” and that he is confident that by doing that he will adequately prepare his students for the high-stakes exams they will take later on.

Some studies show that teachers feel like high-stakes tests simply are not a fair way to measure student achievement. Grant (2000) gives the example of how teachers are uncertain of what tests will look like and how they will be scored. This uncertainty could cause teachers to question the purpose and usefulness of the tests they must prepare for. Some teachers reject the idea that high-stakes tests will help them focus and improve instruction (Mertler, 2011) despite Murnane and Papay’s (2010) assertion that one of the key positives of high-stakes testing is that it will do just that. Jones and Egley (2004) wrote that teachers expressed concerns that data from testing was being used improperly and that a one-time test was not a fair measure for evaluating teachers or students.

Lattimore (2005), in a study of African-American students taking a high-stakes math exam, points out that it is discouraging for students when they hear their teachers say negative things about testing and test preparation. Jones and Egley (2004) note that student motivation can be impacted by the stress and pressure teachers place on them to do well on high-stakes tests. Clearly, school employees are not the only ones affected by the implications of high-stakes testing,

but teachers can play a crucial role in how students feel when they sit down to take a high-stakes test.

Test Preparation Practices

Besides changes to what subjects are taught in schools, the focus on standardized testing also has had an effect on the amount of time and practices used to prepare students for tests. Diamond (2007) provides data to suggest that test prep is extremely widespread among teachers. In a survey of pre-service teachers in New York, Barrett (2009) found that one hundred percent of them planned to use practice tests in their classrooms to help prepare students for state assessments. Some teachers have reported that they feel like all they are doing is preparing students for tests (Shepard, 2000).

There are some concerns about teachers who prepare their students for these assessments. Since the data that is obtained through these tests is used as a measurement of student learning, it is important that the information from that data is valid. However, Perlman (2003) raises some doubts as to whether or not increases in scores are really indicative of student learning. If students have been coached on a narrow range of subjects that are to appear on the exams and a great deal of time has been spent on learning test taking strategies, then the gains in scores may be artificial. The students may actually know less than if they spent their time learning the tested content via regular classroom instruction. Along these same lines, Barrett (2009) expresses the concern that material is being presented out of context just for the purpose of preparing for an exam. Au (2007) writes about

“fragmentation” where content is delivered in small, isolated chunks simply so that it can be used to pass a test.

One of the other consequences of test preparation is that it changes what and who receive attention from teachers. Croft et al (2005) report that teachers spend an increasing amount of time on tested subjects before exams, and that teachers pay an increased amount of attention to specific subgroups. Diamond (2007) refers to teachers who reallocate time and resources to tested subjects, and ignore untested content until after testing is finished. There are recorded cases of administrators telling teachers not to allow their students to study non-tested subjects until after high-stakes assessments had been administered (Hursh, 2008). Jones and Egley (2004) report that teachers do less hands-on work in science and shift to more workbook and textbook work as high-stakes tests approach.

Rather than focusing on content, teachers also spend an increased amount of time covering test-taking skills and giving practice examinations that have the exact same form as the high-stakes tests they are preparing students to take (Olson, 2001). This is done with the intention of increasing test scores, but teachers also have expressed concerns about students not learning how to apply their knowledge to higher-order thinking tasks (Murnane & Papay, 2010). A national survey of teachers found that 40 percent of respondents “reported that they had found ways to raise state test scores without really improving learning” (Abrams, 2004, p.8). This certainly could be problematic because students may be unable to actually use any of the knowledge they have gained outside of the context of the test.

There are a number of test preparation practices that appear to be commonly used by teachers. Teaching test-taking strategies, mirroring the format of high-stakes tests on in class exams, and using sample items from old high-stakes tests are common (Croft et al., 2005; Mertler, 2011; Olson, 2001; Rex & Nelson, 2004). Teachers also will review old content with students as tests approach, use complete versions of tests from previous years, and integrate test review throughout the entire school year (Croft et al., 2005). Lai and Waltman (2008) found that there was no difference between the practices chosen by teachers at high-achieving, mid-level, and low-achieving schools.

Some other common practices in the time leading up to high-stakes testing are holding assemblies to entertain and motivate students, providing breakfast or snacks on the day of assessments, and putting up posters around the school to encourage kids (Fedore, 2006). Teachers also lead their students in stretching before examinations, and they will provide privileges for students as rewards for doing well on tests (Croft et al., 2005). Hoffman and Nottis (2008) found that these extrinsic rewards were effective at motivating students in some cases, but those teachers who preached self-efficacy motivated students quite effectively.

The question of what exactly is motivating teachers to spend time preparing students for tests at the potential cost of being able to acquire and apply knowledge to a variety of other situations is an important one. One major factor is administrative pressure from within the school or school district. Firestone et al. (2002) and Hursch (2008) note the importance of the principal's role and influence on how much test preparation teachers do. District professional development, in

some cases, also was a major source for information on how teachers were expected to prepare students for high-stakes assessments. Grant (2000) states that teachers often collaborate with their peers in coming up with a plan for test preparation. Overall, though, the literature seems to indicate that many of the decisions about what practices to use are often left up to individual teachers.

The pressure associated with high-stakes testing can lead teachers to use unethical practices in preparing students for tests, although it is sometimes difficult to pin down exactly which practices are ethical and which are not (Croft et al., 2005). A teacher's beliefs about their professional responsibilities also may play a role in the kind of practices they use to prepare students for exams (Rex & Nelson, 2004). Lai and Waltman (2008) conducted a survey in Iowa and found that teachers largely were unconcerned with the ethicality of the test preparation practices they used unless those practices clearly were unethical. Despite this, the researchers found that roughly ten percent of the teachers used questions from previous tests that were extremely likely to appear, in the exact same form, on the test for which they were preparing. Mertler (2011) provides some more examples of unethical practices that appear to be used quite frequently. Some teachers will acquire current test questions and use those as they review with students. Teachers also will make copies of old tests when the administrators of the tests strictly prohibit it.

Some researchers attempt to offer suggestions on what kind of test preparation practices can provide students with ample knowledge to do well on tests without being unethical or making teachers feel inadequate. Longo (2010) suggests that rather than spoon-feeding students information that may appear on

high-stakes tests, teachers instead should use inquiry-based instruction in science classes to teach the same content. Rex and Nelson (2004) wrote about an English teacher that taught “real writing” and was confident that his students would learn all they needed to know for the state assessment along the way. Firestone et al (2002) give the example of social studies assessments that are designed with open-ended questions, rather than multiple-choice questions, which provide more of an opportunity for teachers to prepare students without being limited to simple recall of facts.

Summary

Standardized assessments have taken on a much more important role in education in the last decade, since the advent of No Child Left Behind. The stakes are much higher for teachers, schools, and school districts now than they were in the past. There are mixed feelings about these tests among education professionals, but one thing that cannot be denied is that they have changed the way that teachers teach. In many cases curriculum has been narrowed and more time is spent preparing students for mandated assessments. Teachers have had to adapt to these changes and have adopted a variety of tactics for preparing their students for high-stakes exams. Many teachers still are trying to find the best ways to live up to their obligation to educate students, while also satisfying the need to show proficiency on state assessments.

Chapter 3: Design and Methodology

Introduction

In the past, researchers have used a variety of techniques to collect data on teacher beliefs related to standardized tests. Some researchers opted to give surveys or to examine existing survey data related to accountability and standardized testing (Abrams, 2004; Croft et al., 2005; Mertler, 2011). Conducting a survey would make it possible to reach a wider population than other methods allow. Murnane and Papay (2010), however, point out that most available survey data suffers from either low response rates or a lack of a representative sample.

Researchers also have employed the use of in-depth interviews or focus groups as a means of data collection (Grant, 2000; Rex & Nelson, 2004; Valli & Buese, 2007). Murnane and Papay (2010) note that interviews and focus groups are useful because they can provide a great deal of detail about the context in which participants teach, but that they also can suffer due to teachers with only strong opinions about standardized testing volunteering to participate. Other studies employ a combination of both survey data and interviews or focus groups (Firestone et al., 2002; Lai & Waltman, 2008; Murnane & Papay, 2011).

Numerous surveys already have been conducted asking teachers about their attitudes toward standardized testing and other school accountability measures. This type of data may be useful because it can be generalized more broadly than interview data, but in this particular study the context in which the participants teach is important for understanding their choices. The focus of this study is on how teacher attitudes affect the choice of test preparation practices. A survey simply

would not provide the contextual information needed to evaluate why teachers do what they do in the same way that interviews and observations could. Instead, I opted to use a combination of interview data, observations, teacher journals, and test preparation artifacts to create case studies of two teachers who were faced with preparing their students for the Kansas State Science Assessments.

Methodology

I used the information I gained via interviews, observations, and documents to create case studies of two teachers as they prepared their students for the Kansas State Science Assessment. Yin (2003) explains that a case study is an appropriate choice when a researcher is interested in why and how events happen and when the context in which events occur is important for understanding those events. The focus of this study was on the decisions teachers made about preparing students for a standardized test.

Baxter and Jack (2008) point out that case studies are a good choice for researchers when a study aims to cover decision making because the context of the situation often informs the decisions that people make. It would be very difficult to understand the decisions these teachers made without knowing about the context in which they made those decisions. The literature on the subject of high-stakes test preparation points to numerous factors that might affect how teachers choose to handle the state assessment. There is a great deal of important information about the school, classrooms, and community in which these teachers work that informs their attitudes about testing and the test preparation methods that they choose.

