

ABSTRACT

Title of Dissertation: NEGOTIATING THE TERRAIN OF HIGH-STAKES ACCOUNTABILITY IN SCIENCE TEACHING

Isaak Aronson, Doctor of Philosophy, 2007

Dissertation Directed By: Professor Francine H. Hultgren
Department of Education Policy Studies

Teachers interact with their students on behalf of the entire educational system. The aim of this study is to explore how biology teachers understand and construct their practice in a high-stakes accountability environment that is likely to be riddled with tensions. By critically questioning the technical paradigms of accountability this study challenges the fundamental assumptions of accountability. Such a critical approach may help teachers develop empowerment strategies that can free them from the de-skilling effects of the educational accountability system.

This interpretive case study of a high-school in Maryland is grounded in three streams of research literature: quality science instruction based on scientific inquiry, the effects of educational accountability on the curriculum, and the influence of policy on classroom practice with a specific focus on how teachers balance competing tensions. This study theoretically occurs at the intersection of educational accountability and pedagogy. In terms of data collection, I conduct two interviews with all six biology

teachers in the school. I observe each teacher for at least fifteen class periods. I review high-stakes accountability policy documents from the federal, state, and district levels of the education system.

Three themes emerge from the research. The first theme, “re-defining science teaching,” captures how deeply accountability structures have penetrated the science curriculum. The second theme, “the pressure mounts,” explores how high-stakes accountability in science has increased the stress placed on teachers. The third theme, “teaching-in-between,” explores how teachers compromise between accountability mandates and their own understandings of quality teaching. Together, the three themes shed light on the current high-stakes climate in which teachers currently work.

This study’s findings inform the myriad paradoxes at all levels of the educational system. As Congress and advocacy groups battle over the reauthorization of *No Child Left Behind*, they may not pay adequate attention to all the inconsistencies. Educators and researchers must take a critical look at accountability policies. Accountability should promote optimism, responsibility, job satisfaction, avenues for developing pedagogical expertise, and collaboration between teachers and administrators. Only then is it likely to improve educational opportunities for all students.

NEGOTIATING THE TERRAIN OF HIGH-STAKES ACCOUNTABILITY IN
SCIENCE TEACHING

By

Isaak Aronson

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:
Professor Francine H. Hultgren, Chair
Professor Steven Selden
Professor Betty Malen
Assistant Professor Janet Coffey
Professor Randy McGinnis

© Copyright by
Isaak Aronson
2007

DEDICATION

I would like to dedicate this work to my family whose support has made this possible. I could not have achieved this without you.

ACKNOWLEDGEMENTS

As I come to the end of this journey, I am astounded by my own transformation. My view of schools, teaching, and the education has been reshaped by my graduate studies and the dissertation process. My experience at the University of Maryland will forever inform my professional role as an educator, but, perhaps more importantly, it has informed my own sense of place in this world. I am forever grateful to my professors and colleagues for facilitating this transformation.

Thank you, Francine. From my very first class with you, I began to look at education and my role as a teacher from an entirely different perspective. You have been an amazing teacher, a thoughtful and patient advisor, and a diligent guide who has brought me to this final stage of the doctoral process. I most certainly would have still been mired in Chapter VI without you! Thank you, Steve for illuminating the social inequalities in which our system of education is situated and teaching me how to understand my own place within the system. Thank you, Randy for navigating me through the science education literature and facilitating an introduction that led to my first publication in education. Thank you, Betty for being my compass. Your guidance helped keep me on a logical path when I became lost in my own research and writing. Thank you, Janet for helping me find my way through this research. Our meetings at Starbucks kept me focused and on track, both academically and professionally.

Thank you to all the teachers who participated in this research. From our very first lunch meeting, you welcomed me into your school. Through this research, I quickly learned how busy and pressed for time you were. Nevertheless, you always made time for

me and my questions. Thank you to the school district for giving me permission to do my work. I am forever grateful for this opportunity.

Last, but certainly not least, thank you to my family. You have always believed in me even when I did not believe in myself. You have encouraged and supported me throughout all of my endeavors. I love you all very much. I want to express special thanks to my mother, Vita, who helped transcribe all of the interviews and served as an editor for this work. I would not have been able to graduate this semester without your help. Finally, thank you to my wife, Hallie, my physical and emotional support team. Thank you for picking up the slack at home when I was too busy, and, more importantly, thank you for your unwavering support and confidence. I love you so much and could not imagine a better wife and friend.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION.....	1
Competing Horizons in a Curriculum Policy.....	1
An Overview of Accountability.....	5
A National Surge of Accountability.....	6
Science (Accountability) for All.....	14
Accountability in Maryland.....	19
Accountability in Buckley County.....	26
Teaching Accountability.....	30
Overview of the Study.....	34
Purpose of the Study.....	34
Research Questions.....	37
Overview of Methodology.....	38
Theoretical Foundations.....	40
Summary.....	43
CHAPTER II: REVIEW OF THE LITERATURE.....	45
Overview.....	45
Scientific Inquiry-Based Instruction.....	46
What is Scientific Inquiry?.....	47
The Value of Scientific Inquiry in the Classroom.....	54
Limitations of Scientific Inquiry-Based Instruction.....	57
Assessing Scientific Inquiry.....	62
Scientific Inquiry and Accountability.....	66
The State of Accountability.....	67
Curriculum and Test Alignment.....	71
Teaching to the High-Stakes Test.....	74
An International Perspective.....	78
Reducing the Definition of Quality Teaching.....	81
Science Education Accountability.....	83
In-cluding Policies in Teaching.....	84
Defining the Teacher.....	90
Power Struggles.....	92
Aligning Policy with Pedagogy.....	94
Summary.....	97
CHAPTER III: METHODOLOGY.....	99
Introduction to the Study.....	99
Purpose of the Study.....	100
Research Questions.....	101

Methodological Grounding for Case Study Research.....	104
A Critical Perspective.....	107
The Design of the Study.....	109
The Boundaries of the Case.....	113
Sampling.....	116
Data.....	121
Document Analysis.....	122
Interviews.....	124
Observations.....	128
Data Collection and Interpretation.....	130
Strengths and Limitations.....	134
Summary.....	138
 CHAPTER IV: THE SETTING.....	 140
Introduction.....	140
Halbert High School.....	142
The Biology Team.....	143
Teachers.....	146
Ms. Victoria.....	146
Ms. Lydia.....	149
Ms. Khana.....	151
Ms. Calypso.....	152
Dr. Stevens.....	155
Ms. Harris.....	156
Overview of Themes.....	158
 CHAPTER V: RE-DEFINING SCIENCE TEACHING.....	 160
Biology Reigns Supreme.....	160
Environmental Science is Not Biology.....	160
Student Differences.....	166
Pedagogical Differences.....	168
The Changing Curricular Landscape.....	175
An Emphasis on Inquiry.....	176
Mixed Messages.....	184
Disposing of Scientific Inquiry.....	189
Footprints on the Curriculum.....	196
Accountability Lexicon in the Classroom.....	196
Pulling the Curriculum Apart.....	200
Standardized Test Prep.....	203
Summary.....	208
 CHAPTER VI: THE PRESSURE MOUNTS.....	 210
Monitoring Teachers.....	210

No Dentist Left Behind.....	211
Powerless.....	212
Undervalued.....	216
Teaching in the Accountability Vice.....	223
The “Highly-Qualified” Teacher.....	226
The Breakdown of Teaching and Learning.....	229
Time.....	236
The Intensification of Teaching.....	239
Cutting Corners.....	239
Hiding Behind the Structures of Accountability.....	240
An Isolated Community.....	242
The Psyche of Accountability.....	244
Frustration.....	245
Passion and Teaching.....	247
Dividing Teachers.....	250
Summary.....	253
CHAPTER VII: TEACHER-IN-BETWEEN.....	255
In-Between Values.....	257
Avoidance.....	258
Submission.....	265
In-Between Skills.....	273
Competition.....	274
The Eight Hundred Pound Guerilla in the Classroom.....	284
In-Between Pedagogy.....	289
Not Artistic Professionals.....	289
Understanding Pedagogy Within Accountability.....	292
Summary.....	298
CHAPTER VIII: RECONCILING COMPETING HORIZONS.....	300
Re-Viewing Horizons.....	300
Tensions Embedded in Accountability Policies.....	302
Teachers’ Perceptions of Accountability Messages.....	302
Teachers’ Interpretations Shape Their Teaching Practice.....	311
Teaching Adjustments.....	312
Understanding of Science.....	313
Teachers’ Role in the Classroom.....	315
The Passion for Teaching.....	318
The Paradoxes of Accountability.....	321
Must Accountability Fail?	322
The Paradoxical Response.....	329
The Ramifications of Paradox.....	331
Implications of Teachers’ Negotiations Through Accountability.....	335
Reframing the Accountability System.....	336

Researching the Accountability System.....	339
The Future of Accountability.....	343
Summary.....	348
APPENDICES.....	350
Appendix A: Letter of Invitation.....	350
Appendix B: Data Collection Timeline.....	351
Appendix C: Document Review Guide.....	352
Appendix D: Informed Consent Form.....	353
Appendix E: Interview Guide.....	355
Appendix F: No Dentist Left Behind.....	361
REFERENCES.....	364

CHAPTER I: INTRODUCTION

Competing Horizons in a Curriculum Policy

Miss O in-dwells between two horizons – the horizon of the curriculum-as-plan as she understands it and the horizon of the curriculum-as-lived experience with her pupils. Both of these call upon Miss O and make their claims on her. She is asked to give a hearing to both simultaneously. This is a tensionality within which Miss O inevitably dwells as a teacher. And she knows that inevitably the quality of life lived within the tensionality depends much on the quality of the pedagogic being that she is. Here the “third” space, the space in-between, has entered Aoki’s theorization and this theorization will become ever more prominent in the final phase of Aoki’s *oeuvre*. (Aoki, as cited in Pinar, 2005, p. 15)

Teachers in the present accountability climate could be described as negotiating the tensions between these competing curriculum horizons. A central intention of the *No Child Left Behind Act of 2001* is to “hold... [teachers] accountable for improvements in student academic achievement” (United States Congress, 2002a, statute 1620). Although the grand purpose of the legislation, to ensure that *all* students reach academic proficiency by 2014 (United States Congress, 2002b), may be noble, numerous competing horizons have emerged for teachers as a side effect of the policy. *No Child Left Behind* aligns with Tyler’s (1949) technical paradigm of curriculum-as-plan, where teachers implement, with as little interference as possible, what the state outlines as knowledge all students at a particular grade level must possess. From the teachers’ perspective, standards are established by an absolute, non-negotiable authority that is inaccessible to them (Apple, 2006). The locus of control is outside their being because *No Child Left Behind*’s technical paradigm largely ignores the curriculum-as-lived-experience. As a result, teachers may feel torn between their own understandings of quality teaching and the state mandated prescriptions of what they should teach.

Cawelti (2006) outlines three horizons of tension embedded in the implementation of high-stakes accountability policies: a skewed curriculum, professional dissatisfaction, and reduction in the definition of learning. First, although teachers understand that the curriculum must reflect the needs of society, many teachers are compelled to narrow the curriculum as a result of the pressure to show Adequate Yearly Progress on high-stakes tests in reading and math (Cawelti, 2006; Hargrove, Jones, Jones, Hardin, Chapman & Davis, 2000). Second, teachers report that high-stakes accountability forces them to teach to the test, neglect individual students' needs, reduce creativity in the classroom, and bore themselves and their students with practice problems geared toward test preparation (Centolanza, 2004). Further, high-stakes test results may discourage teachers who work hard, yet fail to help students with special needs and limited English proficiency reach the state's definition of proficiency (Cawelti, 2006; Wright, 2006). Finally, *No Child Left Behind* drastically narrows the definition of academic success to single goals and measures. Thus, teachers may be encouraged to tweak the curriculum rather than make holistic changes that they may think are best for their students (Cawelti, 2006). These competing horizons highlight the tensions that teachers may experience in the classroom as they struggle to "fit" *No Child Left Behind* mandates into their own pedagogical practices.

The teachers' struggles in the classroom can be understood by considering the meaning of curriculum. According to *No Child Left Behind*, curriculum is considered to be environment-producing, a construction that encompasses the setting of objectives, classroom design, classroom management, and other atmospheric conditions that influence the transmission of knowledge that is generated from outside the classroom for

implementation by teachers in the classroom (Pinar, 2002). However, Pinar suggests that curriculum is both an environment-producing discipline and a knowledge-producing discipline, in which classroom experiences and interactions *are*, in and of themselves, much of the knowledge that is learned in the classroom. The knowledge is living such that it can be experienced by both teachers and students (Jardine & Rinehart, 2003). This reconceptualization of curriculum beyond the technical paradigm encompasses a more complete understanding of what goes on in the classroom (Pinar, Reynolds, Slattery, & Taubman, 2002). It gives meaning to the classroom tensions around implementing the policy's agenda and the lived knowledge and experiences produced by the teacher and students. Because *No Child Left Behind* attempts to exclude lived knowledge from its measures of student proficiency and, by extension, the teachers' sense of quality teaching, teachers are torn between two competing constructions of the curriculum.

In other words, teachers may be stuck between their own constructions of good teaching and curriculum and *No Child Left Behind's* traditional technical perspective of curriculum implementation. Teachers' understandings of their own profession are influenced by their professional experience, education, and status (Olsen & Kirtman, 2002). Further, schools and school systems are contextualized entities in which specific types of knowledge and understandings are embedded (Clandinin, 1986). Teachers' beliefs, understandings, and construction of their own roles are strongly influenced by their surrounding culture. Their identities are subject to the social control of the institutions in which they work (Bidwell & Yasumoto, 1999), and their relationships with other teachers and administrators (Bryk & Schneider, 2002). At the same time accountability policies seek to influence and to reform the teachers' educational practice

(Cohen & Ball, 1990; Glasnapp & Poggio, 1991). Thus, teachers are likely to undergo complex internal negotiations as they engage in curricular practice. Such a process is likely to be riddled with tensions, contradictions, and inconsistencies.

Thus far, science teachers have been partially inoculated from the more severe strains of educational accountability. Academic proficiency has been defined and measured by state mathematics and language arts standards and high-stakes assessments. For science educators, however, the respite has come to a rapid end. *No Child Left Behind* requires states to develop standardized science tests by the academic year 2007-2008. Adequate Yearly Progress (AYP), which is a measure of a school's progress toward achieving the legislation's goal of "academic proficiency" for every student by the year 2014, may then include the results of the science assessments (United States Congress, 2002a). Failure to make AYP can result in serious consequences for schools, teachers, and students (Flynn, 2002).

The aim of this study is to explore how secondary school biology teachers understand and construct their practice in a high-stakes accountability environment that is likely to be riddled with tensions.¹ De Lissovoy and McLaren (2003) argue that the proliferation of accountability policies, rules, and regulations are crowding out all other discourses in education. I, however, am not studying the hegemonic structures of accountability from the outside; rather, I examine its mechanisms critically from its center by exploring teachers' experiences with it. Teachers offer an *emic* point of view of policy implementation that is embedded in the school's and classroom's culture. Understanding teachers' experiences and the perspectives they hold regarding their work requires a rigorous examination of the possible conflicting messages, implicit and

¹ Each time I refer to biology teachers in this study, I describe secondary school biology teachers.

explicit, present in the policy, how teachers understand these tensions, and how their perceptions of accountability affect the construction of their practice. Teachers' beliefs, knowledge, and experiences are important factors that inform their interactions with students and the pedagogical and science content choices that they make in the classroom (Craig, 2004). By examining science educational accountability through the lens of teachers, I gain a deeper understanding of its effects on the lives of teachers and students.

An Overview of Accountability

How did the present accountability climate come about? How do teachers find themselves within the present educational climate, riddled with accountability discourse – standards, benchmarks, and other competency measures? How has educational accountability become institutionalized in the United States at both the state and national levels? In order to understand accountability in the United States, it is important to understand its guiding forces.

Elmore and Sykes (1992) discuss the importance of tracing “the path of decisions leading to a particular law or regulation” (p. 188). Policies do not simply emerge out of nowhere. Policy decisions are preceded by a process that is imperfect and not necessarily rational. Rather than being analyzed comprehensively, policies enter public discussions on the margins in response to a particular political uproar. Such a process may be riddled with contradictions and inconsistencies. An overview of the process in which a policy is formed is integral to understanding what signals it sends to its recipients and the context in which the messages are received. *No Child Left Behind* is a curriculum policy and therefore, like all policies, needs to be examined in light of its formative roots.

A National Surge of Accountability

Our Nation is at risk. Our once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world. This report is concerned with only one of the many causes and dimensions of the problem, but it is the one that undergirds American prosperity, security, and civility. We report to the American people that while we can take justifiable pride in what our schools and colleges have historically accomplished and contributed to the United States and the well-being of its people, the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people. What was unimaginable a generation ago has begun to occur – others are matching and surpassing our educational attainments. (National Commission on Excellence in Education, 1983, ¶ 1)

Like other major policies, science educational accountability as mandated by *No Child Left Behind* is the byproduct of the convergence of several initiatives that came to fruition over the past decades. Historically, accountability measures have been popular with the general public, and as a result, with politicians (Mehrens, 1998). Politicians have blamed schools for the perception of American inferiority on the international landscape since the Soviet launching of the satellite Sputnik in 1957. The ensuing curriculum reform movement of the 1960s was intended, in part, to help the U.S. regain its perception of technological superiority in the Cold War. “Accountability,” “assessments,” and “behavioral objectives” became the central themes toward which U.S. national educational reform efforts turned (Eisner, 1995; Stake, 1998). Further, in the 1970s, at the state level, significant efforts to increase accountability through minimum competency assessments emerged to add fuel to the accountability movement (McDonnell, 1994).

However, the prominent culture and discourse of education today – based on the language of standards, benchmarks, and proficiency – are rooted in the standards and testing movement which was galvanized in 1983 with the publication of a *Nation at Risk*

by the National Commission on Excellence in Education. The publication was a scathing admonition of the public education system in the United States. Arguing that the U.S. was losing ground to foreign countries due to serious flaws in the education system, the manuscript dealt a serious blow to the public's faith in the public education system (Rudelvidge, 2003). The implication of this document and the standards movement that ensued from it was that until then teachers and teacher educators had no standards (Pinar, 2005).

The Nation at Risk listed thirteen "indicators of risk" that demonstrated how far behind U.S. students were in language arts, math, and science. Simultaneously, the commission identified several indicators of excellence that focused on challenging college entrance requirements, high school graduation standards that are tested with an exit exam, challenging school subject matter, and standardized testing to measure achievement at transition points from one level of schooling to another (National Commission on Excellence in Education, 1983). This coupling of fear of national demise with the remedy of higher educational standards galvanized the policy drive toward the current accountability system (Walberg, 2003)

The report quickly spawned a flurry of school improvement initiatives under the rubric of an "excellence movement" (Amrein & Berliner, 2002; Resnick, 2006). Over 300 national and state task forces were set up to investigate the condition of public schooling in the United States. The research and ensuing reports provided support for ambitious reforms that aimed to improve student performance drastically as measured by standardized assessments. Through the 1980s, many states imposed centrally developed content standards on local districts, implemented minimum competency standardized

testing, increased graduation requirements, and tightened certification requirements for teachers (Eisner, 1995). By the end of the 1980s the educational accountability movement had taken root in the U.S. political discourse, and the model of standardized testing to measure attainment had emerged (Dufour & Eaker, 1992).

Most of the reforms were top-down and excluded educators from the decision-making process (Apple, 1996; Finley, 2000; Mehrens, 1998; Ruppert, 1994; Strike, 1997). Apple (1996) argues that they were based on little understanding of the daily lives of teachers and the already intensified conditions under which they work. The meanings of excellence, achievement, and success which provided the foundation for the reforms were inadequately addressed (Amrein & Berliner, 2002). Further, such standards-based reforms may narrow the curriculum to exclude considerations of difference between localities and students, devalue individual classroom experiences and meaning making, diminish pedagogical innovation, and limit community involvement in education (Eisner, 1995).

In 1989, President George H.W. Bush called an “education summit” with the nation’s governors to address the failures of the decade’s reforms. The governors called for systemic reform, which meant aligning their vision of the chief components of education: academic goals, curricula, instruction, and exams (Walberg, 2003). The summit produced broad performance goals for public schools, including standardized national and state testing (Rudelvidge, 2003). The conference led to the development of the National Council on Education Standards and Testing in 1991 and to the passage of *America 2000*, which included six objectives for schools and students (Wraga, 1999). Under President Clinton, the plan became known as *Goals 2000* which provided

unconditional grants for states to infuse accountability into the system of public education and to develop academic standards (U.S. Congress, 1994a). Although many decisions and responsibilities were left to states and localities, the federal mandates provided the means for spreading centrally conceived accountability measures across the country.

By the 1990s, performance of schools according to indicators that the *Nation at Risk* report cited as evidence of a failing school system essentially remained stagnant, despite the flurry of reform efforts (Amrein & Berliner, 2002). The Report of the Secretary's Commission on Necessary Skills found that the reform initiatives of the 1980's based on trying to get greater results through "tighter curricula, higher certification standards, and more testing of everyone" were largely unsuccessful (U.S. Department of Labor, 1992, p. xvi). Rather than debunking the accountability movement, poor outcomes reinvigorated the policy makers to increase accountability measures (Eisner, 1995). A major political change occurred when Congress reauthorized the Elementary and Secondary Education Act of 1965 (Rudelvidge, 2003). The present understanding of standards and benchmarks emerged from the new law.

The *Improving America's Schools Act* required states to develop content and performance standards for all K-12 schools. States were mandated to make "continuous and substantial" progress towards "proficiency" for all students. Progress would be monitored by regular standardized assessments (U.S. Congress, 1994b). Congress provided no specific deadlines for making progress, nor did it mandate consequences for non-compliance. The legislation, however, did provide the 1997-1998 school year as a deadline for developing state standards, and the 2000-2001 school year as a deadline for implementing standardized assessments. A case study of ten states indicated that the

adoption of state-level standards most often was done rapidly, haphazardly, and was driven by top-down legislative initiatives that excluded schools and teachers from the process (Ruppert, 1994). This report concluded that few states have accomplished the educational analysis necessary to define measures appropriate for systemic decision making and public reporting.

State standards were intended to result in eight broad national education goals by the year 2000. In 1994, Clinton's *Goals 2000* added two goals to the six initially developed in 1989 at the summit called by President Bush. The legislation stated that by the year 2000:

- 1) All children in America will come to school ready to learn.
- 2) The high school graduation rate will increase to at least ninety percent.
- 3) All students will leave grades 4, 8, and 12 having demonstrated competency over challenging subject matter including English, mathematics, science, foreign languages, civics and government, economics, arts, history, and geography, and every school in America will ensure that all students learn to use their minds well, so they may be prepared for responsible citizenship, further learning, and productive employment in our Nation's modern economy.
- 4) The nation's teaching force will have access to programs for the continued improvement of their professional skills and the opportunity to acquire the knowledge and skills needed to instruct and prepare all American students for the next century.
- 5) United States students will be first in the world in mathematics and science achievement.
- 6) Every adult American will be literate and will possess the knowledge and skills necessary to compete in a global economy and exercise the rights and responsibilities of citizenship.
- 7) Every school in the United States will be free of drugs, violence, and the unauthorized presence of firearms and alcohol and will offer a disciplined environment conducive to learning.
- 8) Every school will promote partnerships that will increase parental involvement and participation in promoting the social, emotional, and academic growth of children. (United States Congress, 1994a, pp. 13-14)

Goals 2000 appears to be little more than a broad outline for the future of U.S. public education. Its specific mandates would be difficult to implement. Seidman (1996) argues that *Goals 2000* is doomed for failure because the system of education in the United States is complex and difficult to manipulate. He writes, “The system of education is a vast and complex enterprise comprising all of the many and different ways society educates its citizens. It is useful to distinguish it from the educational system which possesses a logic and laws of behavior of its own and which can be shown to be highly intractable to attempts to reform it by education policy” (p. 1). *Goals 2000*, like many other education policies, generally focuses on the latter system in order to make sweeping adjustments to the whole of society.

Further, Seidman (1996) predicts two unintended consequences of the legislation that would stem from the added socioeconomic inequality created between those students who reach the stated goals and those who do not. First, he suggests that those who do not graduate from high school will be shut out of important non-educational social benefits. Second, he speculates that social benefits will be reduced for those who do complete high school. By not directly addressing social inequalities and injustices, the act is likely to reinforce them (Waters, 1998). At the classroom level, the legislation may reduce intellectual liberty and diversity by limiting students’ ability to make choices about their own education, teachers’ intellectual freedom, and pluralism (Strike, 1997).

Goals 2000 and the *Improving America’s Schools Act* laid out most of the foundation for *No Child Left Behind*. One very important caveat, however, was still missing. The two acts had no mechanism to ensure that states comply with federal mandates (Walberg, 2003). Penalties for failing to achieve the timelines and directives

were absent from federal mandates. Thus, it is likely that political, rather than educational or social, purposes drove states and districts to comply with the legislations' calls for reform (Mehrens, 1998).

States responded to federal calls for accountability. In March 1996, the nation's governors met in Palisades, New York, and called for an "external, independent, nongovernmental effort" to measure and report on each state's annual progress in raising student achievement and improving public schools (Olson, 2006, p.15). Accountability may require simultaneous centralization and decentralization: the centralization of standards at the state level and the decentralization of operational responsibilities to the district or school level. State policymakers were to set goals and measure progress, while local school districts and schools are supposed to develop and execute effective practices (Walberg, 2003). Such an arrangement seems ripe for creating contradictions at the classroom level due to a potentially adversarial relationship between the state and school districts.

In 2002, with the passage of *No Child Left Behind*, the latest reauthorization of the *Elementary and Secondary Schools Act of 1965*, for the first time stipulated real consequences and requirements for the use of federal funds (Flynn, 2002; Walberg, 2003). Specifically, the act required states and local departments of education to:

- 1) Define "academic proficiency" with standardized testing and one other indicator, determined by the states.
- 2) Set objectives to meet annual benchmarks to reach universal proficient achievement by 2014.
- 2) Ensure that students are meeting these objectives with annual testing. All students must be tested by 2005-2006 in grades three through eight in math and reading; science testing must begin by 2007-2008.
- 3) Implement a series of interventions for schools and students who are not making Adequate Yearly Progress. (Flynn, 2002, p. 2)

Funding was now linked to progress toward standards because schools and districts were required to finance specific measures if they failed to make Adequate Yearly Progress (AYP). AYP is the factor that determines which schools are labeled “in need of improvement.” School districts and schools that fail to make AYP toward statewide proficiency goals are “subject to improvement, corrective action, and restructuring measures aimed at getting them back on course to meet state standards” (United States Congress, 2002b, p. 1). From the perspective of many schools and teachers, these interventions are punitive because schools can lose funding and be forced to spend otherwise needed money on improving test scores, and staff may be transferred due to school restructuring or reconstitution (Flynn, 2002).

In terms of pedagogy, since teachers are directly involved with student learning, classroom practice has become more scrutinized by the public due, in part, to the proliferation of school and district data that have emerged from the mandates of *No Child Left Behind* (Jones & Egley, 2004). Further, many schools and districts have responded to high-stakes accountability by being more prescriptive regarding what is taught and how it is taught (Center on Education Policy, 2006). In science education, for example, accountability systems have been shown to shift the primary purpose of education from underlying structure and process of subject matter to the acquisition of skills and facts that are out of context (Wood, 1988).

In the next section, I outline how the accountability movement in education became conflated with the movement to achieve universal scientific literacy for all Americans. I demonstrate how the accountability movement engulfed the scientific

literacy movement to create a massive push for science educational accountability at the national and state levels.

Science (Accountability) for All

All of us have a stake, as individuals and as a society, in scientific literacy. An understanding of science makes it possible for everyone to share in the richness and excitement of comprehending the natural world. Scientific literacy enables people to use scientific principles and processes in making personal decisions and to participate in discussions of scientific issues that affect society. A sound grounding in science strengthens many of the skills that people use every day, like solving problems creatively, thinking critically, working cooperatively in teams, using technology effectively, and valuing life-long learning. And the economic productivity of our society is tightly linked to the scientific and technological skills of our work force. (NRC, 1996, p. vii)

Science education in the United States has been on the forefront of public policy developers' agendas at least since the early days of the Cold War (Eisner, 1995). Achievement of universal scientific literacy has been a goal in the United States since, at least, the late 1950s with the Soviet Union's launching of the shuttle, Sputnik (DeBoer, 2000; Gabel, 2003; Lageman, 2000; Viadero, 2006). Since then, the U.S. has clamored to make science education more effective at attracting more students to a career in science (Bruner, 1960; Cross & Cross, 2005; Finley, 2000; Ravitch, 1995), and fostering non-scientist adults who retain a degree of scientific competency throughout their lives (Arons, 1983; Bybee, 1997; Fensham, 1992; Klopfer, 1969). During nearly fifty years of effort, some gains have been made in scientific literacy, but the engineering and science fields fail to attract and retain students, and science content remains at the fringes of popular public discourse (Viadero, 2006). Most of the curricular reforms were top-down, developed by professional scientists and science educators with little involvement from schools and teachers (Finley, 2000). By the 1970s, the federal support for science

education had somewhat waned to make way for the state led minimum competency testing movement (McDonell, 1994).