Baxter and Jack (2008) note the importance of identifying the unit of analysis

of a study. In this case the unit of analysis was individual teachers. There were two participants and I wrote a case study about each one. I chose the individual as the unit of analysis because I was interested in looking at how the contextual factors related to each teacher affected her choices. This choice allowed me to compare the two cases in an attempt to better understand the different factors that might affect a teacher's attitudes about testing and test preparation.

Case study research has a number of strengths. It allows the researcher to provide a level of detail and depth that can illuminate factors for decision-making that could be difficult to pick up on using other methodologies. Case studies also allow researchers to link causes to outcomes, and to foster new hypotheses and research questions (Flyvbjerg, 2011). For example, other researchers have distributed questionnaires to collect data relating to similar questions that I asked in this study. Those researchers were able to tell how widespread the phenomena were, but their research was limited in that it couldn't answer the "why" questions. Their results may be more generalizable to broad populations, which is traditionally viewed as a strength in research, but I am interested in what caused the phenomena on the individual scale.

Flyvbjerg (2011) also points out some of the weaknesses of case study research. He notes that it is a widely held view that case study research is not generalizable. Much can be learned, however, from individual cases. Flyvbjerg brings up researcher bias as well, saying that many people believe case study research tends to simply verify preconceived notions. He argues that this historically has not been the case and that the same criticisms about bias can be

leveled at other types of studies that are generally viewed as relatively bias-free. None-the-less, it was important for me to be cautious about recognizing my own bias as I analyzed data and drew conclusions.

Sample Selection

This study focused on high school science teachers who were preparing their students for the Kansas State Science Assessment. There are two sections to the assessment. One section covers life science and the other covers physical science. Students are required to take both sections of the exam by the end of their junior year. At my high school, students take the life science portion while taking biology during their sophomore year, and then take the physical science portion during their junior year while taking chemistry, physics, or physical science. Each test contains roughly thirty multiple-choice questions on science topics that are marked as assessed indicators in the Kansas State Science Standards document. Students are required to answer questions that are on the level of knowledge recall, comprehension, and application.

I selected the building I teach in as the site for this study. A number of factors played into this decision. The first factor had to do with gaining access to participants for the study. I knew the teachers in the school and was able to ask them if they would be willing to participate prior to actually making a formal site selection. All of the potential candidates said that they were willing to participate as long as their other obligations allowed them to do so. I also knew the administrators in the building and they were enthusiastic about allowing me to conduct my research there.

Another reason I chose to do this research in the building in which I work is that it allowed me to collect data without interrupting my own professional obligations. I was teaching full-time while conducting this research and choosing the building where I work allowed me to collect data without having to take personal time off from my job to travel elsewhere. Besides convenience, the site was chosen because it was a school that struggled to just make AYP in both math and reading during the past few years. This led to an environment where a great deal of attention was placed on bringing students that were not meeting standards up to par. The state assessments were a hot topic of conversation among many of the staff members at the school, with constant updates coming about scheduling, scores, and the successes of students. With state testing such a part of life at the school, I was confident that preparations for state testing actually occurred so that I could collect the data this study required.

There also were a number of potential downsides to selecting this site. Participants may not have felt that they could be completely open and honest with someone that they saw on a daily basis as a colleague. To combat this, it was important for me to carefully define my role as a researcher, and to separate that from my daily role as a teacher and colleague. I assured the participants that their privacy would be maintained, which was key in gaining the trust of my participants. I wanted them to rest assured that anything they said to me, the researcher, would not lead to professional repercussions.

Managing my time also proved difficult, as I was teaching full time while trying to simultaneously conduct research. My job interfered with my planned

schedule for conducting research several times. For example, there were several instances where I planned to observe a participating teacher, but instead had to fill in as a proctor for state reading testing. If I conducted research outside my place of employment it may have been easier to establish the boundaries between work and research.

The study took place in a high school where four science teachers, myself included, have classes that took one of the state science assessments this year. All three other teachers were willing to participate in the study. While the time constraints of this study were not large, one of the science teachers had a coaching position that made it difficult for them to participate. I selected the two teachers whose schedules worked the best with the demands of my data collection.

I chose two teachers because I performed observations during my two planning periods. Because of the way tests are scheduled at the site, all of the science teachers were testing, and therefore preparing for testing, during the same week. I was not confident that I would have the time to reliably observe more than two classrooms without having to get a substitute to cover some of my own classes. I hoped having two participants would allow for some variety in the types of students they taught, levels of teaching experience, and level of knowledge about the state assessments. Because of my limited pool of potential candidates, I was not able to achieve all of these goals, but they were of secondary concern behind the primary concern of finding participants whose schedules allowed me to collect data. One unforeseen benefit of having two participants, as opposed to just one, was that speaking with and thinking about a participant often led to interesting questions to

ask the other participant. It also highlighted things for which I should look out, when I observed the other participant.

Collection of Data

I used four methods of data collection: interviews with teachers who were preparing students for state science assessments, observations of those same teachers as they prepared their students for state tests, teacher journals documenting the amount of time spent preparing students for standardized tests and the methods they used, and artifacts from test preparation such as handouts and practice materials. Multiple methods of data collection are meant to help ensure rigor by allowing me to “triangulate” what actually happened, using multiple sources of data to paint a clearer picture of what actually occurred (Tracy, 2010).

The idea of “triangulating” in qualitative research is similar to the method used in navigation where three compass bearings are used to pinpoint a traveler’s location. Tracy (2010) cautions, though, that multiple sources do not necessarily lead to the “truth.” For example, what a person says in an interview may differ from what they actually do in practice and deciding what is “true” may not be possible. What the method of triangulation does provide is a deeper understanding of a situation and it may illuminate new problems to explore.

Qualitative interviews were one of the methods I used to collect data for this study. Kvale and Brinkman (2009) describe the qualitative interview as a discussion of a “theme of mutual interest” (p.2) between interviewer and interviewee. In this case, the mutual interest was standardized testing and the conversation revolved around that topic. Qualitative research does not only rely on

“the facts,” but instead relies on the understandings that are specific to individuals (Dilley, 2004). In the case of this research, my understanding of how each individual perceived reality was important. Interviews allowed me to not only hear what the participants had to say, but offered me the opportunity to examine how they said things and to draw conclusions about what they wanted to convey (Dilley, 2004).

I used the approach of the general interview guide. This meant that I had a semi-structured interview, with a list of prepared questions, but the flexibility to ask follow-up questions or to ask questions in an order that made sense based on the direction of the conversation (Turner, 2010). Turner points out several advantages to this approach that I found useful. I was able to develop rapport with the subjects in a way that would have been difficult if I were limited to a very strict protocol. This approach also allowed me to ask follow-up questions to clarify points or to get more information about interesting topics. Taylor also points out the downside that this approach can lack consistency. The general direction of the interviews I conducted, however, were the same, and the information I obtained covered the topics I intended.

I interviewed the participating teachers individually, at a time that was convenient for them and for me. This took place during a shared planning period during the school day or after school, roughly one week before the teachers began to focus on the state science assessments. Interviews were video recorded so that I could produce transcripts, and so that I could make notes about the body language and facial expressions of the interviewees if it seemed relevant to understanding what was written in the transcripts. Penn-Edwards (2004) notes that videos can be

useful tools for coding and interpreting data, but cautions that their use may seem intrusive and can lead to reticence in participants. These provided clues as to how the teachers felt about standardized testing and I might have missed them if I only used an audio recording. Questions focused on the role standardized testing in their teaching practice, how they personally felt about high stakes testing, and how they planned to go about preparing their students for state standardized tests. The interviews were limited to thirty minutes, and each of them lasted roughly twenty minutes. The interview protocol can be found in Appendix A.

As the participating teachers began to prepare their students for the state assessments, I observed their classes where that test preparation took place. I was not able to observe every instance of the teachers doing test preparation because of my own teaching obligations. I was able to observe each teacher three times, for fifteen to twenty minutes on each visit. I took notes during each visit, and each observation was videotaped. This allowed me to watch the video of the class again so that I could make sure that I did not miss anything noteworthy or to clarify what was in my notes at a later time.

Within the context of qualitative research, observation does not mean that I simply observed what happened, but that I approached watching the participants in a systematic way (DeWalt and DeWalt, 2004). In this case I created an observation protocol, found in Appendix B, which had specific instructions for how I should carry out the observations.

I also identified some key themes in the literature for which I wished to watch for evidence. Those themes later were used when I coded data from all of the

sources I collected. The themes in which I particularly was interested were methods teachers used for test preparation, ethics of test preparation, amount of pressure teachers felt about achieving good scores, and the culture of the school as it related to standardized testing. I did not limit myself, though, to these areas. I tried to be as thorough as possible so that later I could examine the data more closely.

Polkinghorne (2005) notes that not everything that is observed necessarily is significant, but that significance also can become apparent later on when you revisit the data.

Polkinghorne (2005) also points out that observations can be very useful because they are used to supplement data from interviews. What I saw helped me understand the oral comments that I heard from the teacher. I was able to learn more about how seriously the teacher took the preparations for the test and how motivated the students were for the exam.