In the 1983 publication of the *Nation at Risk*, the National Commission on Excellence in Education issued a dire warning: The United States “once unchallenged pre-eminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world” (p. 5). The educational accountability movement of the 1980s and 1990s provided additional fodder for those who called for universal scientific literacy for all Americans (Collins, 1995). As it was conceived, accountability in science education could ensure that all students achieve a particular degree of competence in science before they leave school (Ravitch, 1995). The presumption is that the more students know when they leave school, the more they will retain throughout their lives. Thus, many organizations, like the American Association for the Advancement of Science and the National Research Council, that were working toward expanding scientific literacy embraced the accountability movement (Collins, 1995; Leonard & Penick, 2005).

In response to this conflation of movements and national calls for standards, scientists and science education leaders began to push for the development of national standards for science (Bybee & Ferrini-Mundy, 1997; Collins, 1995; Finley, 2000). Standards in science education emerged from two national organizations, the American Association for the Advancement of Science and the National Research Council. The first widely publicized standards were the *Benchmarks for Science Literacy* (AAAS, 1993), and the next came three years later with the *National Science Education Standards* (NRC, 1996). Both documents grew out of a synthesis of research on learning, curriculum, and

science instruction (Collins, 1995; Fuhrman & Massell, 1992; Oakes, 1987; Sarason, 1990). They present a vision of a scientifically literate populace by outlining what students need to know, understand, and be able to do by high school graduation. Both documents outline standards for teaching science, professional development of teachers, assessment, science content, school programs, and state and local systems (Bybee & Ferrini-Mundy, 1997; Leonard & Penick, 2005). In other words, the explicit purpose of the documents is to provide states, districts, and schools with a research-based model of science instruction that can be used to improve science teaching and learning.

Both the *Benchmarks for Science Literacy* (AAAS, 1993) and the *National Science Education Standards* (NRC, 1996) were funded by large grants from the National Science Foundation, an independent U.S. federal government agency responsible for promoting science (Collins, 1995). Both documents provide a detailed definition of scientific literacy, the “knowledge and understanding of scientific concepts and processes required for personal decision making, participation in civic and cultural affairs, and economic productivity” (NRC, 1996, p. 22). They share the following general themes:

- 1) They emphasize the education of all students.
- 2) They emphasize student comprehension, rather than mere memorization of a series of facts or set of procedures.
- 3) They recommend developing a depth of knowledge about fundamental mathematical and scientific content and processes.
- 4) They emphasize content more than curriculum. That is, the documents do not define the order, structure, and organization of science programs. These decisions are left to states and local school districts.
- 5) They recommend a comprehensive, coherent, and integrated approach to science education. (Bybee, 1997; Bybee & Ferrini-Mundy, 1997)

In terms of how science is taught, both documents recommend that students learn science through the scientific inquiry process in order for them to understand how to do science,

not just to understand its findings (Leonard & Penick, 2005). Both documents list exactly which skills each student should learn.

According to the documents, scientific inquiry is central to learning science effectively. The National research Council (1996) defines scientific inquiry as

...a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations. (p. 23)

The American Association for the Advancement of Science (1993) defines scientific inquiry as

...more complex than popular conceptions would have it. It is, for instance, a more subtle and demanding process than the naive idea of “making a great many careful observations and then organizing them.” It is far more flexible than the rigid sequence of steps commonly depicted in textbooks as “the scientific method.” It is much more than just “doing experiments,” and it is not confined to laboratories. More imagination and inventiveness are involved in scientific inquiry than many people realize, yet sooner or later strict logic and empirical evidence must have their day. Individual investigators working alone sometimes make great discoveries, but the steady advancement of science depends on the enterprise as a whole.
(<http://www.project2061.org/publications/bsl/online/ch1/ch1.htm#B>, ¶ 31)

In other words, the documents call students to learn both scientific content knowledge and the process skills that are involved in doing science.

The *Benchmarks for Science Literacy* and the *National Science Education Standards* are non-binding, and they are not officially linked to any state or local department of education (AAAS, 1993; NRC, 1996). Nevertheless, most state departments of education claim that they were integral to the development of state

science standards (Loucks-Horsley & Bybee, 1998). *Goals 2000* mandated that all states develop specific proficiency standards for all subject areas in grades K-12. In response, states created competency goals and objectives for teaching and learning science (Collins, 1995; Finley, 2000). The standards contain the concepts and theories, strands, skills, and processes upon which all science instruction should be based (AAAS, 1993; NRC, 1996). In addition, state science curricula explicitly define and illustrate the connections between the *National Science Education Standards*, the *Benchmarks for Science Literacy*, and state science education standards (McFadden, 2001). Maryland claims to have adhered to both documents strictly for developing both its science standards and assessments (MSDE, 2001).

Policy development and implementation do not always match policymakers' stated goals. The literature on public policy distinguishes between policies as statements of intent and policies as actions (Elmore & Sykes, 1992; Ripley, 1985). This distinction suggests that policy statements are about what ought to be and policy actions are what actually happen. An examination of policy implementation reveals that "what government does [is] distinguished from what it says it is going to do, sometimes with many conflicting and unclear voices" (Ripley, 1985, pp. 40-41). Thus, it is likely that the *Standards* and *Benchmarks*, and the messages embedded within them, may not, in fact, have been integral to the development of state standards (despite claims made by states), and state accountability systems are likely to send vague and conflicting messages to their recipients.

In many cases, the development of the standards excluded teachers (Finley, 2000), and the documents subscribe to a technical curricular perspective where teachers merely

deliver standards to the students (Pinar, Reynolds, Slattery, & Taubman, 2000). Consequently, teachers are left to implement passively an outsider's curriculum. However, there may be an inherent contradiction between what the *Standards* and *Benchmarks* espouse and how they intend science education to be implemented. Teacher involvement and integration into the curriculum are integral to successful scientific inquiry-based instruction (Hammer, 1997). Teachers facilitate, question, and entice students to further their development of knowledge and skills. As such, it may be inappropriate to exclude them from the development of the curriculum. If teachers were largely excluded from the process, are calls for scientific inquiry-based instruction embedded in the documents simply a façade for further propagation of the technical curricular perspective? Are teachers expected to ignore this contradiction? Thus far, research shows that accountability measures have driven some teachers into complacency with fear (Craig, 2004; McNeil, 2000; Pringle & Carrier-Martin, 2005). What if teachers were to become empowered? What if teachers “become as curious about the *productiveness* of our continuously remodeled ignore-ances, lacks of fit, and limitations of knowing as we have been about how to achieve full and complete [competency]?” (Ellsworth, 1997, p. 53, emphasis in the original).

Accountability in Maryland

A few years after leaders at the national level began advocating for educational accountability measures, states began to respond. They acceded to national calls by accepting funding to align state curriculum frameworks with standardized tests (Wraga, 1999). In some cases, including Maryland, states may have even been ahead of federal policy (McDonnell, 1994). In response to arguments that assessments using traditional

outcome based standardized tests led to educational practices that overemphasize basic skills (Resnick & Resnick, 1992), Maryland began to develop performance-based standards and assessments (Stone & Lane, 2003).

In 1990, the Maryland State Board of Education established the Maryland School Performance Program (MSPP) to fulfill the recommendations made by the Governors' Commission on School Performance (MSDE, 2000). With MSPP, Maryland became one of the first states in the nation to develop standards in all subject areas and an assessment to measure attainment of those standards (Mintrop & Trujillo, 2005). Boards composed of state and local school system content supervisors developed specific subject area standards. These Learning Outcome Development Committees developed the Maryland learning outcomes in reading, mathematics, writing, language usage, science, and social studies (MSDE, 2000). The learning outcomes were used to develop MSPAP assessment objectives, test item specifications, and the test items themselves (Cerrillo, Hansen, Parke, Lane, & Scott, 2000).

In order to test whether students mastered curricular outcomes, Maryland developed the Maryland State Performance Assessment Program (MSPAP), "criterion-referenced performance tests in reading, mathematics, writing, language usage, science, and social studies for students in grades 3, 5, and 8" (MSDE, 2000, p. 8). MSPAP claimed to measure how well students "relate and use knowledge from different subject areas and how well they apply what they have learned to solve real world problems" (p. 8). It also claimed to emphasize "higher order skills such as supporting an answer with information; predicting an outcome and comparing results to the prediction; and comparing and contrasting information" (p. 8). The prevailing assumption underlying the

use of performance-based assessments is that they encourage the use of instructional strategies and techniques that foster reasoning, problem solving, and communication (National Council on Education Standards and Testing, 1992). All of these characteristics are integral to scientific inquiry-based instruction (AAAS, 1993; Hammer, 1997; Furtak & Ruiz-Primo, 2005; NRC, 1996). But do teachers enact these instructional practices in science classrooms?

The explicit goal of MSPAP was to change *teaching* (emphasis mine) in Maryland, which effectively made teachers a target of the testing policy. The MSPAP's mission was to measure the progress of schools and districts. These measurements were intended to induce higher standards and "fundamental changes in instruction" (Koretz, Mitchell, Barron, & Keith, 1996, p.vii). MSDE suggests that MSPAP sets the learning standards, and, thus teachers should teach to the standards by teaching to the test.

MSPAP guides instruction by measuring students' understanding of rigorous content and their ability to apply what they learn to real-world problems. Teachers improve students' performance on MSPAP by teaching students to analyze what they read, apply skills and knowledge to solve problems, integrate knowledge from different content areas, and work independently and in groups. In this sense, "teaching to the test" is a good instruction, the kind of instruction that results in understanding and not mere rote recall of isolated facts. (MSDE, 2002, ¶ 4)

In effect, the test became the curricular guide rather than the assessment. Teachers began using MSPAP-like questions and focusing on MSPAP-type reasoning in class (Stone & Lane, 2003). Rather than teaching standards, teachers began teaching the test.

MSPAP included teachers in the scoring process. Goldberg and Roswell (2000) report that many of the 650 to 700 teachers who scored the MSPAP tests each summer identified the project as an opportunity to obtain professional development not otherwise available through system- or state-based activities. Maryland's Local Education Agencies

(LEAs) regarded the experience so highly that they attached financial and other incentives for teacher participation and lobbied vigorously for the opportunity to host one of the four regional sites at which summer scoring occurred each year. What value did teachers and LEAs see in these summer programs? Rather than buying-in to the pedagogical value of high-stakes testing, this emphasis on scoring may indicate that teachers and LEAs wanted the opportunity to get an inside look at the MSPAP tests, so they could prepare their students for the kinds of questions that would be asked.

In 2002, Maryland began to revamp its accountability structure according to *No Child Left Behind*. While the law was more prescriptive than previous legislation, individual states allegedly were given broad freedoms to develop specific definitions for proficiency and ways to assess it (United States Congress, 2002). In 2003, MSDE implemented the Maryland School Assessment (MSA), an outcomes based assessment, to replace the MSPAP. The stated reason for the switch was compliance with *No Child Left Behind*'s requirements. As opposed to the MSPAP, the MSDE (2004a) claims that the MSA is aligned with the state content standards (MSDE, 2004a). Stated reasons, however, do not always match actual policy purposes in implementation (Elmore & Sykes, 1992; Kingdon, 2003; Ripley, 1985). Perhaps MSDE could not handle the complexities of performance based assessments, or they finally responded to state wide resistance to the MSPAP from schools and parents (Gowen, 2002; Shulte, 2002). In any case, MSDE decided to simplify its testing program with the MSA. The MSA reduced curricular coverage to math and language arts (and eventually science) and expanded the number of students tested to include all grades 3 through 8 (MSDE, 2004a).

In addition to the MSA, Maryland developed the High School Assessment (HSA) to comply with the high school testing requirements of *No Child Left Behind*. Like the MSA, the HSA is allegedly aligned with the learning goals for each content area that it tests. Across the United States, alignment of standards and assessments, however, has proven to be more challenging than state agencies claim. Although numerous strategies for aligning state tests to standards have been developed (AAAS, 1993; Leffler, 2004; Mid South Regional Resource Center, 2004), high-stakes tests have been found to be poorly aligned with state standards (Resnick, Rothman, Slattery, & Vraneck, 2004). Further, Kulm, Wilson and Kitchen (2005) reveal that student thinking in answering test questions often does not correspond to the intended standard. In other words, although a state may claim that a test question aligns with a particular standard, students may not actually apply the standard to answer the question. Finally, as with most components of *No Child Left Behind*, the notion of alignment subscribes to a technical curricular perspective. By predicting a specific structure for standards and assessments, alignment ignores and diminishes any emerging components of the curriculum. Is it reasonable or desirable to expect that everything in the classroom will be rigidly aligned with a test or a standard? What do teachers think about the notion of curricular alignment? Is their pedagogy compromised by such imposing structures? Educational experiences are filled with both conscious and unconscious desires about what the curriculum addresses (Ellsworth, 1997). By fixating on alignments of standards, tests, and pedagogy, education strains to hone its structures to achieve particular outcomes. Instead, Ellsworth argues that education should be open and free to possibilities where teachers and students can explore “previously unimagined ways of thinking and knowing” (p. 58).

The format of the HSA is similar to the MSA, except the HSA covers four core subjects: English, government, algebra/data analysis, and biology. The English and algebra sections currently fulfill the requirements of *No Child Left Behind* that all students are tested in mathematics and language arts at least once in high school. Starting with the 2007-2008 academic year, the biology section will fulfill the science testing requirement. Beginning with the graduating class of 2009, students will be required to earn a satisfactory score on the HSA in order to earn a Maryland High School Diploma (MSDE, 2005). Thus, the HSA expanded accountability into the high school and will add a severe penalty (beyond a national mandate) for inadequate student performance. In effect, the accountability noose is set to tighten further in 2009. How will teachers respond? How will they deal with such consequences?

In order to facilitate alignment between learning goals and assessments, MSDE developed the ironically named Voluntary State Curricula for science, language arts, social studies, and mathematics, which MSDE would like implemented by all the schools in the state. Each curriculum defines what students should know and be able to do at each grade level. Each document includes content standards or broad, indicator statements which explicate the corresponding content standards and objectives to “provide teachers with very clear information about what specific learning should occur” (MSDE, 2004b, ¶ 1). The curriculum supposedly is aligned with the MSA, and teachers are encouraged to comply with the Voluntary State Curricula to avoid the consequences of student failure on the high-stakes test. The MSDE currently is developing Voluntary State Curricula for the HSA. It already has developed drafts for nine subject areas, but not developed drafts in science.