One concern with performing observations was that I wanted to be certain that what I was observing was indicative of what happened on a typical day in the observed teacher's classroom. I observed fifteen to twenty minutes of two classes with each teacher in which they were not preparing their students for tests, roughly one week before I did the test preparation observations. The purpose of these observations was to make both the teacher and students more comfortable with my presence in the room, and to inform me of what was and was not typical in that classroom. This procedure helped me establish what was normal and abnormal when I observed test preparation and it was meant to make my presence less of an abnormality for the participating teachers. The protocol for these initial

observations was the same as for the observations where I was seeking data on test preparation, and used the same protocol found in Appendix B.

Since it was not possible for me to observe every class where test preparation took place, I collected any artifacts related to the test preparation in the participating teachers' classrooms as well. Yin (2003) recommends in case study research that a variety of sources be used to collect data. Among his recommended methods are the collection of documents and artifacts. Creswell, Hanson, Clark Plano, and Morales (2007) suggest that participants in the study can write about their experiences.

In a classroom setting, it is common for teachers to provide paperwork for their students or to impart information using a computer, such as through the use of PowerPoint or similar software. Prior (2003) makes the point that it is not only important what documents contain, but how they are used. Documents can provide insight into the relationship between the producer of the content, in this case the teacher, and the consumer of the content, in this case the students. For this reason, I tried to observe the use of documents as much as possible. When I was unable to observe, however, I asked teachers to provide copies of the artifacts they used and I asked each participating teacher to keep a journal documenting the amount of time they spent on test preparation, what they did, how they used any materials, as well as their impressions of how their students received the information. Information on what teachers were asked to do in their journals can be found in Appendix C.

After state science assessments concluded, I conducted follow-up interviews with each of the participating teachers. There were several reasons for conducting

these interviews. First, I wanted to know if the teachers thought that the methods they used were effective. I wanted to get their reflections about what went well, what did not go well, and what they might change in the future. Once again, I videotaped these interviews to extract as much information from the interviews as possible. These interviews also were limited to thirty minutes, and they both took roughly fifteen minutes each. Information about the interview protocol for the post-test interviews can be found in Appendix D. Upon completion of these interviews, the participating teachers were provided with a copy of the transcripts from each of their interviews, both pre-test and post-test, so that they could verify that the information in the transcripts was accurate. Tracy (2010) calls this practice member reflection, although it is more commonly referred to as a member check. This practice helps give the analysis credibility by allowing an opportunity for additional data to emerge and for participants to elaborate on details.

Analysis of Data

After I collected data from interviews, observations, journal entries, and artifacts I had to create a system to analyze it. This is commonly referred to as content analysis, which Bowen (2009) describes as “the process of organizing information into categories related to the central questions of research” (p.32). There are a number of different approaches to content analysis, but I elected to use an approach developed by Mayring called qualitative content analysis.

Mayring (2000) describes the qualitative content analysis approach as using both inductive categories and deductive categories to organize data. Inductive categories emerge from the data through examination. Deductive categories are

developed based on theory or previous research and are the starting point for the data analysis (Zhang and Wildemuth, 2009). Thomas (2006) cautions, though, that deduction can be dangerous because it is easy for a researcher to be blinded to emerging themes by their preconceptions about what they will find in the data.

I elected to use the following steps for conducting a qualitative content analysis adapted from Zhang and Wildemuth (2009). The steps I undertook are as follows:

- 1) Prepare the Data - I transcribed the videotapes of both the pre-test and post-test interviews. I transcribed the audio literally, then went back through the video to write notes about emotion, hesitation, and any other potentially relevant action that I would have missed using only a literal transcript. This was meant to help me interpret the data later (Penn-Edwards, 2004).
- 2) Define the Unit of Analysis – There are a number of ways to approach classifying and categorizing text. One commonly used method is to do word frequency counts. However, there are some limitations to this method such as missing out on synonyms and use of words with more than one meaning (Stemler, 2001). I choose to identify themes as my unit of analysis, allowing me to find relevant chunks of data of any length that demonstrated a theme. This is the typical unit of analysis on qualitative analysis of content (Zhang and Wildemuth, 2009).
- 3) Develop Categories and Coding Scheme – I used both deductive and inductive themes to identify categories for coding the data. The deductive themes were based on the existing literature on the subject of standardized test preparation. The literature pointed to methods teachers used for test preparation, ethics of test preparation, amount of pressure teachers felt about achieving good scores, and the culture of the school as it related to standardized testing. I developed the interview

questions and observation protocols with these themes in mind. The inductive theme that I recognized during my reading of the interview transcripts and other data was teacher knowledge about the actual test, both as it related to content on the test and what were acceptable preparation practices for the test. I formally defined the themes and set up the rules for coding using the format recommended for the qualitative content analysis method by Mayring (2000). The coding agenda is found in Appendix F.

- 4) Test the Coding System – I tested my coding system on one of the pre-test interviews that I did with a participant. I found that the themes I identified seemed acceptable for coding the rest of the data except for the theme related to school culture. This theme did not appear in the interview on which I tested the coding system. I decided to still look for it in the other documents, but ultimately it did not show up in any significant way in the data I collected. As such, it was not included in the results section of the study.
- 5) Code the Text - I employed a color-coding scheme in which each theme was marked on a copy of the transcripts and other documents using a different color. This helped with the recognition of patterns both within the evidence gathered from a single participant as well as between multiple participants. Throughout the analysis process, I took notes on the themes that appeared and my analysis of the data I collected. These analytic memos were important reminders for when I reflected on the design and implementation of the study (Maxwell, 2005). Thematic memos (Rossman & Rallis, 2003) that I took throughout the coding process also served as important reminders of how the themes I found in the data related to each other and to the research questions.
- 6) Assess Consistency – Zhang and Wildemuth (2009) point out that as time goes by the coders may not stay consistent. I went back and rechecked all of the data to make sure that my understanding of the definitions of each theme stayed consistent throughout the coding process. I also used this

as an opportunity to check for any errors in coding.

- 7) Draw Conclusions - The data provided me with information that I used to formulate a picture of how standardized testing fits into each of the participants' job as a teacher. I looked at each individual teacher, but also for patterns in the responses I received and in the methods I observed across teachers (Weston et al., 2001). Patton (2002) offers a concise summary of what this process entailed. "Interpretation means attaching significance to what was found and making sense of the findings, offering explanations, drawing conclusions, extrapolating lessons, making inferences, considering meanings, and otherwise imposing order" (p.480). An important part of this phase of the research was for me to decide what data was relevant and what data provided illumination of the research questions. Organization of multiple sources of data and drawing conclusions from that data were key for providing answers to the research questions. The color-coding and analytic and thematic memos provided me with a backbone for the management of the data I collected.
- 8) Report Findings - I used the data and its analysis to write a case study about the teachers who participated in the study. The analysis of data led to the production of a case study about the teachers. The focus of the finished product was on the contextual factors that influenced the teachers' attitudes about standardized testing and how those affected the teachers' choices about test preparation practices. The themes I found throughout the analysis process provided the backbone of the story of the teachers and how they fit into the testing and test preparation processes. I combined some of the codes into the broader categories of attitudes and practices for the purposes of writing the case study.

One concern that I had to keep in mind throughout the analysis process was the bias that I brought to the table as the researcher. There were several methods that I used to make sure that my interpretation of the data was reliable and valid.

First, I used multiple methods of data collection to “triangulate” the meaning of the data I collected, which offered a deeper understanding of the topic than a single type of data might have (Tracy, 2010). Second, I used member checking where I presented the transcripts of interviews to the participants, as well as a summary of my notes to them so that they could offer any corrections or additional insights related to my analysis. This allowed me to establish the validity of my findings through the “lens” of the participants themselves, giving them the opportunity to provide feedback about whether their reality was represented accurately (Creswell & Miller, 2000). Finally, I kept a reflective journal throughout the process then used those reflections to create a section in my results about my role as the researcher (Creswell & Miller, 2000).

Ethical Considerations

There are a number of ethical considerations for which I had to account during this study. Before beginning the study, I obtained approval from the school where the study took place. There was no official procedure in place at the district level, but I obtained permission from my building principal to conduct the study. The principal provided a letter to the KU Human Subjects Committee granting me permission to conduct research at the school.

I also needed to ensure that the science teachers were willing to participate. Each participating teacher signed an informed consent form. A copy of this form can be found in Appendix E. The form laid what they were committing to in terms of the process involved in being a part of the research study. Along with the form, I

communicated with them about what was going to happen, as well as answering any questions that they had.

Privacy of the participating teachers especially was important for me to maintain. In the interviews the teachers may have said things that they would not want associated with their names in a public forum. I took precautions to make sure that my subjects are not identifiable in the final, written product of the study.

Before I was done analyzing the data, however, it was important for me to maintain records in which the statements of participants could not be associated with their identities. I did this by assigning pseudonyms to each participant, and by changing any other identifying information. It was important, though, to be careful not to manipulate contextual factors when I did this because context is extremely important to this study.

Another aspect of the privacy issue was my place as both a researcher and as a colleague of participants. There might have been issues with participants feeling unwilling to be truthful with me because of my standing as one of their colleagues. Thus, it was very important for me to explain to them my role as a researcher, and to assure them that anything they said to me would not be shared with anyone else in any way where their identity could be ascertained. I had to conduct myself very professionally in both roles, as researcher and colleague, and worked to maintain the trust I was able to build between researcher and subject.

The privacy of students also was an important consideration. While students were not the prime focus of this study, it seemed inevitable that they would be an important part of what happened during observations. The same considerations

used for the participating teachers applied for students as well. In addition, videotaping was an important issue when students were involved. The school district had consent forms on file from parents, which allowed me to videotape these observations. A copy of the form, which parents were required to sign at enrollment, can be found in Appendix E. Furthermore, any video I took with students in it was only viewed by me, the researcher. If a student did not wish to be videotaped, or for some reason there was not a consent form on file for a student, I had plans in place to make sure that the video arrangements were such that that student did not appear on camera. If this were not possible I would have abandoned videotaping that particular class in order to not violate the privacy of any students. This situation did not arise, though. All video files were erased upon completion of the data analysis, as specified in the human subjects application.

Chapter 4: Results

Introduction

I used the data I collected from the pre-test and post-test interviews, the classroom observations, and the documents provided along with teacher journals to create a case study about each teacher. I coded the data based on five identified themes: methods of test preparation, pressure teachers felt about achieving good scores, ethics of test preparation, school culture related to testing, and issues related to inexperience with the state assessments. The theme related to school culture provided almost no data, and all of the data I coded for that theme I also coded for the pressure theme. This left three themes that informed the decisions the teachers made related to the methods for test preparation they chose. In each case study I relate pressure, ethics, and issues with inexperience to test preparation to establish a picture of how each teacher's attitudes about testing affected what they did to prepare their students for testing.

The Site

The teachers in this study both work in a suburban, Midwest high school. The community is part of a large metropolitan area, but the community itself has a small town feel. It primarily is a working class community. A significant portion of the students in the school has parents who have high school diplomas but never went to college. Many students perceive the expectations at the high school to be lower than at other suburban schools in the area, and this likely is a function of the amount of wealth, or lack thereof, in the community compared with the surrounding area. Almost half of the students at the school fall under the state's "economically

disadvantaged” label, but you wouldn’t know it just by looking at the students or the modern facilities.

The school itself is small compared to many of the other schools in the area. There are about 700 students in grades nine through twelve. Each year about 150 of those students take the Kansas State Assessment covering life science, and about the same number take the corresponding physical science test. Most students at the school only take the three required science courses required to graduate, and the story is much the same in other academic subjects as well. There are only a few dozen students that take the higher-level science classes, physics and anatomy, during their time at this high school, and there are no honors or AP classes offered.

The school made AYP in math and reading the last two years, although just barely in both subjects. The number of students scoring proficient on the science exams has been similar to the numbers in reading and math; about 85% of the students passed the science assessment each of the last two years.

Katie

Katie is a second year science teacher in her mid-twenties. During my first interview with Katie I asked her to tell me about herself and the population of students she serves. She graduated from a nearby high school and went to college in the area, studying with the intention to become a high-school biology teacher. She has a long tradition of educators in her family and she felt like becoming a teacher was her destiny. She now teaches about ninety students, mostly sophomores, divided into four biology classes. Almost all of these students took the life science state assessment this year. In her biology classes Katie’s students have a wide array

of ability levels. A few in each class are labeled as gifted while she also has a few in each class that have IEPs for various disabilities.

She also has been tasked with teaching one section of chemistry, mostly juniors, which is a subject that she was not very comfortable teaching, having completed only two introductory college chemistry courses. This year it has been a bit easier than last year in terms of planning and content, but she has had to deal with over thirty students in her one section of the class. These students took the physical science state assessment this year. A significant portion of this class is high-achievers, and the difference in ability is not as wide as in her biology classes. A few factors played a role in Katie's decision about which test preparation methods to use. Her attitudes about testing and test preparation were influenced to varying degrees by the types of pressure she felt to achieve good scores, ethical considerations, and her identity as a relatively inexperienced teacher.

Pressure to achieve good scores. The pressure associated with having students perform well on the state assessment played the biggest role, of the themes identified in this study, in how Katie prepared her students for the exam. Katie indicated in our first interview that the greatest source of pressure that she felt was personal. She said, "I put pressure on myself because I am so competitive and I want my students to be the best." She spoke of the pride she felt when her students did well last year, and how she took it personally, blaming herself, when a student did poorly.

Katie said a number of things, both before and after the tests, that indicated that outside influences were not a significant source of stress for her. I asked her in

both interviews if she felt like other teachers, administrators, or community members expected her classes to perform well. She pointed out that no teachers commented about her scores, that she did not think anyone outside the school knew about how well her students did, and that she did not feel that anyone in a position of power in the school district was putting any pressure on her to perform. Another factor in the lack of stress was reflected in that Katie said she felt that the state assessment was just something that had to be done. She said:

I kind of just feel like it is a hoop we have to jump through. I mean they need to assess us somehow and make sure we are doing our jobs. But, it is kind of annoying hoop to jump through, that this is the only way they can measure us.

Despite these feelings, the pressure that Katie put on herself to have her students achieve high scores seemed to guide how she approached preparing for the test.

When I observed Katie as her classes prepared for the state assessment, she seemed to take the whole process very seriously. She produced an eleven-page study guide that she gave to her students one week before the test. The class divided the packet into sections, and each person was assigned a page or two. In examining the study guide I could tell that Katie put a lot of time and effort into creating it. It was very detailed, covering all of the assessed state biology indicators, with very numerous detailed questions on each page. I watched as the students worked independently on the packet. The questions were difficult for many students, requiring them to consult their biology textbooks. The students worked diligently, though, mirroring Katie's demeanor.

During my observation of the class, the students spent about forty-five minutes working independently, and then paired up with someone who had the same assigned questions. They verified their work, and then got up in front of the class the next day to share the answers they had come up with. Again, the atmosphere was very business-like, especially from Katie. I also observed the third day of preparation. The students finished going over the study guide then moved on to a review game. Throughout this activity, even over several days, Katie took preparing her students for the state assessment very seriously. She claimed there was no external pressure; that it was all self-imposed. It seemed that Katie genuinely wanted her students to succeed because she took pride in their success.

The state science assessment is situated in a somewhat awkward position that may explain some of Katie's feelings. The science assessment does not have the same kind of high-stakes repercussions that the reading and math tests do. Even though science assessments are required under ESEA, only the reading and math scores are used for determining AYP. Yet, the test is administered just like its more important counterparts; it is scheduled during the same block of time in the spring, and it follows the same rules and procedures. Katie's responses to interview questions, contrasted with her behavior in classroom observations, indicate that she knows that, ultimately, her students' performance will not have a big effect on the school. Her test preparation was quite serious and intense because, despite her annoyance with the exam, she used it as a personal measuring stick to see if she is doing a good job and because she wanted to see her students succeed.

Ethics of test preparation. During my first interview with Katie I asked her if she felt like test preparation was unauthentic or unethical. She replied, "I always feel a little uncomfortable just throwing a bunch of information at the kids...but I want to have something for them to review." Katie's journal entries showed that she spent one week solely working on test preparation with her students, a significant amount of time when considering a comment Katie made in the pre-test interview about feeling like she needed to move faster to cover all of the tested content before the test. She did not think that what she was doing, the study guide and review game, was unethical. She did, however, shy away from giving practice tests, which she said was something she did her first year of teaching. Katie told me in the interview before testing that she choose her test preparation practices in consultation with a more experienced biology teacher at the school, and the practices he had used in previous years helped inform what Katie saw as normal and acceptable in terms of test preparation.

While I was observing the students present the study guide questions they had answered to the rest of the class, there were a number of occasions where students asked questions like, "What should we write?" and "Do we need to write cellular respiration?" It seemed that the students were quite concerned with just having the right answer and instead of just giving them the answer, Katie encouraged them to "think about it" and "reason it out." This coincided with Katie's thought in our interview, "...it's not optimal to sit in class and just spoon-feed them a bunch of information."

Identity as an inexperienced teacher. During one class observation, once Katie's class finished going over the extensive study guide, they moved on to playing a Jeopardy-style review game. Katie numbered them off into teams of three or four, and they answered multiple-choice questions related to the tested indicators. It was at this point that I noticed the atmosphere in the room start to change. Students became restless; they had been preparing for the state assessment for three straight days and, although they took it quite seriously before, their patience appeared to be wearing thin.

In our interview after the state assessment was over, Katie recognized that the review game had not gone over well with her students. She said, "I don't think [the review game] was as effective as doing the study guide...some kids treated it as if it was something they didn't have to do." Katie gave a mixed message about the study guide. In our post-test interview she said it was "very effective," but just a few minutes later she said, "I feel like the results probably would have been similar," when I asked her if she thought not preparing would have changed her results. This seems to indicate that Katie was not completely confident in the practices she chose. She, however, did not choose her test preparation strategies on her own. She admitted that because she was a relatively new teacher she had chosen the test preparation practices in collaboration with the other, more experienced, biology teacher. In our interviews it was clear that many of the practices she chose for test preparation were greatly influenced by her colleagues.

Katie expressed in our first interview that she "want[s] to improve the scores next year," and "hope[s] they would go up as I become a better teacher." As a young

teacher Katie relied upon others in her department to provide guidance about how to get students ready for the state assessment. Katie said that she would use the data from this test, though, to “change the way I prepare my students.” Before her information about the tests came from other people, but as she learns more it appears she plans to adapt her test preparation practices to better fit her teaching style and the needs of her students.