One of the ramifications of the litany of curricular requirements is that teachers may feel overwhelmed by the new curriculum, confused by it, and/or ignore it (Finley, 2000). On the other hand, several researchers found that with both the MSPAP and the MSA, most teachers respond by teaching to the tests (Cerrillo, Hansen, Parke, Lane, & Scott, 2000; Parke, Cerrillo, Levenson, O'Mara, Hansen, & Lane, 1999; Stone & Lane, 2003).

The Maryland State Department of Education embraced the national calls for accountability, and it did it well according to the paradigmatic framework of *No Child Left Behind*. Maryland is considered to be a leader in the implementation of accountability measures (Mintrop & Trujillo, 2005). According to the Maryland State Department of Education (2006), after the 2004-2005 school year, Maryland earned an "A" for a fifth straight year from *Education Week* on "its K-12 standards and its system of student testing and school accountability" (p. 57). Further, Maryland placed second on a ranking of AP test scores, and the state reports significant improvement on MSA scores. The Maryland public school system has embraced the standards movement. Because of its successful position in the arena of standards and accountability, it is likely to activate many actors and mechanisms to ensure continued success as measured by the accountability system. Therefore, Maryland likely provides a revealing model for an in-depth examination of science instruction within the accountability system because it is likely to be what Patton (1990) calls an information-rich case.

Accountability in Buckley County¹

In 1999, Buckley County Public Schools (BCPS) responded to Maryland's educational accountability policy with a multiyear initiative designed to "raise the level of student achievement to rigorous standards of academic performance."² In 2004, the initiative was extended through 2009. BCPS is in line with the federal and state accountability programs. The Board of Education's academic priorities include two goals that directly address accountability:

- 1) Align rigorous curriculum, delivery of instruction, and assessment for continuous improvement of student achievement.
- 2) Use student, staff, school, and system performance data to improve student achievement.³

To take it a step further, BCPS developed a strategic plan with four goals, the first two of which directly address the accountability policy. The first goal for the district is to ensure success for every student. According to BCPS, the first milestone to fulfillment of the goal is that all students in the district will achieve or exceed state proficiency standards in mathematics, reading, and writing as measured by state wide assessments. The second goal is to provide an effective instructional program. The first benchmark that measures this goal is that by the end of 2nd grade all students must be prepared to meet state standards in math and science. Federal and state mandated testing in math and language arts begins in third grade. Clearly BCPS has fallen in line with the accountability policy and the consequent influence on instruction. Both in its rhetoric and policies, BCPS reflects the key high-stakes accountability principles of *No Child Left Behind*.

¹ In order to protect the anonymity of this study's participants, I use a pseudonym for the school district in which this study takes place.

² This citation is from a document produced by the district, but I have excluded it from the references in order to protect the district's identity.

³ This citation is from a document produced by the district, but I have excluded it from the references in order to protect the district's identity.

According to a 2005 annual report BCPS made “significant progress” in all four of the school systems’ strategic goal areas. In other words, the district progressed in its technical implementation of accountability measures. BCPS is one of the state leaders on standardized test passing rates. Overall, in 2004, the district saw every one of its subgroups meet Maryland’s Annual Measurable Objectives (AMO’s). Out of 190 schools, 170 (90%) met the 2004 AMO’s in every category, and 188 schools (99%) made Adequate Yearly Progress (AYP) for each of the racial and ethnic groups (Stevenson, 2005, p. 4).

If Buckley County wants to maintain its reputation for high passing rates on standardized tests after science testing is included in the battery of assessments for the 2007-2008 school year, it must develop a science curriculum that is closely aligned with state standards; it should ensure that its own standards and curricula match the state tests; and it should prepare teachers to implement the new curriculum in classrooms across the district. This process is likely to reduce the professional space for teachers even further. In the midst of such proliferation of technical frameworks, teaching with any “sense of coherence or continuity or community begins to appear to be virtually impossible” (Friesen, Clifford, & Jardine, 2002, p. 116).

In order to facilitate the implementation of a successful science testing program, Buckley County Public Schools (BCPS) spent the school year of 2005-2006 revising its science curriculum. The process fully reflects a technical curricular perspective in which curriculum is “developed” outside the classroom. Initially, with the help of high school science teachers from across the district, BCPS developed a curricular framework for each science discipline – biology, chemistry, physics, earth/space systems, and matter

and energy – taught in high school. Although this process involved teachers, they may not feel ownership of the framework. Teachers were asked to play by the rules of accountability, but the “intensification” of their work and the top-down power structure in the education system may have left many of them de-skilled (Apple, 1992). Therefore, when surrounded by those who are considered experts and asked to subscribe to particular structural frameworks, teachers may feel unable to offer their own authentic imprint on the curriculum.

BCPS has implemented a three step process of curriculum development. In each subject area, the county developed frameworks which contain the goal, enduring understandings of the subject, content description, instructional approach, and a collection of the subject indicators that students should know upon completion of the course. The frameworks have been used to develop curricular blueprints in each subject area. Science blueprints describe the major elements of subject area units and are intended to serve as a basis for the development of lesson plans. The blueprints allegedly match enduring understandings, essential questions, indicators, and assessment strategies by topic. Currently, the county is revising comprehensive draft curriculum modules to match the frameworks and blueprints. The new science curricula are being implemented during the 2006-2007 academic year. According to BCPS, they “align with and exceed the Maryland School Performance Program Core Learning Goals”.¹ Clearly, over the last few years the district has been making a concerted effort to prepare for the implementation of the high-stakes science test.

¹ This citation is from a document produced by the district, but I have excluded it from the references in order to protect the district’s identity.

The outline of BCPS's response to *No Child Left Behind* shows that the county is moving down the technical path of curriculum frameworks. Its buy-in to accountability structures makes BCPS an informative district for this study because teachers are likely to be riddled with tensions due to conflicting curriculum practices and beliefs in relation to the demands of accountability. It fits Flyvbjerg's (2001) definition of a critical case. Because Buckley County is responding vigorously to state and national science accountability mandates, a study within the county is likely to uncover more nuances about the implementation of science accountability than a more "typical" district that has less to lose from non-compliance. Further, as a relatively affluent district, Buckley County Public Schools is as likely as any school district to have the proper resources to reach the benchmarks established by *No Child Left Behind*.

Teaching Accountability

If, as I propose, teaching is more than instructing, more than training, more than method and more, much more than instrumentation, then what is it? The "it" teaching is is the third thing, neither this nor that. It is to encourage being "like" letters in running water. (Doll, 2000, p. 146)

According to curriculum theorists, three paradigms exist for examining curriculum: technical, interpretive, and critical. The technical perspective, most notably attributed to Ralph Tyler (1949), constitutes what Pinar, Reynolds, Slattery, and Taubman (2000) describe as the heritage of the field, when curriculum studies was dominated by implementing bureaucratic processes. According to the technical perspective, curriculum entails the "concrete ever changing tasks of curriculum development, design, implementation, and evaluation... [which are] largely theoretical; being directed at school people who want to know 'how to;' it [is]...practical" (Pinar, as cited in Pinar, Reynolds, Slattery, & Taubman, 2000, p. 212). The knowledge in this

perspective is observably gained and measured, and the relationship between actors in the educational system is based on causality and functionality (Aoki, 2005). Teachers are essentially excluded and de-skilled by this perspective of curriculum because they are simply media for the larger, more powerful educational actors (Apple, 1992). Although this perspective is the heritage of the curricular field, the current U.S. educational accountability system reflects this technical curricular paradigm.

The interpretive paradigm moves beyond curricular development as implementation in the technical sense. Rather, it is characterized by an effort to understand the curriculum (Pinar, Reynolds, Slattery, & Taubman, 2000). Rather than being rooted in the social sciences alone, this perspective draws upon the human sciences, and, as such, it seeks to understand curricular meanings from the perspective of the people involved in the curriculum process (Ellsworth, 1997). Knowledge is experientially meaningful, contextually authentic, and intersubjective (Aoki, 2005). I use this perspective to examine science educational accountability by seeking to understand how teachers make meaning of their experiences with it.

The critical perspective of curriculum builds upon interpretation. It seeks to contextualize the curriculum socially, politically, and economically and examine it through the lens of social justice (Pinar, Reynolds, Slattery, & Taubman, 2000). The critical perspective is rooted in postmodernism which refutes grand narratives as explanations for reality, rejects universal reason as a foundation for human affairs, and problematizes representation and context (Giroux, 2002). The critical perspective's interest is in action to improve the "human condition by rendering transparent tacit assumptions and hidden assumptions and by initiating a process of transformation

designed to liberate man” (Aoki, 2005, p. 100). As such, it is also called the emancipatory curricular perspective, because by offering a critique of society, it encourages emancipatory opportunities for participants engaged in the curriculum process (Pinar, Reynolds, Slattery, & Taubman, 2000). I use this perspective to challenge some of the potentially dis-empowering structures of high-stakes accountability.

In terms of education policy, teaching and learning are understood and valued differently by different education stakeholders. Stakeholders’ perspectives on teaching often reflect their positions in the educational system and their organizational affiliations (Haughey, 1993). Those who develop accountability plans are likely to have a very different interpretation and understanding of instruction than teachers who implement them in the classroom. For example, test developers have been found to have little faith in teachers’ abilities to interpret results and to adjust their practice in response to testing outcomes (Rudman, 1989). This notion comports with the business metaphor in which curriculum producers offer something to the curriculum consumers. Thus, skills of a teacher-as-implementer are reduced to techniques oriented toward efficiency and instrumentalism (Aoki, 2005).

Using Tyler’s (1949) technical rationale, states and districts attempt to simplify instruction into a set of inputs and outputs by defining it according to standards, curricular guidelines, learning outcomes, and assessments that influence teachers’ pedagogical perspectives (Louis, Febey, & Shroeder, 2005). Teachers, on the other hand, may hold a more complex and problematic perspective of accountability and instruction because of their regular interactions with students, parents, and the greater school community (Invernizzi, Landrum, Howell, & Warley, 2005). Further, students and their

families may have different educational objectives than those laid out by accountability policies. Perpetuated by societal structures and forces that espouse the importance of technical skills (Apple, 1996), the primary purpose of schooling for many students is to ensure the development of essential academic and social skills necessary to advance in a career after graduation and to function properly in our society (Rosenbaum, 2001). Thus, teachers are faced with the difficult task of adjusting their instruction in a way that meets the expectations of the state, while addressing the intellectual and practical learning needs of their students. The negotiation between these two goals may be complex and fraught with contradictions. Students represent a daily and urgent need, while the teachers' professional goals may be largely dependent on the state's long-term mandates.

Good teaching practice should be evaluated by broader indicators than student test scores because teaching is a non-linear, complex endeavor. Even if the standardized assessment is deemed to be an appropriate single measure of student learning (an assumption that is held by the accountability policy), quality teaching goes beyond student performance. A high-stakes test only measures outcomes, willfully ignoring teaching itself (Aoki, 2005). Medley (1982) makes a distinction between "teacher competency" and "teacher effectiveness." Because the teaching process is complex, it is impossible to correlate particular teaching practices with specific student outcomes, especially on one specific measure like a standardized test. Further, "teacher competency" and "effectiveness" still are rooted in linear connections between students and teachers. In other words, student behaviors and methods of learning guide decisions about how to teach (Aoki, 2005). The relationship between students and teachers, however, is more complicated and non-linear. Its essence cannot simply be characterized

by teachers knowing what students need to know because they possess more knowledge, authority, or experience (Ellsworth, 1997). The difference is more holistic because students and teachers occupy a different “*location* within the pedagogical structure of address” (p. 62, emphasis in the original). The complicated connections between these locations are where and how learning takes place. Therefore, assessing teaching based on a limited measurement tool that ignores teaching by examining it solely through an outcomes model is likely to send messages to teachers that conflict with their pre-understandings, conscious and subconscious, of the meaning of teaching.

Because quality teaching means different things to different people, it is extremely difficult to standardize. Although teachers may find creative outlets, a system of accountability where the aim is to achieve proficiency for every student demands some standardization of teaching practice implicit in the policy (Brees, 2003). The logic follows along the lines of...If all students must achieve x, then all teachers must do y (or a combination of y, z, and q). Ironically, the actual components of teaching are not addressed by accountability plans. Productive discussions about the details of quality teaching practice are rarely held in official arenas because it is difficult to agree on strategies for teacher improvement (Scanell & Metcalf, 2000). Thus, the policies outline particular outcomes for students without explicating what teachers must do to ensure that students meet the standards. Nevertheless, accountability plans have been shown to influence teaching by inducing teachers to teach to the test (Cerrillo, Hansen, Parke, Lane, & Scott, 2000; Parke, Cerrillo, Levenson, O’Mara, Hansen, & Lane, 1999; Stone & Lane, 2003). Such a construction of teaching ignores the unpredictability of a lived-curriculum. Therefore, tensions are likely to arise between the mandates of high-stakes

testing and teachers' understandings and best practices. According to the rhetoric of accountability, students cannot succeed without quality teaching, and teachers are the street-level bureaucrats (Lipsky, 1980) who are responsible for implementing the accountability system. Therefore, an interpretive examination of teachers' tensions within the science accountability system may challenge the foundations, assumptions, and structures of accountability policies.

Overview of the Study

Tensions appear to describe the life of an academic. (Selden, 1991, p. 26)

Without offering a public explanation of all the science disciplines covered by the Maryland's science content standards (earth/space science, chemistry, physics, and biology), the Maryland State Department of Education included only biology on the HSA. Because biology is the sole branch of science included in the state's science accountability system, high school biology classes should be most affected by the new high-stakes test and the curricular restructuring accompanying it. In other words, high school biology teachers should experience the consequences of Maryland's science accountability system more directly than other high school science instructors. Other science subjects may have content standards that students are supposed to acquire, but the state does not formally assess students on whether they acquired that knowledge. Therefore, non-biology high school science teachers have fewer consequences or "stakes" associated with their students' learning.

Purpose of the Study

The aim of this study is to explore how biology teachers understand and construct their practice in a high-stakes accountability environment that is likely to be riddled with

tensions. Teachers find themselves along the forefront of educational accountability. They are the actors who interact with their students on behalf of the entire educational system. Thus, from the perspective of both students and the school system, teachers are directly responsible and accountable for transmitting the state biology standards to their students. In the words of the *No Child Left Behind* Act, they are responsible for ensuring that all of their students “achieve academic proficiency” (United States Congress, 2002a). This complicated social position within the system of education is the focus of this study, which seeks to uncover and interpret how science teachers negotiate the complicated and possibly competing pressures associated with educational accountability.