Summary. The factors that affected how Katie felt about the state science assessment and influenced the test preparation practices she chose were self-imposed pressure to achieve high scores, ethical concerns, and her inexperience as a teacher. Katie indicated that she did not feel like other people in the school or community were putting pressure on her to produce high scores. She did, however, say that she took pride in her students doing well and that her competitive nature factored into how she approached preparing for the test. She said that she did not want to feel like she was “spoon-feeding” information to her kids in an unauthentic way, but felt that review for the test was necessary. Because of her inexperience, she also relied a colleague to provide guidance about what had worked for him in the past. In our first interview Katie said that she “had not seen [the data] used in any way,” and that “it kind of defeats the purpose of taking the test.” Despite this feeling, though, I observed Katie’s serious approach and thorough review. The themes discussed here provide insight into how Katie ultimately decided to prepare her students for state assessments.

Sarah

Sarah is a first-year teacher in her mid-thirties. In our first interview I learned how Sarah came to be a science teacher, and about the students she teaches. She spent the last fifteen years coaching volleyball at the university level and teaching some college courses in biology. Last year she worked as a paraprofessional at a middle school, but this is her first time managing her own classroom. Sarah is a popular teacher because of her positive attitude and her peppy personality.

Sarah teaches two sections of anatomy. Her students in those classes mostly are seniors that already have taken both state science assessments. Here three other classes, though, are integrated science. This course is an alternative to chemistry and generally is taken by students that have low academic ability or low motivation to do science. Most of these students took the state physical science assessment this year. Despite having the difficult job of motivating students who normally do not do well in school, Sarah is optimistic about her students. She encourages them and wants to see them succeed. Sarah's decisions about the test preparation practices she used were influenced by the types of pressure she felt about testing, the ethicality of particular practices, and her uncertainty about what practices to choose because of her inexperience as a teacher. I analyzed the data and organized it based on those themes.

Pressure to achieve good scores. Sarah indicated in our pre-test interview that she did not know very much about the state assessment. She assumed that the scores were not terribly important to administrators in the building because no one

had spoken to her about testing. She did express, though, that she was putting a lot of pressure on herself, especially because she was a new teacher. She said that she was using the results as a way to gauge how well she had done in her first year of teaching.

Sarah said that she was nervous about testing, especially because she did not know how much effort her students would put into it. She said, though, that she really had not thought much about testing prior to a few weeks before it was set to begin. When I asked her if testing caused her any stress she said, “No it doesn’t. But that is kind of my personality too...I’m spending time on the standards. That is kind of where I’ve been.” Testing had not been a focus for Sarah, but as the date was approaching, she shared that she had realized that she really did not know much about what was happening.

Sarah was optimistic that the results on the state assessment would help guide how she teaches next year. She was very interested in using the data from the tests, even though she said that she had not seen any data from previous years or seen any data from last year used in the building in any way. She wanted to use the data as a personal measure of how well she had done during her first year of teaching. When I observed her, she communicated her optimism through her encouragement of her students. She tried to build their confidence, even though many of them appeared to me to be unenthusiastic about testing.

Sarah’s stated lack of stress about the tests did not show through in the actual test preparation that I witnessed. Sarah’s original plan that she told me about in the first interview was “ten minutes each day we are just going to review

concepts that we have already talked about.” In reality, though, Sarah spent much more time on test preparation than she originally planned. In her journal about what she did each day, she wrote that she spent more time than she expected on all but one of the days she did test preparation. One day when I observed her class, she spent the entire hour answering review questions on the internet with her students. Despite feeling a lack of pressure from outside sources, it seems that the pressure that Sarah put on herself to achieve high scores caused her to stray from her original plan, putting more emphasis on test preparation than she originally planned.

Ethics of test preparation. Sarah also was quite concerned with the ethics of preparing her students for the state tests. She said, “Honestly, I am a little worried about how much I can really prepare them for the tests. I don’t know ethically how much time I really can spend preparing.” She said that because of this uncertainty that she would rely on colleagues to help her figure out what to do. She also talked about the reality of needing to review old concepts, but wishing that she had “a little more direction” on exactly what she could do.

In our first interview, she also talked about how she thought she could overcome some of her ethical qualms, because she was building the confidence of her students. She said that she wanted them to sit down to take the test and “feel comfortable and confident” and that “confidence will go a long way for them if we have spent some time preparing.” When I observed her classes as they prepared for the assessment, my sense was that Sarah was focusing more on building confidence than she was on actually making sure her students remembered information. When one student got a question right, Sarah said, “Awesome! See I knew you guys would

remember this stuff.” She was constantly encouraging, even when students got answers wrong. On another day when a student was struggling with the formula for density, Sarah helped them walk through the problem until they reached a point where the student could handle it the rest of the way. When the student finished the problem, Sarah gave them a high-five, even though she had done most of the work. After the test, in our interview, Sarah reflected on her test preparation, saying, “I don’t feel like I over-coached.”

Identity as an inexperienced teacher. Sarah’s identity as a first-year teacher came up numerous times in both interviews. In our first interview she talked about how she did not really know what to expect from the state assessment beyond her knowledge of the state standards. Despite her lack of awareness about the exam, in my observations Sarah seemed very confident when telling her students about the exam. During one observation, she led her students through some sample test questions from a different state. She made them aware that she did not know exactly what the questions might look like on the test, but that “this is something that you need to know.” She acted in front of her students as a strong leader despite her private lack of knowledge about the test.

Sarah relied heavily on the advice of colleagues as she prepared for state assessments. In the middle of our first interview, she asked me several questions about what the tests looked like and when we would get the results. She also expressed some frustration about not knowing much about earth and space science, which is covered on the state physical science assessment but not in her class. She said she would have to “touch base with colleagues that teach that [earth and space

science]” so that she could consult, not about test preparation tactics, but about content.

During the interview after testing was done, Sarah admitted that she still was not extremely confident about the state assessment. When we talked about the rules prohibiting discussion of the tests, even after students were all done, she remarked that “I still don’t have a true understanding of what questions are on there.” Despite her lingering uncertainty, she said she felt like she had learned enough to make some adjustments for next year.

Summary. Sarah’s biggest influences in how she felt about testing and the practices she chose as a result of those feelings were self-imposed pressure for good scores, ethical concerns, and her inexperience as a teacher. Sarah said that she had not heard about testing from anyone outside her department, which made her feel like there was not much pressure on her to achieve high scores. Even after the test, she told me, “I don’t get the sense that anyone cares what happens.” She did want to use the test as a way to evaluate her job as a first-year teacher, though, and she ended up spending more time on test preparation than she had planned because she felt the need to get her students ready for the exam.

The most influential factor in Sarah’s decision about test preparation was her concern about doing something unethical. She planned to spend only a few minutes each day reviewing old content, while continuing to teach her class like normal. In the end, I observed her spending much more time reviewing than she planned, but after the test, she said she still felt like she had not done too much. Finally, Sarah expressed a lot of uncertainty about the test, both before and after it took place,

because of her inexperience dealing with the state assessment. She relied on her more experienced colleagues to provide insight into what they had done in the past and what they thought were effective test preparation strategies.

My Role as the Researcher

Creswell and Miller (2000) suggest that in qualitative research, the researcher should keep a reflective journal, so that he can make observations about his role in the research and disclose the “lens” through which he sees his work. As I prepared my own classes for the state physical science assessment I wrote about what I did to prepare my students, and how I played a part in determining what my colleagues, some of them participants in this study, did to prepare their students. My own feelings about the state assessment were that it was something that we had to do, but that preparing for it is not my primary purpose. In my chemistry and physics classes, I use the state standards document as a starting point, but I aim to go beyond it’s scope in an effort to prepare my students to study science in college. I do, however, feel some pressure from one individual, my department chair and mentor, to achieve high scores on the state assessment. I prepare my students for the state assessment because of this pressure.

Throughout this research the participants in this research occasionally asked me for advice about what they should do to prepare their students for state assessments. I shared what knowledge I had with them, although I do not know the extent of my influence on the decisions that ultimately made. I can say that they did not use any test preparation materials that I created myself. As I worked on this research I found that my role as a colleague was impossible to separate from my

role as a researcher. I think, though, that my familiarity with the context in which the participants of this study were working helped me understand their motivations and decisions better than I would have been able to understand if I was an outsider. Despite my involvement with the participants in this research, I did my best to set aside my personal beliefs and biases.

Chapter 5: Conclusions

This study looked at two science teachers as they prepared students for the Kansas State Science Assessment. The purpose of the study was to examine how a teachers' attitudes about the assessments affected their preparation of students for those exams. The teachers in this study expressed different emotions about the tests. Katie saw them as a "hoop to jump through," although she also put pressure on herself to get good scores. Sarah was excited for the tests because she wanted to get feedback about how she had done in her first year of teaching and how she could improve in the future. Both teachers used test preparation methods that are documented in the literature (Croft et al., 2005; Fedore, 2006; Mertler, 2011; Olson, 2001; Rex & Nelson, 2004). Katie used an extensive study guide covering all of the content from the entire year. She also used review games. Sarah used daily questions that mirrored those that might appear on the exam as a means to review formulas the students saw months earlier. The teachers also reflected on their choices. They were generally happy with them, although they expressed some doubt about their effectiveness.

Research Question 1

What are teachers' attitudes in regard to standardized tests?

In the cases examined in this study, there was a range of feelings about standardized tests. This coincides with the literature on the topic, which suggests that feelings vary (Mertler, 2011; Murnane & Papay, 2010). One teacher described the test as "annoying," and observed that she had not seen the data being used in her school. I still observed her taking preparation for the testing seriously, though,

because she put pressure on herself to get good scores. She said in an interview that because she was “competitive” she wanted her students to score well. She also said that she wanted to be able to assess how much her students had learned, but she recognized that the state assessment was not necessarily the best indicator of student learning.