An examination of the effects of high-stakes science testing on classroom practice presents a major gap in current educational research. Because the federally mandated system of accountability in this country is new, most states have only recently begun implementing science accountability measures (Haury, 2001). Thus, the research literature on this topic is limited. The research conducted on accountability has mainly focused on math and language arts, the content areas for which the *No Child Left Behind* policy has mandated testing since the 2002-2003 academic year, so I draw on this literature to situate my study. Because *No Child Left Behind* ordered science testing beginning in the 2007-2008 academic years, the school district only recently has begun to prepare for the impending high-stakes tests. If they must work within the frameworks of accountability, educators deserve to have the tools to function within the system and to understand its structural and technical framework. An understanding of the accountability system may allow them to negotiate through it without suffering some of its more severe consequences, while being able to address the practical needs of students and other

educational stakeholders. This research seeks to provide educators with insights into how educational accountability influences the science classroom.

As the policy's ultimate implementers, teachers are integral to the success or failure of *No Child Left Behind*. Related research has shown that the current high-stakes standardized testing model may not be well aligned with the science pedagogy that is advocated by national, state, and local school districts (Stecher, 2002). The system is prescriptive and outcomes based, while all three levels of the education system emphasize the importance of scientific inquiry-based instruction, which is deeply focused on the processes of doing science (Hammer, 1997). Thus, teachers may find themselves negotiating their way through competing and possibly irreconcilable institutional signals that deal with exactly how science should be taught and how the discipline is constructed, primarily as a set of processes or as a content base.

This study examines the effects of the accountability system through both an internal and a structural lens. Because the standardized tests have high-stakes attached to them, teachers are likely to succumb to the overt pressures presented by the assessments. In order to protect themselves from the harsh ramifications associated with failure to meet the demands of the system, at a bare minimum, teachers are likely to develop coping mechanisms to negotiate their way through a maze of conflicting structures. Therefore, this study examines how teachers negotiate through the system to maximize the quality of their own practice and their students' learning within the frameworks of accountability. By questioning the technical paradigms of accountability this study challenges the foundations of accountability. Such a critical approach may help teachers develop empowerment strategies that can free them from the de-skilling effects of the

educational accountability system (Apple, 1996). A study that examines how teachers confront this challenge can provide invaluable insights into science accountability programs that continue to be spreading throughout the United States.

Research Questions

The overarching research question that is the focus of this study asks: How do high school biology teachers negotiate the explicit and implicit messages regarding high school biology accountability policies governing their work? This question is examined and elucidated through three auxiliary questions and sub-questions:

1. What, if any, tensions are embedded within and across the national science standards, Maryland State science standards and the high-stakes assessment, and the district's biology curriculum?
2. What, if any, tensions do biology teachers perceive in these messages or between these messages and their construction of teaching?
3. How do biology teachers' efforts in interpreting the policy messages, both implicit and explicit, shape their teaching practice?
 - a. How, if at all, do biology teachers adjust their teaching in response to the accountability system?
 - b. How, if at all, do biology teachers adjust their understanding of the discipline of science to conform to the messages they receive from the accountability system?
 - c. How, if at all, do biology teachers construct and understand their role in the classroom in response to the accountability system?
 - d. How, if at all, does the accountability system influence biology teachers' passion for their profession?

These questions are intended to inform the complicated relationship between education policies and teaching practice. Through an interpretive analysis of these questions, using case study methodology I seek to reveal the tensions, conundrums, paradoxes, and compatibilities that exist between how teachers understand and construct their practice and externally imposed policy regulations on their performance.

Overview of Methodology

In order to develop a specific strategy for investigating this phenomenon, it is important to develop a proper epistemological understanding of the methodology to be employed in order to contextualize the knowledge gained from the study. All research methodologies have theoretical perspectives that guide the processes and procedures associated with them (Crotty, 1998). In other words, in addition to having its own set of techniques and strategies for conducting research, each methodology bears its own set of assumptions about the nature of knowledge, its presuppositions and foundations. When conducting investigations, researchers hold specific assumptions about whether and how knowledge and understanding are socially embedded. A particular methodology allows the researcher to uncover and examine knowledge through a particular paradigmatic lens. Thus, examining the same issue with multiple methodologies bearing different theoretical perspectives may render very dissimilar results.

Case study research is one of the research methodologies that can be used for an interpretive study of teachers' experiences. As a set of research methods, a case study is broadly defined and used in a multitude of social and educational settings. One of the greatest strengths of a case study design is its call for an "intensive description and analysis" of a specific case to gain an "in-depth understanding of the situation and [its] meaning" (Merriam, 1998, p. 19). Case studies allow for an examination of a phenomenon as it is manifested in an actual context. Yin (2003) suggests that the "case study method allows investigators to retain the holistic and meaningful characteristics of real-life events" (p. 2). In other words, case study methods allow researchers to examine the case *in vivo*, like the behavior of a living organism in its natural habitat. Thus, case

study methodology affords me the opportunity to study the effects of science accountability measures on teachers within a particular school, district, and state environment.

The power and significance of an opportunity to study a phenomenon embedded within a real-life case cannot be taken for granted. A researcher's epistemological perspective is integral to understanding the particular habitat in which a phenomenon resides and how the phenomenon can be uncovered within a particular social context. According to the constructivist paradigm, which holds that knowledge and meaning are "constructed rather than discovered" (Stake, 1995, p. 99), an investigator must seek "culturally derived and historically situated interpretations of the social life-world" (Crotty, 1998, p. 67). As researcher who holds a constructivist epistemological perspective, I hold the perspective that all knowledge is socially embedded. Therefore, an interpretive approach to research allows me to gain a greater understanding of the phenomenon itself *and* its meaning to those who experience it. In other words, the phenomenon cannot be fully understood in and of itself. Its meaning must be filtered through the lens of individuals who experience it. It is the interaction between the research and the participants that leads to the generation of data (Guba & Lincoln, 1989).

In this study, I use interpretive case study methodology to uncover science instruction within the confines of educational accountability. An interpretive case study allows me to study the process of how teachers negotiate their way through a complicated multi-level system of curriculum and accountability directives within the social context of the school. Merriam (1998) says that in interpretive research, "Education is considered to be a process and the school is the lived experience" (p. 4) because interpretive work is

concerned with examining contextual and situational knowledge (Aoki, 2005). Thus, an interpretive case study can help uncover teaching within the greater context in which it takes place. Further, interpretive case study methodology can help illuminate the process of education and its meaning from the teacher's perspective (Merriam, 1998). The understandings gained from the research can be used to develop connections and themes about science teachers' experiences with accountability for further consideration by actors involved in the accountability system.

In this study, I use three sources of data – document review, semi-structured interviews, and classroom observations as outlined by Merriam (1998), and I use Merriam's (1998) and Stake's (1995) protocols for their implementation. A careful document review is used to answer the first auxiliary question regarding the messages embedded in national, state, and district policies. I hold semi-structured interviews with biology teachers to gain insight into the second and third auxiliary questions. I seek to gain an understanding of the tensions and conflicting messages that teachers perceive in the science accountability system. Further, I attempt to gain an understanding of how teachers may be adjusting to the science accountability system. Specifically, I focus on its effect on teaching, teachers' understanding of the discipline of science, teachers' understanding of their role in the classroom, and teachers' passion for the profession. I use classroom observations to gain further insights into the third auxiliary question. In general, my observations in this study focus on insights into teaching practices that are a response to the HSA, behaviors that are indicative of how teachers perceive their role in the classroom and within the accountability system, and indicators that characterize or run counter to scientific inquiry-based instruction.

Theoretical Foundations

The theoretical foundations for this study are developed from three broad and interconnected areas of study. This study theoretically occurs at the intersection of educational accountability and pedagogy. In general terms, according to the critical or emancipatory curricular perspective, pedagogy is a way for teachers to be with their students that leads them to a new place of knowing. Together they take a transformative pedagogical journey, which brings them toward some point of deeper understanding or enlightenment. Bell hooks (1994) writes, “Combining the analytical and the experiential is a richer way of knowing” (p. 89). According to hooks, an active unity of being and understanding leads to a powerful social transformation. Thus, teaching is a personal and social commitment between individuals that is defined and executed by them.

Educational accountability powerfully interacts with pedagogy because it attempts to influence teachers’ aims. Rex and Nelson (2004) describe two forms of educational accountability: externally mandated accountability and accountability that comes with responsibility toward one’s own students. In terms of the latter, good teachers hold themselves accountable for student learning. They bear the responsibility for the journey that they take with their students and for the outcome of that journey. A responsible teacher understands the power of pedagogical interactions, and, as a result, thoughtfully engages with students in making meaning. In this sense, teaching is like using nuclear energy responsibly. Good teachers hold themselves accountable for using pedagogy for constructive, rather than destructive, purposes. Increasing external accountability measures may be counterproductive when they compete with teachers’ sense of internal accountability (Rex & Nelson, 2004).

On the other hand, teachers are held accountable for student performance by reform oriented policies designed to improve student achievement, as defined by state standards and measures. Because this mandated notion of accountability applies to thousands of teachers, its definition and understanding of pedagogy is narrower and more general than the one held by individual teachers regarding their own students (Louis, Febey, & Shroeder, 2005). In the sense of external accountability, teaching is public, in addition to being personal. Since teachers make meaning of their relationships with students individually, understanding their own responsibilities is a complex and multilayered process (Invernizzi, Landrum, Howell, & Warley, 2005). Externally mandated accountability, on the other hand, may reduce teaching to a few specific indicators of “Adequate Yearly Progress” (in the case of Maryland, test scores and attendance).

These two types of accountability offer potentially opposing perspectives on teaching and the responsibility of teachers. Hansen (2001) writes:

...the practice of teaching as it has come down to us through time and human effort, does not constitute a hardened, unchangeable endeavor to which teachers must bend themselves unquestionably. Rather, it is a living practice. It evolves as a result of the initiative and imagination of teachers, part of whose task is to respond (but not to “react”) to external pressures and social demands. (p. 9)

Hansen describes an interplay between externally mandated accountability emanating from an educational policy and teachers’ sense of professional responsibility for their practice. Although teachers have their own notions and understandings of teaching, their conceptualization of pedagogy does not exist in a vacuum. Teachers’ pedagogical or content knowledge, beliefs, and experiences (Olsen & Kirtman, 2002), as well as the school’s social and professional environment (Clandinin, 1986) all influence the

curriculum. Further, high-stakes accountability policies impose rules and pressures to which teachers must adhere in order to avoid penalties. Serious tensions between external guidelines and teachers' internal conceptualizations of teaching may arise. Nevertheless, these pedagogical tensions are somehow addressed by teachers before the "final product" is manifested in the classroom. This study examines how teachers live these tensions that result from the signals put out by science educational accountability. In light of these complexities, I attempt to unpack the process that teachers undergo as they navigate their teaching through a high-stakes environment that likely places many controls on their pedagogy.

Summary

I intend for this study to contribute to the limited existing body of research on the effects of accountability on teaching practice in science. High-stakes assessments have been shown to place serious challenges on teachers, such as influencing their pedagogy and attitudes toward teaching. This study helps uncover the experiences of science teachers as districts and schools attempt to comply with the mandates of *No Child Left Behind*. Because teachers are directly responsible for student learning, they are largely responsible for implementing the policy's mandates. In other words, by its own technical logic, and according to both the interpretive and critical curricular perspectives, the success or failure of high-stakes accountability ultimately rests in the hands of the teachers. Thus, an understanding of how teachers negotiate through the policy's implicit and explicit signals is extremely relevant in today's educational climate. Further, if our schools are to achieve their ultimate goal of providing the best education possible, educational stakeholders must be aware of how the overarching accountability climate is

weighing in on teaching. Ultimately, the aims of accountability are irrelevant if we do not know how they are manifested in the classroom. We can only describe the benefits and drawbacks of the policy and challenge its technical foundations when we have gained a true understanding of its effect on teaching.

In Chapter II, the literature review, I focus on three relevant strands for this research. First, I delineate how quality science teaching based on scientific inquiry is outlined in policy documents in order to lay a foundation for understanding what is cast as quality science instruction in science standards documents. Second, highlighting the perspectives of curricular theorists, I focus on accountability and the current research on its effects on classroom practice. Finally, I outline the current body of research on how teachers balance competing tensions in the classroom. Specifically, I focus on how the research may relate to high-stakes accountability mandates. These threads of knowledge provide the theoretical foundation for this case study and its findings. In Chapter III, I provide the methodological grounding and description of my study, and in Chapter IV, I outline the context in which this study takes place. In Chapters V-VII, I develop the themes that outline tensions biology teachers experience in the current accountability climate. In Chapter VIII, I conclude with an exploration the implications of this study for science teachers and the accountability system itself.

CHAPTER II: REVIEW OF THE LITERATURE

Overview

The purpose of this literature review is to provide theoretical foundations for the research questions to be investigated in order to demonstrate how this study will advance what we already know about the influence of educational accountability on science teaching and curriculum. For a qualitative study, the literature review may be framed in a manner consistent with the methodological assumptions (Creswell, 1994). Thus, the literature is used to inform the research questions and the framework of the study, but it does not necessarily direct the findings as it would in the quantitative paradigm. Instead, the research is guided by the data, the documents and participants involved in the study. The literature is used to provide a context for the data. It does not necessarily provide a model for collecting or analyzing the data (Stake, 1995).

This chapter situates the study through a synthesis of background research in the following three areas: how quality science teaching based on scientific inquiry is outlined in policy documents in order to lay a foundation for understanding what is cast as quality science instruction in science standards documents, accountability and its effects on the curriculum, and a grounding in research on how teachers process competing messages in the classroom, especially in response to high-stakes accountability policies. The review of research and writing on science instruction shows that scientific inquiry based instruction generally has been accepted as the preferred method of teaching science by many educational researchers, and it is promulgated by science education standards and policy documents. The review of research and writing on accountability shows that high-stakes testing policies have mixed results in the classroom, and they promote a technical

paradigm of teaching and learning. Finally, the review of research and writing on how teachers process competing tensions in the classroom reveals that teachers synthesize contradictions and integrate them into their pedagogy, but not without complications and glitches throughout the process that can influence the curriculum in significant ways.

Scientific Inquiry-Based Instruction

Both the Benchmarks for Scientific Literacy (AAAS, 1993) and the National Science Education Standards (NRC, 1996) make lofty pronouncements which call for the attainment of scientific literacy for everyone. The documents underscore that society is changing at an increasingly rapid pace. Science, technology, and mathematics will play an integral role in the change – causing, shaping, and responding to it (NRC, 1996). As a result, all people must be prepared to negotiate their way through the social and technological transitions slated to occur throughout their lifetimes. According to the National Research Council (1996), “Everyone needs to use scientific information to make choices that arise every day. Everyone needs to be able to engage intelligently in public discourse and debate about important issues that involve science and technology. And everyone deserves to share in the excitement and personal fulfillment that can come from understanding and learning about the natural world” (p. 1).