The other teacher was excited for the tests because she wanted to use the data to verify her teaching practices and make adjustments. This was her first time giving the state assessment as a classroom teacher, and she expressed a lot of uncertainty in our pre-test interview about the tests. She noted that she had not seen how the data from last year’s test was used, if at all, but that she hoped to use this year’s data to help her better serve her students.

Research Question 2

Did the attitudes of the teacher play a role in how they prepared their students for the standardized test?

The interview responses of both participants indicated that they hoped to use the data from the state assessment to either confirm or to improve the effectiveness of their instruction. They both said that beyond their own personal use, though, they had not seen the results from the test used, both before and after the test occurred. The pressure that the teachers felt related to their students performing well on the exam was personal. Despite the indications that the exam did not seem to be important to other people in the school or community, I observed both teachers take test preparations quite seriously. One teacher created a very extensive study guide that she used to help her students review for the test. The

other teacher spent much longer than she had anticipated spending on test preparation in an attempt to review all of the material the exam covered.

Both teachers also expressed some concerns in the pre-test interview about feeling like they were “spoon-feeding” information to their students right before the test. After the exam, though, one teacher said that all she did was “review some material” and that she didn’t “feel like I over-coached.” The practices that I observed appeared to be meant to simply refresh material that the classes had not covered in several months rather than an attempt to “cram.”

The teachers I spoke with and observed in this study both were relatively inexperienced teachers. In the interview before testing, they both talked about how they relied on colleagues to help them choose which test preparation practices to use. This seemed to stem mostly from their lack of knowledge about the exam, with one teacher saying “I... don’t know what is on the test.” The other teacher talked about trying to find “what works” and about making changes to what she does next year as she “learn[s] more about how to prep them [the students].”

The choices that these two teachers made seemed to be influenced by the self-imposed pressure to produce good scores, the ethical considerations concerning reviewing material, and their feelings of uncertainty as inexperienced teachers. Other factors also may have played a role in the decisions of the teachers in this study, but any other factors were not apparent from the data that I collected.

Conclusions and Implications

There were a number of factors that affected how the teachers in this study felt about the state assessments that their students were required to take. Both

teachers expressed that they felt some pressure, but the source of that pressure was surprising. The pressure was self-imposed rather than from outside sources. Contextual factors might explain why there was not much outside pressure on the teachers. Math and reading count for AYP, while science does not. There also are documented cases of administrator checks for compliance with district-approved methods and personnel decisions based, at least in part, on test scores (Barrett, 2009; Valli & Buese, 2007). Neither of these were present in this situation, though. The self-imposed pressure had two results. First, it meant that the teachers only felt responsible to themselves when it came to their scores. Second, it meant that the teachers still took the tests seriously, even though they did not feel there would be repercussions if their scores were low.

Generalizing these results is problematic because they rely on the many contextual factors. The decisions that the teachers in this study made were based on the specific situations in which they worked. If the administrators in the building had demonstrated concern about scores, if the scores were printed in the local newspaper, or if the teachers involved had bad experiences with state tests in the past, it is very possible they would have made different decisions about how to prepare their students for the tests.

One theme that emerged in the data was the uncertainty that these young teachers had about the state assessment. Sarah, the first-year teacher, had never seen the assessment before and had no idea what the questions looked like or how many there were. She also did not know the depth of knowledge tested. The standards documentation provided by the Kansas Department of Education was the

only guidance she had in terms of what might appear on the exam, but the amount of information in that document is sparse. An exam that is as important enough to give to all students in the state during their high school career ought to be more transparent. Obviously cheating is a concern if too much information is provided. The CETE TestBuilder provides some example questions, but teachers have to rely on district administrators or colleagues to make them aware of its existence. It is not clear where the responsibility falls in regard to educating teachers about the tests and available preparation resources, but it would be advisable to do something to increase awareness about what these important tests look like.

Another theme that emerged was the uncertainty teachers had about what were acceptable test preparation practices. It is not always clear what is and is not ethical when it comes to test preparation (Croft et al., 2005). The teachers in this study received no guidance about what was appropriate, besides the warnings in the test instructions that they were not allowed to discuss test questions even after the test. Instead, the teachers had to use their own professional judgment, which could lead to problems with test validity if some teachers deem unethical practices to be acceptable.

This study has helped me reflect on my own choices about test preparation. I now find myself asking “why” questions more often when I make a choice to use a particular strategy. I also reflect more about whether the choices I made were effective and how they can be improved for next time. I also have come to realize my importance as a more experienced teacher than some of my colleagues. I did not realize that, despite still being somewhat new to the teaching profession myself, my

less experienced colleagues look to me for guidance. I have a better understanding of the importance of making wise choices for my own classroom so that if others choose to emulate me, they will be better off.

Recommendations

Uncertainty about the state assessment was one theme that emerged during the course of this study. Both participants were inexperienced teachers and the only guidance they had about the state assessment came from their peers. For this reason, I would recommend the following: On the school and district level, test coordinators should make sure that specific guidelines are in place about what are acceptable test preparation practices. The Kansas State Department of Education requires that each school district appoint a test coordinator, who must attend state training, and that each building also has a test coordinator, who is trained by the district test coordinator. These individuals are responsible for training teachers on test security using materials provided by the state department of education. These materials outline what is strictly prohibited, such as asking questions that appear on the test, but leave many gray areas for teachers to grapple with on their own. For example, a fact sheet on appropriate testing practices says that “comprehensive reviews” are prohibited, but it is not clear what qualifies as “comprehensive” (KSDE, 2012). A more thorough conversation at the school or district level than what the current state-provided materials offer could help teachers resolve issues about what practices they should choose.

On the state level, there are not many resources for teachers to understand what the state assessment looks like. The state standards document indicates which

standards will be tested, and the CETE TestBuilder website has some sample questions, but beyond these resources teachers do not have much information about the exam. I had to inform one teacher about the length, format, and scope of question that appear on the state assessment, and even I was not certain about much of what I said since teachers are prohibited from reviewing the test as their students take it. The state should provide some sort of summary about the test to give teachers a better understanding of exactly what their students should expect.

The teachers in this study wished to use the state assessment as a way to improve their instruction. When the scores for the state assessment are reported to the teacher, though, they only see one number, a score out of 100 for the entire test. It would be much more useful for a teacher to see a more detailed report of the scores, broken down by standard. This would allow teachers to recognize weaknesses in their instruction and to focus on topics that their students scored poorly on. This report could be done either by student or as a summary for entire classes or schools. The feedback that teachers get right now, though, is not particularly useful in identifying instructional weaknesses.

Further research should be done into how teachers make choices about test preparation. A body of research exists about what practices teachers use, but most of it focuses on “what” rather than “why.” In this study, the teachers relied heavily on the input of colleagues when they chose tests preparation strategies. It would potentially be very useful for teachers to have more information, supported by literature, about what practices are effective and ethical. It also could be useful to identify further the sources of pressure that teachers feel that relate to the choices

in the practices they use, so that administrators can help to alleviate stress on science teachers preparing students for state testing.

References

- Abrams, L. M. (2004). Teachers' views on high-stakes testing: Implications for the classroom. Education Policy Studies Laboratory, Arizona State University.
- Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory Into Practice, 42*(1), 18-29.
- Alon, S., & Tienda, M. (2005). Assessing the "mismatch" hypothesis: Differences in college graduation rates by institutional selectivity. *Sociology of Education, 78*(4), 294-315.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher, 36*(5), 258-267.
- Barrett, B. (2009). No Child Left Behind and the assault on teachers' professional practices and identities. *Teaching and Teacher Education, 25*(8), 1018-1025.
- Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report, 13*(4), 544-559.
- Bernstein, B. (2000). *Pedagogy, symbolic control, and identity*. [Revised edition]. London: Taylor and Francis.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27-40.
- Buchmann, C., Condrón, D., & Roscigno, V. (2010). Shadow education, American style: Test preparation, the SAT and college enrollment. *Social Forces, 89*(2), 435-462.
- Byrd-Blake, M., Afolayan, M. O., Hunt, J. W., Fabunmi, M., Pryor, B. W., & Leander, R. (2010). Morale of teachers in high poverty schools: A post-NCLB mixed methods analysis. *Education and Urban Society, 42*(4), 450-472.
- Common Core State Standards Initiative. (2012). Frequently asked questions. Retrieved from www.corestandards.org/resources/frequently-asked-questions
- Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice, 39*(3), 124-130.