Quality science education plays a central role in facilitating the development of a scientifically literate society. In order to fulfill the demands of the National Science Standards, all students should be able to demonstrate high levels of scientific performance. The national standards call science curricula to focus on encouraging the attainment of scientific knowledge and processes like reasoning, creative thinking, decision making, and problem solving (AAAS, 1993). What does such a curriculum look

like? How should science curricula and teachers address the call for increasing scientific literacy? How can secondary school science teachers most effectively facilitate the development of the skills and processes called for by the scientific and science education communities?

Both the *Benchmarks for Science Literacy* and the *National Science Education Standards* were developed by educators and scientists from a synthesis of research on science curricular materials, learning, and best classroom practices. These documents remain the most extensive and comprehensive compilation of science teaching practices to date and are routinely cited in research articles on the topic (Leonard & Penick, 2005). Both documents recommend scientific inquiry-based instruction as the most effective way to teach all students scientific content and processes. The documents argue that scientific inquiry promotes the learning, understanding, and retention of scientific concepts, and it cultivates a mastery of the process of doing science. The standards suggest that inquiry-based teaching must promote the formulation of scientific questions, crafting hypotheses, understanding different variables and experimental designs, collection and analysis of data, and making valid inferences from the data. Further, the documents call students to understand the nature of scientific knowledge. They must realize that it is empirical, testable, verifiable, refutable, and publicly open for discourse (AAAS, 1993; NRC, 1996).

What is Scientific Inquiry?

Others see the classroom as an arena, not only for student exploration but also for teacher exploration, of the students understanding and reasoning, of the subject matter, of what constitutes progress toward expertise and how to facilitate that progress. For them, successful instruction depends on teachers' often unanticipated perceptions and insights. One might call this discovery teaching. (Hammer, 1997, p. 485)

One of the major challenges to fostering quality scientific inquiry-based instruction is the diversity of ideas around what constitutes scientific inquiry (Harwood, Reiff, & Phillipson, 2002). Thus, it is important to develop a clear understanding of scientific inquiry before discussing its values in the classroom. The national science standard documents portray scientific inquiry as a teaching approach, science process skills, and science content (Lederman & Niess, 2000). The National Research Council (1996) refers to scientific inquiry as "...the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work. Inquiry also refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world" (p. 23). If the goal of science is to discover theories and laws that explain natural phenomena, then the skills, logic, and dispositions that lead to those discoveries can be considered aspects of quality scientific inquiry (Zachos, Hick, Doane, & Sargent, 2000). Therefore, all characteristics of scientific inquiry-based instruction must contribute to the development of skills that are useful in making scientific discoveries about natural phenomena.

Two distinct features of the discipline contribute to the development of scientific knowledge. On the one hand, scientists must have a solid grasp of the concepts, theories, laws, and principles that have been developed through former investigations. In other words, scientific investigation must be based in prior scientific knowledge in order to cross the boundary legitimately from currently known understandings to yet unknown potential discoveries (Harwood, Reiff, & Phillipson, 2002). On the other hand, scientists must comprehend the processes of scientific investigations. They must understand the

actual methodology that is used to make the discoveries that build on the current knowledge base (Zachos, Hick, Doane, & Sargent, 2000).

Traditional science instruction begins with lecture, is followed by reading and worksheets, and then a lab to verify the lesson. Traditional pedagogy calls students to think deductively from cause to effect. This pedagogy contrasts with how most scientists inductively reason through their work. Thus, while this type of instruction may promote the learning of scientific knowledge, it has been shown to be ineffective for teaching the scientific process (Branford & Donovan, 2005; Haury, 1993; Kinney, 1989; Light, 1990; Lin, 1998; Lord, 2001; Schmidt, Gillen, Zollo, & Stone, 2002). Inquiry-based instruction reflects the constructivist paradigm of learning, which holds to the epistemological assumption that students make meaning from their experiences and previous knowledge (Crotty, 1998; Guba & Lincoln, 1982; Stake, 1994), and their understanding can be promoted by specifically designed instruction (Anderson & Helms, 2001; Duit & Treagust, 1998). Within this epistemological paradigm, students construct their own understandings, using their experiences and existing knowledge, and thereby they view the world in a way that is both logical and useful to them (Sinatra & Pintrich, 2003).

The constructivist understanding of knowledge forms the epistemological foundation of the scientific process. The explicit intent of scientific inquiry-based instruction is to foster the learning of scientific process skills that are grounded in students' scientific knowledge base (Zion, Michalsky, & Maverich, 2005). Scientific skills are at least as important as pre-discovered knowledge. Thus, inquiry-based instruction runs a course opposite to traditional scientific instruction. It is lived instruction that cannot be accommodated merely by using Tyler's (1949) planned

approach to curriculum. In general, the scientific-inquiry based learning cycle consists of three phases, originally called exploration, invention and discovery (Karplus & Thier, 1967). Through the use of laboratory or other forms of exploration, the first phase exposes students to the concept to be developed. During the second phase, the students develop the concept from the data through classroom discussion. In the final phase, students explore the usefulness and application of the discovery. In this particular sequence, these three phases may be necessary to develop conceptual understanding of the scientific topic (Tobin, Tippins, & Gallard, 1994).

With scientific inquiry-based instruction, teachers must help students focus on the process of an investigation and not just on getting the right answer (Harwood, Reiff, & Phillipson, 2002). In science the correct answer is only as valid as the methods used to arrive at it. Through staying focused on the inquiry process rather than its result, students learn to describe objects and events, ask questions, construct explanations, test those explanations against current scientific knowledge, and communicate their ideas to others (AAAS, 1993). They are then able to identify their assumptions, make them explicit, use critical and logical thinking, and consider alternative explanations to their predisposed notions or to their new discoveries (Waterman, 1998). Thus, with scientific inquiry-based instruction, students may develop a holistic understanding of science, which allows them to understand the nature, origin, and fluidity of scientific knowledge.

The National Science Teachers Association (2004), which calls for kindergarten through college teachers to embrace scientific inquiry, characterizes inquiry-based instruction as an approach that causes students to explore and to use the knowledge they gain to raise questions and to develop explanations for phenomena of the natural world.

In a study of scientists' notions of scientific inquiry, Harwood, Reiff, and Phillipson (2002) found that the most important characteristics of a science investigator are the ability to make connections between data, the aptitude to connect different disciplines around an investigation in order to ground it in a multidisciplinary context, a detailed focus on the process of the investigation rather than its outcome, and solid analytical skills.

As a result of a pedagogy that supports these types of behaviors, students should be able to ask questions that can be addressed with scientific investigations, design and conduct investigations that can answer scientific questions, analyze and interpret data, and think critically and logically to create explanations based on evidence and theory. The NSTA (2004) argues that by engaging in these activities, students will understand the following integral components of the scientific process: scientific investigations do not follow a fixed sequential series of steps; different kinds of questions may call for different kinds of scientific investigations; scientists' perceptions of the world are based on evidence that they collect; scientists must be skeptical when assessing their own work and the work of others; and the scientific community only seeks to understand and explain phenomena that are empirically grounded and logically consistent.

Because science research is necessarily and explicitly grounded in the findings of other members of the scientific community, scientific inquiry must be dialogically based. Thus, scientific inquiry-based instruction is sometimes referred to as cooperative learning (Kinney, 1989; Lord, 2001; Lin, 1998; Nesbit & Roger, 1997; Watson, 1991), which is characterized by five guidelines (Johnson & Johnson, 1991):

1. Constructive interdependence between members of the class or group.
2. Promotion of interaction between all students in the class or group.

3. Individual accountability where students cannot depend on the work of others for their own success.
4. Development of positive social skills such as leadership, constructive collaboration, effective and efficient communication, and decision-making.
5. Maintenance of effective working relationships between students engaged in activities. (pp. 19-20)

These principles can be achieved pedagogically by establishing mutual goals between students, maintaining appropriate group sizes, fostering complimentary roles in groups and within the classroom, and facilitating ongoing discussions about goals and processes. Such pedagogy can help students understand the nature of scientific knowledge, which allows them to understand that in addition to being empirical, testable, verifiable, and refutable, science is public and open for discourse in the scientific community and in the broader society (Waterman, 1998).

Lederman (1998) defines scientific inquiry as “the systematic set of approaches used by scientists in an effort to answer their questions of interest” (p. 2). Because scientific investigations are grounded in research questions, teachers who subscribe to scientific inquiry must facilitate students to ask questions that push them to think and talk about relevant topics that will help them progress in their learning. In other words, scientific inquiry-based instruction must begin with appropriate questions. Students may be excellent at generating general questions, but the scientific significance and pedagogical potential of a student’s question can be lost if the teacher fails to facilitate the transition from casual curiosity to scientific inquiry (Lucas, Broderick, Lehrer, & Bohanen, 2005). In order to engage students in inquiry behaviors, teachers must nurture and support student questions and dialogues that are productive guides to scientific inquiry.

Furtak and Ruiz (2005) have developed what they call the informative questioning cycle that helps teachers verify that students are progressing toward specific learning goals during the inquiry process. The authors contend that properly constructed questions that are posed at an opportune moment can help redirect and improve the quality of students' learning during the investigation process. The cycle involves three steps. The teacher begins by eliciting a response from students that reveals the level of understanding students have about a particular topic. Next, the teacher reflects the student's response back to her or asks another follow-up question to help the student explicate gaps in understanding. The third step involves taking some form of action to help the student move toward using an activity or previously known concept to resolve the gaps. Essentially, the purpose of the cycle is to help turn students' implicit doubts or quandaries into explicit scientific questions that the student can then work to resolve with scientific methods.

Explication of internal ideas is one of the greatest challenges to quality scientific inquiry-based instruction (Roehrig & Luft, 2004). Students may be able to generate justifications intuitively to support their actions or points of view, because these skills are often learned early in childhood. However, students often regard their actions or beliefs as unproblematic, even self-evident. As a result, students are often metacognitively unaware of gaps in their own logic and have a difficult time explaining their perspectives and understandings (Lucas, Broderick, Lehrer, & Bohanen, 2005). Scientific inquiry demands a deliberate separation between belief and evidence. Students must be aware that what they believe is only scientifically correct if it is based on evidence. The scientific inquiry process calls for developing a logical connection between students'

beliefs and scientific truths. It takes students beyond a perspective based solely on common sense by helping them generate evidence, theories, and hypotheses to justify their understandings.

The Value of Scientific Inquiry in the Classroom

Most scientists and science educators consider scientific inquiry-based instruction the most effective method for teaching all students science. Both the *National Science Education Standards* and the *Benchmarks for Scientific Literacy* argue that scientific inquiry promotes the learning, understanding, and retention of scientific concepts, and it cultivates a mastery of the process of doing science (AAAS, 1993; NRC, 1996). Thus, both documents claim that scientific inquiry-based instruction is the best method for teaching science to all students. In fact, numerous studies have shown that scientific inquiry based instruction improves science achievement, in both content knowledge and process skills, relative to traditional teacher-centered instruction, which includes lecture, reading, worksheets, and labs that validate the lesson (Branford & Donovan, 2005; Haury, 1993; Kinney, 1989; Light, 1990; Lin, 1998; Lord, 2001; Schmidt, Gillen, Zollo, & Stone, 2002; Watson, 1991; Zion, Michalsky, & Maverich, 2005). These studies suggest that students engaged in scientific inquiry behaviors not only have a deeper understanding of the scientific material, but they also retain the information they learn for a much longer period of time. Students' misconceptions about scientific principles may not always be dispelled with inquiry, but evidence suggests that when they are dispelled, they are dispelled more effectively (Bransford & Donovan, 2005). Overall, development of scientific inquiry skills in the classroom seems to promote a greater and more lasting understanding of scientific knowledge and methods.

In addition to increasing the depth of students' scientific knowledge base, scientific inquiry-based instruction can result in a paradigmatic shift in how students view the discipline of science. Lin (1998) finds that in addition to learning scientific facts, concepts, theories, and principles more effectively, students gain a much greater understanding of the nature of science. That is, students switch from an absolute view that favors a single truth, to a more relative view of science knowledge. They develop an understanding of the empirical nature of the discipline. As a result of cooperative learning strategies, students realize that scientific knowledge is constructed by the scientific process and community. Since knowledge is empirically based, not absolute, scientists have the opportunity to generate new knowledge with the scientific process. In other words, the scientific process precedes scientific knowledge, which is only as valid as the methods used to develop it (Zachos, Hick, Doane, & Sargent, 2000).

Students working in groups to perform challenging tasks in biology have been found to speak more often, ask more scientific questions, and be more engaged with the subject matter than students in traditional teacher centered biology classrooms (Lin, 1998). Students in cooperative groups engage in activities such as interaction with each other, sharing ideas, seeking additional information from outside sources, and making decisions collectively (Light, 1990). Light finds that shifting responsibility for learning onto the students provides learners with the opportunity to help each other through science content, and gives students a much greater appreciation for the subject matter.

Lord (2001) reports that when a teacher calls upon a student in a traditional biology class, the student becomes the focus of attention for the entire room. An error made by the student can become the subject of ridicule, which instills in students a fear of

being called on or speaking in class. In this manner, traditional pedagogy extinguishes classroom dialogue because students develop a fear of making mistakes. In contrast, Lord finds that in a cooperative learning situation, the focus of attention is diffused among the whole group that is talking. When the group (or individual within a group) makes a statement to the class, it represents the collaborative work of the team and, therefore, no single individual is held publicly accountable. In addition, the group members usually have an opportunity to ruminate over an issue before they present a statement about it to the whole class, which diminishes the prospect that a mistake will occur in the first place. Thus, if a correction is warranted, it is more likely to serve as a teaching tool for the whole class or group, instead of being a public admonishment of a singled out student.

This type of collective responsibility for knowledge can be particularly important in mixed racial and ethnic classrooms (Kinney, 1989). Because students in cooperative groups are involved in exploring issues and interacting with each other on a regular basis, they are likely to develop solid working relationships and friendships with one another. Thus, they may become sensitized to, and more understanding of, problems faced by other students. In addition to developing an improved social consciousness, Kinney finds that cooperative learning groups in multi-ethnic biology classes significantly raise achievement over traditional teacher-centered instruction.