- Creswell, J. W., Hanson, W. E., Clark Plano, V. L., & Morales, A. (2007). Qualitative research designs: Selection and implementation. *The Counseling Psychologist*, 35(2), 236-264.
- Croft, M., Waltman, K., Middleton, K., & Stevenson, E. (April, 2005). The impact of school-level accountability on local test preparation practices. Paper presented at NCME annual meeting, Montreal, Canada.
- Darling-Hammond, L. (2007a). Race, inequality and educational accountability: The irony of 'No Child Left Behind'. *Race Ethnicity and Education*, 10(3), 245-260.
- Darling-Hammond, L. (2007b). The flat earth and education: How America's commitment to equity will determine our future. *Educational Researcher*, 36(6), 318-334.
- Dee, T. S., & Jacob, B. (2011). The impact of no child left behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418-446.
- DeWalt, K. M. & DeWalt, B. R. (2011). *Participant observation: A guide for fieldworkers*. (2nd ed.). Plymouth, UK: AltaMira Press.
- Diamond, J. B. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education*, 80(4), 285-313.
- Dilley, P. (2004). Interviews and the philosophy of qualitative research. *The Journal of Higher Education*, 75(1), 127-132.
- Fedore, H. (2006). De-stressing high-stakes testing for NCLB. *The Education Digest*, 71(6), 23-28.
- Firestone, W. A., Monfils, L., Camilli, G., Schorr, R. Y., Hicks, J. E., & Mayrowetz, D. (2002). The Ambiguity of test preparation: A multimethod analysis in one state. *Teacher's College Record*, 104(7), 1485-1523.
- Flyvbjerg, B. (2011). Case study. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (301-316). Thousand Oaks, CA: Sage.
- Foster, S. (2012, July 19). Kansas receives NCLB waiver. *The Topeka Capital-Journal*.
- Fuller, B., Wright, J., Gesicki, K., & Kang, E. (2007). Gauging growth: How to judge no child left behind? *Educational Researcher*, 36(5), 268-278.
- Gallagher, C. (2003). Reconciling a tradition of testing with a new learning paradigm. *Educational Psychology Review*, 15(1), 83-99.

- Grant, S. G. (2000). Teachers and tests: Exploring teachers' perceptions of changes in the New York state testing program. *Education Policy Analysis Archives*, 8(14), 1-26.
- Guilfoyle, C. (2006). NCLB: Is there life beyond testing? *Educational Leadership*, 64(3), 8-13.
- Harris, D. N. (2007). High-flying schools, student disadvantage, and the logic of NCLB. *American Journal of Education*, 113(3), 367-394.
- Hoffman, L. M., & Nottis, K. E. K. (2008). Middle school students' perceptions of effective motivation and preparation factors for high-stakes tests. *NASSP Bulletin*, 92(3), 209-223.
- Hursh, D. (2008). *High-stakes testing and the decline of teaching and learning*. New York: Rowman and Littlefield.
- Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in Chicago Public Schools. *Journal of Public Economics*, 89(5), 761-796.
- Jacob, B. A., & Levitt, S. D. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*, 118(3), 843-877.
- Jones, B. D., & Egley, R. J. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing. *Education Policy Analysis Archives*, 12(39), 1-29.
- 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. *The Phi Delta Kappan*, 81(3), 199-203.
- Kansas State Department of Education. (2011). Public schools and districts not making AYP 2011. Retrieved from <http://www.ksde.org/Portals/0/Title%20Programs%20and%20Services/2011%20Not%20Making%20AYP%20Schools%20and%20Districts.pdf>
- Kansas State Department of Education. (2012). Kansas assessment fact sheet: Appropriate testing practices 2012-2013. Retrieved from <http://www.ksde.org/Default.aspx?tabid=420>
- Kansas State Department of Education. (2012). Kansas ESEA flexibility waiver overview. Retrieved from <http://www.ksde.org/Default.aspx?tabid=5075>
- Kvale, S., & Brinkman, S. (2009) *InterViews: Learning the craft of qualitative research interviewing* (2nd ed.). Thousand Oaks, CA: Sage.

- Lai, E. R., & Waltman, K. (2008). Test preparation: Examining teacher perceptions and practices. *Educational Measurement: Issues and Practice*, 27(2), 28-45.
- Lattimore, R. (2005). African American students' perceptions of their preparation for a high-stakes mathematics test. *Negro Educational Review*, 56(2), 135-146.
- Longo, C. (2010). Fostering creativity or teaching to the test? Implications of state testing on the delivery of science instruction. *Clearing House*, 83(2), 54-57.
- Maleyko, G., & Gawlik, M. A. (2011). No Child Left Behind: What we know and what we need to know. *Education*, 131(3), 600-624.
- Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.). Thousand Oaks, CA: Sage.
- Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research*, 1(2).
- McMillan, J. H., Myran, S., & Workman, D. (1999). *The impact of mandated statewide testing on teachers' classroom assessment and instructional practices*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
- Mertler, C. (2011). Teachers' perceptions of the influence of No Child Left Behind on classroom practices. *Current Issues in Education*, 13(3), 3-35.
- Murnane, R. J., & Papay, J. P. (2010). Teachers' views on No Child Left Behind: Support for the principles, concerns about the practices. *Journal of Economic Perspectives*, 24(3), 151-166.
- National Center for Education Statistics. (2011). NAEP-TIMSS linking study. Retrieved from <http://nces.ed.gov/nationsreportcard/pdf/about/schools/2011472.pdf>
- National Center for Fair and Open Testing. (2013). Confirmed cases of test-cheating. Retrieved from <http://fairtest.org/sites/default/files/CheatingReportList.pdf>
- National Commission on Excellence in Education. (1983). A nation at risk: A report to the nation and the secretary of education. Washington, DC: US Department of Education. Retrieved from <http://www2.ed.gov/pubs/NatAtRisk/index.html>

- Neal, D., & Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics*, 92(2), 263-283.
- Olson, L. (2001). Overboard on testing? *Education Week*, 20(17), 23–30.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Penn-Edwards, S. (2004). Visual evidence in qualitative research: The role of videorecording. *The Qualitative Report*, 9(2), 266-277.
- Perlman, C. L. (2003). Practice tests and study guides: Do they help? Are they ethical? What is ethical test preparation practice? In J. E. Wall & G. R. Walz (Eds.), *Measuring up: assessment issues for teachers, counselors, and administrators* (pp.387-396). Greensboro, NC: ERIC Counseling and Student Services Clearinghouse.
- Polkinghorne, D. E. (2005). Language and meaning: Data collection in qualitative research. *Journal of Counseling Psychology*, 52(2), 137-145.
- Prior, L. (2003). *Using documents in social research*. Thousand Oaks, CA: Sage.
- Rex, L., & Nelson, M. (2004). How teachers' professional identities position high-stakes test preparation in their classrooms. *Teachers College Record*, 106(6), 1288-1331.
- Rossmann, G. B., & Rallis, S. F. (2003). *Learning in the field: An introduction to qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.
- Sadovnik, A. (2006). Toward a sociology of educational change: An application of Bernstein to the US 'No Child Left Behind' Act. In R. Moore, J. Beck, M. Arnot, & H. Daniels (eds.), *Knowledge, power and educational reform: Applying the sociology of Basil Bernstein* (pp. 196–210). London: Routledge.
- Shepard, L. (2000, April). The role of assessment in a learning culture. Presidential Address presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Stemler, S. (2001). An overview of content analysis. *Practical Assessment, Research & Evaluation*, 7(17).
- Strauss, V. (2013, April 1). Atlanta test cheating: Tip of the iceberg? *The Washington Post*.

- Sunderman, G. L., Tracey, C. A., Kim, J., & Orfield, G. (2004). *Listening to teachers: Classroom realities and no child left behind*. Cambridge, MA: The Civil Rights Project at Harvard University.
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237-246.
- Tracy, S. J. (2010). Qualitative quality: Eight “big-tent” criteria for excellent qualitative research. *Qualitative Inquiry*, 16(10), 837-851.
- Turner, D. W. (2010). Qualitative interview design: A practical guide for novice investigators. *The Qualitative Report*, 15(3), 754-760.
- US Congress. (1994). Public Law 103-227, Goals 2000: Educate America Act. Retrieved from <http://www2.ed.gov/legislation/GOALS2000/TheAct/index.html>
- US Congress. (2001). Public Law 107-110, No Child Left Behind Act of 2001. Retrieved from <http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf>
- US Department of Education. (1991). America 2000: An educational strategy. Retrieved from <http://www.eric.ed.gov/PDFS/ED327985.pdf>
- US Department of Education. (2009). Race to the top program executive summary. Retrieved from www2.ed.gov/programs/racetothetop/executive-summary.pdf
- US Department of Education. (2010). A blueprint for reform. Retrieved from www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf
- Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, 44(3), 519–558.
- Vogell, H. (2011, July 6). Investigation into APS cheating finds unethical behavior at every level. *The Atlanta Journal-Constitution*.
- Wageman, J., & Bartoszek, K. (2004). *Adequate yearly progress and No Child Left Behind: What could AYP mean for your child, school, and community?* Fargo, ND: NDSU Extension Service.
- Weston, C., Gandell, T., Beauchamp, J., McAlpine, L., Wiseman, C., & Beauchamp, C. (2001). Analyzing interview data: The development and evolution of a coding system. *Qualitative Sociology*, 24(3), 381-400.
- Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Zhang, Y. & Wildemuth, B. M. (2006). Qualitative analysis of content. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science*. Westport, CT: Libraries Unlimited.

Appendix A

Pre-assessment Interview Protocol

Instructions to Interviewer

Read the following instructions. Do not proceed with the interview until you have obtained informed consent from the participant. During the interview, appropriate probe questions may be asked between the questions provided here. Take notes during the interview to help facilitate analysis of the answers provided by the interviewee.

Instructions (to be read to interviewee):

We would like to videotape this interview in order to help produce a transcript which you will be allowed to view and verify once it has been produced. The video will only be viewed by the researcher, and will be destroyed once we have completed a transcript of this interview. Please sign the release form.

Please also read and sign this form, which meets our human subject consent requirements. It states that you are voluntarily participating in this research and that your privacy will be maintained throughout this research process.

This interview will last no longer than one hour.

We would like to speak with you about the state science assessment, which your students will take this year. Particularly we are interested in how you prepare your students for standardized assessments.