While developing science content knowledge and process skills, scientific inquiry-based instruction has been shown to improve interdisciplinary skills, a characteristic which Harwood et al. (2002) determine to be part of the bedrock of scientific inquiry-based learning and science in general. In other words, the scientific process incorporates other disciplines, especially language arts. Researchers have

identified the language used in science explanations to be central to students' understanding (Sutton, 1992; Yore, Bisanz, & Hand, 2003). Further, exploratory writing, written field observations, rich description of phenomena, and written discussions of data are all essential components of scientific inquiry. Research suggests that science teachers have benefited from research on the writing process, and writing teachers have been shown to benefit from advancements in the theory and practice of scientific inquiry-based instruction (Ryan & Walking-Woman, 2000).

Nesbit and Rogers (1997) present several cooperative learning lessons that exhibit scientific inquiry methods. They specifically integrate science with reading and writing. Their research finds that these lessons significantly improve students' literacy and writing skills. Although research on students with special needs is extremely limited (McGinnis & Stefanich, 2007), special education students have been shown to benefit from scientific inquiry-based instruction (Schmidt et al., 2002). The pedagogy may help children with learning disabilities improve their achievement in both science and the language arts. Additionally, students may exhibit a sustained interest in science, a greater ability to focus in class, and more positive social interactions with peers.

Limitations of Scientific Inquiry-Based Instruction

Research has demonstrated that scientific inquiry-based instruction presents certain challenges and limitations in science classrooms (Bell, Blair, Crawford, & Lederman, 2003; Haefner & Zembal-Saul, 2004; Reiff, Harwood, & Phillipson, 2002; Roehrig & Luft, 2004). A fundamental challenge that often arises for teachers is the tension between engaging students in their own scientific inquiry and ensuring that the scientific content knowledge is covered. In a scientific inquiry-based classroom, teachers

and laboratory materials should guide students toward the discovery of the intended content. In practice, however, students often make “incorrect” discoveries, or their discoveries take a long time (Hammer, 1997). In order to be effective facilitators of discovery, teachers must be comfortable with uncertainty, or at least accepting of their own discomfort with it.

A further challenge exists due to the breadth of scientific inquiry techniques. Scientific inquiry encompasses all process skills that scientists practice in the field. The techniques used in scientific research are extraordinarily diverse (Reiff, Harwood, & Phillipson, 2002) and there is no consensus on the exact nature of science (Osborne et al., 2002). In other words, a single and concrete method for doing science does not exist. Nevertheless, Osborne et al. argue that the fluidity of scientific knowledge is in itself an argument for focusing on scientific process skills over scientific knowledge because they are more lasting and durable. Methods produce knowledge. In other words, through the implementation of scientific techniques and processes, scientific knowledge evolves. Therefore, scientific methods are of primary importance due to the fact that knowledge only comes as a result of them.

However, time constraints, resource limitations, and an inadequate foundation in the content and knowledge of the scientific processes have been shown to prohibit many teachers from effectively using a variety of scientific inquiry methods in their pedagogy (Roehrig & Luft, 2004). As a result, teachers, like scientists, must make choices in pedagogical techniques based on their own comfort and empirical realities in the classroom. Hammer (1997) suggests that teachers must be flexible when instruction goes “awry.” He argues that when a curriculum fails to go as planned, it is because the teacher

“misperceives the students’ participation or is unable to respond to what she or he does not perceive” (p. 503). As Hammer implies, communication between teachers and students is complex and difficult to predict, especially during the inquiry process. Many teachers may be unprepared for some of these challenges and, thus, may be reticent to engage in quality, open-ended scientific inquiry-based instruction.

Some inquiry-based teaching methods have been shown to be more effective than others in the classroom. Metacognitive techniques coupled with asynchronous science based activities in groups have been found to be the most effective at teaching both content knowledge and process skills (Zion, Michalsky, & Maverich, 2005). Zion et al. found that students learn best when they develop their own research questions and investigations and continuously scrutinize the meaning of their findings and observations. It may be reasonable to assume that the more students “do” science, the better scientists they become. However, Bell, Blair, Crawford, and Lederman (2003) find that apprenticeship programs, which pair students with scientists working in the field, only slightly improve students’ understandings of scientific inquiry methods, while critical concepts like flexibility of scientific methods and creativity remain unchanged. Ironically, mentors are innacurately convinced that their mentees learned a great deal about the scientific process through the apprenticeship (Bell et al., 2003). These findings seem to demonstrate the importance of proper pedagical preparation, in addition to knowledge of the scientific process, when teaching students science.

Therefore, another major obstacle to proper implementation of scientific inquiry-based instruction is the inadequate preparation of teachers. If a teacher is insufficiently prepared in the discipline, he or she is unlikely to be able to follow, let alone to facilitate,

complex discussions about inquiry-based findings. Students engaged in scientific inquiry often meander through unfamiliar arguments and express ideas using non-technical terms and expressions (McDermott, 1990). Therefore, a teacher who is just slightly ahead of the students in disciplinary knowledge would be unable to follow and push students through the discovery process (Roehrig & Luft, 2004).

Secondly, teachers who are trained in science usually are unfamiliar with inquiry-based pedagogical methods. Roehrig and Luft (2004) identify five main characteristics of secondary science teachers that are integral to proper enactment of inquiry-based instruction: a true understanding of scientific process skills and scientific inquiry, subject content knowledge, pedagogical content knowledge, teaching beliefs, and concerns about classroom management. Scientific inquiry requires an in-depth content and pedagogical knowledge base. Most science teachers learn their content area through traditional pedagogy and never practice as scientists in the field (Osborne, Collins, Ratcliffe, Robin, & Duschl, 2003). Thus, many teachers are unfamiliar with scientific inquiry and have almost no experience with its implementation. Specific training in scientific inquiry methods has been found to improve teachers' understanding of both scientific process skills and scientific inquiry pedagogy. Prospective teachers who undergo such education become more accepting of approaches to science teaching that encourage students' questions about observed phenomena (Haefner & Zembel-Saul, 2004).

Scientific inquiry-based instruction is not incompatible necessarily with traditional science pedagogy. Haigh, Frances, and Forret (2005) argue that scientific inquiry should be paired with traditional lecture style instruction in science classrooms to maximize content knowledge. They contend that in order for scientific findings to be part

of the scientific community's knowledge base, it must be presented in the paradigms and lexicon to which scientists subscribe. In other words, findings must be grounded properly within the existing knowledge base. Thus, students must possess a certain critical mass of knowledge in order for them to notice, think about, and understand scientific concepts and patterns, and then to communicate them properly with the greater community. In order for students to work comfortably with both scientific knowledge and process, the two must be intertwined pedagogically rather than kept separate (Bransford & Donovan, 2005). Hugh, Frances, and Forett (2005) seem to suggest that scientific inquiry-based instruction is more effective at teaching process skills, but traditional pedagogy may teach scientific content more efficiently.

Perhaps, the most costly obstacle to scientific inquiry-based instruction is the current science curriculum materials that promote traditional teacher-centered pedagogical approaches. Germann (1996) finds that biology textbooks rarely, if ever, enable students to use their knowledge and experiences to construct scientific questions, solve problems, investigate natural phenomena, or develop answers or generalizations based on findings. In fact, science textbooks often promulgate misconceptions about the scientific process by portraying it as a linear, step-by-step method that all scientists follow in their investigations. Reiff, Harwood, and Phillipson (2002) found that the scientific method portrayed in most science textbooks is completely inconsistent with scientific inquiry methods practiced by scientists. Rudolph (2005) explains that the linear method has been publicly disseminated since the early 1900s, when scientists were trying to enlist public support for a fledgling discipline. He argues that the focus on

accumulation of facts reflects a paradigm of late 19th century science and that it does not represent current science practice.

Assessing Scientific Inquiry

Assessment in science classrooms is determined as much by politics as it is by educational theory or psychometrics (Tamir, 1998). Within the current climate of educational accountability, the paradigm of large scale assessment of student learning is becoming increasingly dominant. Current standardized testing strategies may be inadequate for assessing scientific inquiry-based learning (Haury, 2001). At present, no generally recognized, systematic methodology exists for assessing the students' competence in conducting scientific inquiry (Zachos, Hick, Doane, & Sargent, 2000). However, Principled Assessment Design for Inquiry (PADI) is using technology to develop an evidence-centered approach to assessment design that can be used to measure inquiry-based learning (Mislevy, 2004). Developers of large scale assessments of inquiry-based learning face several challenges and barriers in their efforts.

One of the obstacles to large scale assessment of inquiry may be epistemological. Traditional science assessments may be appropriate epistemologically for measuring a student's science knowledge base, but insufficient for determining how well a student demonstrates scientific process skills (Zion, Michalsky, & Maverich, 2005). The epistemological grounding for an evaluation should be appropriate for the phenomenon to be evaluated (Aoki, 2005). Therefore, a quality evaluation methodology would remain germane to scientific inquiry-based instruction and the constructivist paradigm on which it is based (Berlack et al., 1992; Gipps, 1994). Further, an assessment tool should allow

researchers to assess different inquiry-based teaching methods and programs to illuminate what processes best facilitate student learning.

Multiple measures are needed to evaluate the depth and breadth of learning that is expected from scientific inquiry-based instruction (Herman, 1997). Science performance assessments can access scientific process skills that cannot be addressed with traditional written tests, but they bring their own challenges. They often do not adequately probe cognitive complexities (Baxter & Glazer, 1997); they are expensive to develop, administer, and score (Collins, 1993; Ruiz-Primo & Shavelson, 1996). Doran, Boorman, Chan, and Hejaily (1993) developed a large scale assessment of scientific inquiry skills, but after implementation in Ohio, they found it to need further refinement.

Zachos, Hick, Doane, and Sargent (2000) have developed what they call a framework for a universal assessment of scientific inquiry-based pedagogy. First, they identify a diverse set of indicators for measuring competence in conducting scientific inquiry. Then, they propose a set of indicators of success in attaining scientific concepts as a result of direct investigations into natural phenomena. In order to determine a causal link between inquiry behaviors and the scientific products of those behaviors, they test the relationship between each inquiry-based behavior and its success in making a scientific discovery. They find that comparative reasoning, the coordination of theory with evidence, and the disposition to search for underlying principles strongly correlates with success in making scientific discoveries. The authors do not address how their assessment guidelines might be administered as part of a large scale assessment system. Scientific inquiry-based instruction is not standardized. Perhaps, inquiry-based

instruction, because it is grounded in individual students' construction of learning and discovery, cannot be evaluated readily on a large scale.

Unsurprisingly, large scale assessment of scientific inquiry has presented significant challenges to departments of education in the current climate of educational accountability (Hickey, DeCuir, Hand, Kyser, Laprocina, & Mordica, 2002). Most large scale assessments are limited in the type of learning they can assess (Ketterlin-Geller, McCoy, Twyman, & Tindel 2003). This limitation is clearly a tremendous obstacle for evaluating scientific inquiry-based learning, which is fluid and expansive in its style and scope. One of the greatest challenges to designing an evaluation method for scientific inquiry-based learning is that it may paradigmatically contradict the epistemological tradition to which most assessment developers and the *No Child Left Behind* legislation subscribe (Pinar, 2005). The traditional input/output model for evaluation runs contrary to scientific inquiry based-instruction which focuses on the process of learning, not its outcomes. Traditional assessment is linear and preset, where students are ordained "to dance the same, to paint the same, to sing the same, to act the same" (Aoki, 2005, p. 418). Scientific inquiry requires students to be original, to develop knowledge, not to replicate it. Any effective evaluation strategy must embed evaluation into the curriculum process, rather than assess curricular outputs, so that it both reflects the learning and moves the learner to further understandings.

Despite the epistemological challenges, efforts at developing large-scale assessment of scientific inquiry-based learning are underway (Harden, 2005; Mislevy, 2004). In 2004, the Oregon State Department of Education developed a scoring rubric for assessing scientific inquiry-based learning. The scoring guide measures four dimensions

of scientific inquiry: forming a question or hypothesis, designing an investigation, collecting and presenting data, and analyzing and interpreting results (Harnden, 2005). These dimensions can be evaluated according to three integrated theoretical domains based on scientific process skills. First, students must use scientific reasoning skills when formulating questions, developing investigations, and analyzing results. Second, the reasoning skills must be based on scientific knowledge that students can use effectively. Finally, students must possess written, oral, and visual social process skills to communicate with one another while engaging in scientific inquiry (Ruiz-Primo & Furtak, 2004).

As part of the Oregon state testing system, students must submit inquiry work samples for evaluation in the fifth, eighth, and tenth grade in order to fulfill the state's interpretation of the requirements of *No Child Left Behind*. Each school district establishes guidelines regarding who performs the actual assessment, but the classroom teacher is most likely responsible for this evaluation. Some of the challenges present in using the scoring guide are as follows: the rubric seems to be more effective for assessing tasks in the physical sciences; the guide has been found to be biased against qualitative data; teachers have had difficulty transitioning from traditional recipe-style labs to student developed inquiry investigations; and many teachers are not sufficiently proficient in scientific inquiry methods (Harnden, 2005). Further, this scoring system relies on teachers evaluating their own students. Thus, it may be less effective or impossible in most other states, including Maryland, where teachers are not directly involved in carrying out the assessments.

In order to address the perceived limitations of having teachers assess their own students, an instrument using video analog technology was developed. Hickey, DeCuir, Hand, Kyser, Laprocina, and Mordica (2002) have developed two separate video assessment strands, summative and formative. Although they find that that video assessment can be used to compare groups of students, they express skepticism that their techniques could be implemented in a large scale accountability system. They find that in order to assess inquiry behaviors, the evaluator must have both personal and cultural knowledge of the students. Otherwise, there is too much room for misinterpretation of behaviors. As a formative assessment strategy, Hickey et al. find that students and teachers are able to learn a great deal about classroom scientific inquiry behaviors by observing videos of investigations and other collaborative interactions. Like Oregon's method, their approach may be ineffective as a large scale summative assessment where teachers are not involved in assessing their own students.

Scientific Inquiry and Accountability

Thus far, little research has been conducted on how educational accountability has influenced science instruction. Further research needs to focus on how science accountability measures mandated by *No Child Left Behind* will change the way that science is taught and learned. This review of literature offers a set of markers for examining science instruction as called for by national and state standards. This study examines how high-stakes standardized testing affects how teachers engage in scientific inquiry based instruction. In an effort to understand how teachers negotiate through competing policy signals, I seek to gain insight into whether the current manifestation of educational accountability is compatible and reconcilable with scientific inquiry-based

instruction. Further, I am interested in how accountability influences inquiry-based innovation in the classroom; how it influences teachers' efforts at incorporating laboratory investigations, student cooperation, and scientific writing into the curriculum; and whether accountability contributes to a shift in how teachers understand science and science instruction.