1. How big of a role do the state science standards play in your practice as a teacher?
2. How do you prepare your students for state science assessments?
3. How much time do you spend on test preparation?
4. How long have you been teaching? How many times have you given state science assessments?
5. Tell me about any concerns you have in regard to your students meeting standards on the state science assessment.
6. Describe any pressure you feel from anyone to have your students meet standards? (Administrators, fellow teachers, parents, community, school board?)
7. How does this pressure influence your teaching?
8. How do you think your students feel about the state standardized tests?
9. How does the culture of your school play into how you feel about state standardized testing?
10. What could be done to better use the time devoted to state science assessments?
11. How is the data from state science assessments being used in your school? What about outside your school?
12. What have you been doing so far this year to prepare your students for the state science assessments?
13. What kind of test preparation do you plan to do in the weeks/days leading up to the state science assessments? Why have you chosen those approaches?

14. What role do the other members of your department play in how you prepare your students for state science assessments?
15. How do you decide what content to cover when you prepare for state assessments?
16. What approaches to test preparation might be more effective than what you have used in the past or plan to use in the future?
17. What kinds of assessment do you think would tell you the most about what your students have learned? How well do the state assessments align with the types of assessment approaches that you value?
18. What kind of rewards do you see when your students do well on state assessments?

Appendix B

Observation Protocol

Instructions to Observer

You will be observing a high school science class as they prepare for state science assessments. The focus of this observation should be on the practices that the teacher employs in order to prepare their students for state testing. Take note of any materials used, the methods employed, the attitudes the teacher shows, and the response of the students.

Guiding Questions

- Does the teacher use any of the following approaches to test preparation:
Practice tests, review games, or student-directed study?
- Is review done individually by the students, in small groups, or as a whole class?
- Does the teacher use technology in their test preparation practices?

Teacher_____

Date_____

Hour_____

Class_____

Number of Students_____

Description of the Classroom:

Description of the Teacher:

Descriptions of the Students:

Lesson Chronology (Note test prep activities):

Materials Used:

Time / Observation

Comments

Appendix C

Teacher Journal Protocol

Instructions for Teachers

As you prepare your students for the state science assessment, I would like you to record any test preparation activities you do, whether I am there to observe your class or not. Any handouts you use should be provided along with this journal. Also, if you use any visual aides or internet sites, please provide a copy of the web addresses where they can be found, or a copy of any computer files you use. Please provide the following information so that I have an accurate record of all of the test preparation that took place in your class.

Date:

Class:

Please provide a brief description of the test prep activity:

How long did you spend on this activity:

Please provide any notable impressions you had about how the activity went:

Other information (including relevant websites):

Appendix D

Post-assessment Interview Protocol

Instructions to Interviewer

Read the following instructions. Do not proceed with the interview until you have obtained informed consent from the participant. During the interview, appropriate probe questions may be asked between the questions provided here. Take notes during the interview to help facilitate analysis of the answers provided by the interviewee.

Instructions (to be read to interviewee):

We would like to videotape this interview in order to help produce a transcript which you will be allowed to view and verify once it has been produced. The video will only be viewed by the researcher, and will be destroyed once we have completed a transcript of this interview. Please sign the release form.

Please also read and sign this form, which meets our human subject consent requirements. It states that you are voluntarily participating in this research and that your privacy will be maintained throughout this research process.

This interview will last no longer than forty-five minutes.

We would like to speak with you about the state science assessment, which your students just took. Particularly we are interested in how you prepared your students for the state science assessment.

1. What did you do to prepare your students for the state science assessments?
2. What would you say was your most effective test preparation strategy?
3. What would you say was your least effective test preparation strategy?
4. How will you change the way you prepare your students for the state science assessment next year based on what happened this year?
5. How much of an effect do you think your test preparation had on your students' test scores?
6. What were you disappointed about in regards to the results of your students' tests?
7. What kind of performance on the state tests do you think would satisfy your administrators? What about the community?
8. How aware do you think people inside of the school are of how your students did on the state tests? What about outside the school?
9. How does your students' results on the state tests in comparison to the results other teachers' students affect you?
10. How does your attitude about state testing change based on whether testing is coming up or is already completed?

Appendix E

INFORMED CONSENT STATEMENT

(Name of the Study)

INTRODUCTION

The Department of Curriculum and Instruction at the University of Kansas supports the practice of protection for human subjects participating in research. The following information is provided for you to decide whether you wish to participate in the present study. You may refuse to sign this form and not participate in this study. You should be aware that even if you agree to participate, you are free to withdraw at any time. If you do withdraw from this study, it will not affect your relationship with this unit, the services it may provide to you, or the University of Kansas.

PURPOSE OF THE STUDY

The purpose of this study is to look at the factors that affect how a teacher prepares their students for the Kansas state assessments in science.

PROCEDURES

You will be asked to participate in two interviews, both lasting no more than half an hour each. You will also be asked to allow the researcher to observe you teaching at least five classes over the course of several weeks as you prepare students for state science assessments. Both the interviews and the observations will be videotaped. You will also be asked to keep a journal of test preparation activities you use in your class. The total amount of time you will commit to this study is estimated at six hours total (1 hours for interviews, 4 hours of observation in the classroom, 1 hour keeping the journal).

VIDEO AND AUDIO RECORDING

Both the interviews and classroom observations will be videotaped. The purpose of these recordings is so that the researcher (Chris Klager) can transcribe the audio from the interviews and can make a detailed account of the test preparation done during the classroom observations. No one else will have access to the recordings, and they will be erased when the transcripts have been made and you have had a chance to verify their accuracy. You can choose not to be recorded, and you have the option of stopping the recording at any time.

RISKS

No risks are anticipated in this study.

BENEFITS

Potential benefits of this study include an increased awareness of the factors involved in decisions of how to prepare students for state science assessments. This awareness could benefit you or other teachers in similar situations by making them more aware of how they make instructional decisions.

PAYMENT TO PARTICIPANTS

You will not be paid for your participation.

PARTICIPANT CONFIDENTIALITY

Your name will not be associated in any publication or presentation with the information collected about you or with the research findings from this study. Instead, the researcher(s) will use a study number or a pseudonym rather than your name. Your identifiable information will not be shared unless required by law or you give written permission.

Permission granted on this date to use and disclose your information remains in effect indefinitely. By signing this form you give permission for the use and disclosure of your information for purposes of this study at any time in the future.

REFUSAL TO SIGN CONSENT AND AUTHORIZATION

You are not required to sign this Consent and Authorization form and you may refuse to do so without affecting your right to any services you are receiving or may receive from the University of Kansas or to participate in any programs or events of the University of Kansas. However, if you refuse to sign, you cannot participate in this study.

CANCELLING THIS CONSENT AND AUTHORIZATION

You may withdraw your consent to participate in this study at any time. You also have the right to cancel your permission to use and disclose further information collected about you, in writing, at any time, by sending your written request to:

Chris Klager
100 N. McDaniel
Bonner Springs, KS 66012

If you cancel permission to use your information, the researchers will stop collecting additional information about you. However, the research team may use and disclose

information that was gathered before they received your cancellation, as described above.

QUESTIONS ABOUT PARTICIPATION

Questions about procedures should be directed to the researcher(s) listed at the end of this consent form.

PARTICIPANT CERTIFICATION:

I have read this Consent and Authorization form. I have had the opportunity to ask, and I have received answers to, any questions I had regarding the study. I understand that if I have any additional questions about my rights as a research participant, I may call (785) 864-7429 or (785) 864-7385, write the Human Subjects Committee Lawrence Campus (HSCL), University of Kansas, 2385 Irving Hill Road, Lawrence, Kansas 66045-7568, or email irb@ku.edu.

I agree to take part in this study as a research participant. By my signature I affirm that I have received a copy of this Consent and Authorization form.

_____ Date
Type/Print Participant's Name

Participant's Signature

I agree to be videotaped while being interviewed and while my classroom is observed as part of this research.

Participant's Initials

Researcher Contact Information

Chris Klager
Principal Investigator
100 N. McDaniel
Bonner Springs High School
Bonner Springs, KS 66012
316 841-6048

Dr. Jim Ellis
Faculty Supervisor
346 Joseph R. Pearson Hall
University of Kansas
Lawrence, KS 66045
785 864-9847

Appendix F

Coding Agenda

Category	Definition	Example	Coding Rules
C1: Methods of test preparation	Any practice implemented with the intention of increasing score on the state assessment beyond 'normal' classroom instruction.	Study Guide Practice Test Review Game Etc.	Must include some sort of action, something actually has to be done. Cannot just be thinking about test prep – see C2, C5
C2: Test preparation ethics	Issues related to whether or not a test preparation practice is morally or professionally acceptable.	How much preparation is acceptable? Is this particular practice okay?	Must relate to moral or professional acceptability.
C3: Pressure on teachers	Factors that cause teachers to worry or stress about state assessments.	Repercussions for low scores Loss of instructional time Trying to fit all of the tested content in on time.	Must cause anxiety for teacher.
C4: School Culture	Shared beliefs inside of the school.	Positive or negative feelings about testing held by the faculty	Must be shared beliefs, widespread throughout school or segment of school, ex. Administrators
C5: Uncertainty about tests	Any ambiguity where the teacher is unclear on what will happen or what should happen.	What will the test look like? What am I allowed to do with the test?	Must be related to questions the teachers still has unanswered related to tests.