The State of Accountability

Should we aspire for excellence in schools? Of course we should. But in aspiring for excellence we need to weigh with care what understanding of excellence we are calling upon teachers and students to excel. (Aoki, 1990, p. 5)

In order to gain an understanding of how educational accountability is influencing the science curriculum, it is important to have a foundational grounding in accountability and its effect on education. As a set of policies, educational accountability is pervasive throughout the public education system in the United States, and accountability policies have been found to penetrate through the boundaries of school systems more than other state education policies (Malen & Muncey, 2000). The district accommodation of state and federal educational priorities has led to changes in curriculum at the school and classroom levels. The proliferation of accountability policies, rules, and regulations seems to be crowding out all other discourses in education (De Lissovoy & McLaren, 2003). Educational accountability aligns with Tyler's (1949) notion of curriculum implementation, where teachers implement, with as little interference as possible, what the state outlines as knowledge all students at a particular grade level must possess.

The accountability movement in the United States has been growing exponentially on two fronts since the late 1960s. On the one hand, the federal government has fostered accountability since the passage of the *Elementary and Secondary Schools*

Act (ESEA) in 1964. The ESEA required the testing of all students receiving Title I funds, and it mandated that the results of the tests be reported to the public.

Simultaneously, at the state level, policymakers began to develop competency tests that sometimes influenced grade promotion or graduation (Linn, 2005a). High-stakes standardized testing, however, really began to take off in the 1980s and beyond. Erickson (1986) documents an increase in accountability systems – including curricular management by objectives, continual and regular achievement testing, and more federal pressure on curricular and goal uniformity. Within twenty years, nearly every state has used high-stakes testing in some form (Hoffman, Asap, & Paris, 2001).

On multiple occasions, the ESEA has been restructured to redefine how the tests are administered and the results should be reported. In 1994, the federal accountability movement gained strength with the passage of the *Improving America's Schools Act*, which mandated the alignment of state content standards with assessments and required the testing of all students (Linn, 2005a). Nevertheless, the underlying premise of the act was undisturbed. The federal funds allocated by the ESEA to students in poor communities were not linked to the results of standardized tests (Rudelvidge, 2003). In 2001, the *No Child Left Behind Act*, the most recent restructuring of the ESEA, altered that principle. This latest growth spurt in educational accountability has married the standardized tests to federal funds. States and/or districts who refuse to test their students risk losing federal education funding. Currently, all schools must test each student in mathematics and language arts, and by the 2007-2008 school year, schools will be required to test students in science.

In addition to nationwide expansion, educational accountability has come to influence the curriculum powerfully at the school level. As the amount of testing has increased, so have the consequences associated with performance. By 1980, attempts at curriculum revision were already found to affect materials and teaching activities, while ignoring the wishes, habits, and needs of schools, teachers, and students (Atkin & House, 1981). Some of the other consequences are questionable labeling and grouping strategies to inflate test score statistics, general treatments and interventions on student masses rather than individuals, and exertion of psychological pressure on teachers and students to bring them into policy compliance (Craig, 2004). Current accountability schemes, as outlined by *No Child Left Behind*, are essentially systems for motivating teacher and school performance through severe and escalating consequences for failing to reach Annual Measurable Objectives (Herman, Baker, & Linn, 2004). For students in twenty-six states, including Maryland, high school graduation is or soon will be dependent on successful performance on state assessments (CEP, 2005).

Advocates of high-stakes testing are what Doll (2000) would call “blockheads.” They make positive assertions on the premise that good tests will inevitably drive good instruction, and good test results necessarily equate with quality education (Finn, 1993; Grant, 2001; Ravitch, 1995). As such, they are “uncritical, unquestioning, inauthentic, and exclusive” (Doll, 2000, p. 10). They refuse to see beyond the process-product model of curriculum and ignore the fact that teaching is not equivalent to transmission. Teaching and learning produce knowledge that is lived and experienced by both students and teachers (Jardine & Rinehart, 2003). This curriculum-produced knowledge (Pinar, 2002) cannot be predicted by standards or captured by a summative high-stakes assessment.

Interestingly, not only might the proponents of accountability be viewed as “blockheads,” but the system itself might be reflective of a “blockhead” mentality because it ignores the lived nature of the curriculum. Accountability, itself, is concerned mostly with the inputs (standards) and outputs (test scores) of education.

Because *No Child Left Behind* mandates states to develop their own accountability plans, systems across the country are different from one another. Because accountability systems in the United States are organic byproducts of separate socio-political movements, little consistency exists across states. Nevertheless, states do share some important similarities. Robert Linn (2005a) has identified the following seven assumptions that all U.S. accountability systems share:

1. Education quality is sub-par.
 2. Student learning must be improved.
 3. Student outcomes, not process, should be used to measure school quality.
 4. Content standards elucidate what should be taught and learned.
 5. Schools are directly accountable for the learning of all students.
 6. School accountability will motivate students and teachers.
 7. Public information provided by the system can improve teaching and learning.
- (pp. 1-2)

Although the manifestations of these assumptions differ greatly across states, LEAs and schools, the assumptions have serious implications, both positive and negative, for what occurs in the classroom (Stetcher, 2002).

Before considering the effects of accountability at the school and classroom levels, it is important to reflect on the greater structural foundations of accountability in order to develop a context for its implementation. Apple (2001) suggests that accountability policies are an educational manifestation of neoliberal and neoconservative free-market reforms. One of the features he highlights about accountability is its regulatory nature which relies on alignment of curriculum, standards,

and assessments. Curricular alignment has been organized around a concern for external oversight, regulation, and judgment of performance (Mentor, Muschamp, Nicholls, Ozga, & Pollard, 1997).

Apple (2001) takes this market driven construction of accountability further by suggesting that it is being colonized by parents who “possess what is seen as appropriate economic, social, and cultural capital” (p. 190). Thus, he argues that the accountability establishment can be linked to the neoconservative wish to return to a lost past of high standards, discipline, and real knowledge. Additionally, Apple suggests that in terms of social justice, families who do not possess the appropriate economic, social, and cultural capital are paradigmatically disadvantaged by the system. In a Marxian analysis of accountability, De Lissovoy and McLaren (2003) argue that by reducing learning to a test score, policy makers seek to make the knowledge of individuals commensurable. Then, knowledge can be handled and manipulated as a commodity that can be compared across students, schools, and districts. This reduction of knowledge to currency may reify the consciousness and creativity of students and teachers.

Curriculum and Test Alignment

Curriculum is a messy and unpredictable event that constantly exceeds *both* understanding *and* misunderstanding. (Ellsworth, 1997, p. 46, emphasis in the original)

States differ in their interpretation of federal policy mandates and in the formulation of their own policy responses to them (McDonell, 1994). Thus, implementation of accountability policies at the state level has been far from consistent. States have scrambled to align curricular standards with newly developed assessments. Alignment has historically posed a challenge for curricular developers, and assessments

have been plagued with criticisms regarding their validity and reliability since the 1980s (Mislevy, 2004). Current research has shown that existing state science assessments do not cover intended standards, especially those pertaining to scientific inquiry skills (Porter, 2002; Herman, Webb, & Zuniga, 2003). The National Research Council has published a framework for developing coherent state assessment systems (Wilson & Bertenthal, 2005). The framework envisions alignment between teaching, assessment, and standards based on a shared vision of goals at the classroom, school, state, and district levels.

Before examining the specific challenges associated with alignment, it is important to consider some of the epistemological concerns with the notion of curricular alignment in order to challenge some of its assumptions. The concept of curricular alignment is embedded in the technical curricular paradigm (Pinar, Reynolds, Slattery, & Taubman, 2002). The technical perspective of curricular alignment implies that teaching and learning are consistent and homogeneous across teachers, schools, and districts (De Lissovoy & McLaren, 2003). Obviously, this is impossible. Aoki (2005) distinguishes between curriculum-as-plan and curriculum-as-lived-experience. The designers of curriculum-as-plan consider teachers to be installers of a prescribed curriculum that is perfectly aligned with their standards, aims, and evaluation measures. The curriculum-as-lived-experience is what actually happens in the classroom. The lived curriculum is fluid, unique, and active. As such, it cannot be aligned to something created outside the classroom. Therefore, the notion of alignment is a construction of curriculum that exists outside the classroom in tests, standards, and other policy documents. It is likely to

remain in these places because curricular alignment can never really be embodied in the classroom.

The manifestations of this inherent contradiction are evident. Analyses of assessments have uncovered inconsistencies among assessment goals, developed tasks, and scoring criteria. Although some of the issues have been resolved, serious disagreements remain regarding how assessments and state standards can and should be aligned (Herman, Baker, & Linn, 2004). Most recently, validity and reliability of state assessments have been questioned due to tremendous variations between states in requirements for meeting Adequate Yearly Progress, as well as variations in scores on state tests and the NAEP exam (Linn, 2005b). Some efforts are underway to re-conceptualize assessments to match the subject standards better (Porter, 2002; Webb, 1999; Wilson & Bertenthal, 2005). Ironically, it is far from clear whether state standards or the high-stakes assessments drive the curriculum (Grant, 2001).

Another threat to the reliability and validity of accountability systems results from sampling complications. A particular sample of test-takers may not be representative necessarily of the general population. Although states have latitude in setting the minimum number of subpopulations that need to be tested in order to have their data reported, *No Child Left Behind* requires schools to report the test results of ethnic subgroups and the test results are used to determine whether schools have made AYP. Thus, scores from certain groups of students may be over-emphasized (Linn, 2005b). For example, in schools with small minority populations a few students may constitute an entire ethnic subgroup, which greatly inflates the value of these particular students' scores with respect to the remaining population in the school (Koretz, 2005). As a result,

teachers may be compelled to focus their instructional efforts toward small groups of students who are under-represented in the school population.

In science education, PADI is using technology to develop an evidence-centered approach to assessment design that can be used to measure inquiry-based learning (Mislevy, 2004). Science high-stakes assessments may be inherently inconsistent with the *National Science Education Standards*, which call for less emphasis on external assessments and standardized tests. The National Research Council (1996) endorses science learning through inquiry, deemphasizes science as a pre-established canon of knowledge, and advances science as a way of knowing and understanding. High-stakes tests, on the other hand, often focus on a broad body of predetermined knowledge and may treat science as a disconnected set of facts (Wideen, O'Shea, Pye, & Ivany, 1997). Therefore, high-stakes testing, itself, may contradict a principle of the national science standards.

Teaching to the High-Stakes Test

The demand for accountability, and its focus on efficient management, is a response to a strong mistrust of educators, especially teachers (Apple, 2001). Ironically, rather than completely controlling teaching, the accountability system may send mixed messages to teachers. Grant (2001) found that teachers are forced to choose between teaching curricular standards or the content of high-stakes assessments. Within accountability systems, test scores may become the bottom line, the single measure of educational success for students and teachers (McNeil, 2000). As such, standardized high-stakes assessments ascend to what may be considered the highest position in the classroom. Rather than being a tool of assessment, the test becomes a central feature in

the classroom. Thus, for teachers of a particular discipline, a conflict may arise when test preparation becomes a subject matter in and of itself, rather than a mode for evaluating students (Rex & Nelson, 2004).

Craig (2004) uses a high school principal's metaphor to describe the role of educational accountability in schools. The principal refers to standards and high-stakes testing as a dragon in his school's back yard. The school must "appease the dragon or face the consequences when it rears its fire-breathing head" (p. 1230). High-stakes assessments have been shown to affect what teachers teach more than content standards (Grant, 2001; Hamilton, 2003; McDonnell, 2004; Stetcher & Hamilton, 2002); therefore, when standards and tests are not aligned, teachers teach to the test rather than to the standards. The result is a backward relationship between assessments and curriculum standards. The influence of standards on curriculum is diminished, and assessments, whose purpose should be to measure student performance, instead drive the curriculum (Grant, 2001).

High-stakes tests may narrow the focus of the curriculum, cause teachers to teach to the test, and unnecessarily overburden teachers and students (Haury, 2001). Testing and test preparation infiltrates the curriculum at the expense of substantive work. Valli and Chambliss (2007) find that a teacher in a student centered classroom environment facilitates learning by building on students' experiences, prior knowledge, and culture backgrounds. However, in a reading intervention class focused on test preparation the researchers find that the teacher constructs "her role narrowly, as a test coach attempting to train students to perform well on the state assessment by staying close to a test-preparation script" (p. 71).

In a high-stakes accountability environment, district and school level discussions of success are often limited to analyses of scores on standardized tests (Booher-Jennings, 2005). When student learning becomes “subservient to testing, curriculum becomes bounded by what is testable, instruction wrapped around in bubble-in worksheets, and human worth (that of teachers, students, and institutions) determined by the rankings and ratings produced by the accountability marketplace” (Craig, 2004, p. 1240). A study in Texas, finds that teachers spend 8-10 hours per week on test preparation. In fact, the teachers report that they feel pressured to spend even more time on the tests (Hoffman, Assaf, & Paris, 2001). Teachers often reallocate their time to material that is tested more often, and they coach students on the specific test items or on strategies that can help them pass the standardized tests (Koretz, 2005). Koretz finds that instructional overemphasis on standardized tests may cause score inflation, which may result in misinterpretations of results.

Unfortunately, summative high-stakes tests are ineffective pedagogically, because teachers do not receive results in time to be informed about their students’ progress. Further, tests generally focus on measuring the amount of learning that has taken place. Thus, they usually provide little insight into how instruction may be improved (Black & Dylan, 2005). In other words, they evaluate the planned curriculum, not the lived curriculum. Perhaps this is why Firestone, Mayrowetz, and Fairman (1998) find that high-stakes tests influence the subject matter that teachers include in their courses, but they have limited influence on how the teachers teach the subject matter to their students. Thus, Black and Dylan (2005) find that an effective accountability policy would require the integration of both formative and summative assessments that include the active

participation of teachers. Boudett, Murnane, City, and Moody (2005) suggest strategies for using assessment results to improve classroom practices, such as finding patterns, communicating with students, and building on the reported results. Unfortunately, most accountability systems are set up such that teachers cannot readily follow the authors' advice due to the timing of the assessments and how their results are reported.

In most accountability systems in the United States, test results are reported in aggregate, so teachers may be unable to evaluate their ability to meet students' individual needs. Further, annual test results arrive too late for teachers to use them in the classroom with students whose knowledge is assessed. It may not be valid to infer that each year's students are the same as the previous year's cohort. If accountability systems do not incorporate regular classroom assessments that can be used to improve instruction for the students who are taking them, appropriate teacher decision-making regarding individual students cannot be assured (Herman, Baker, & Linn, 2004). Thus, when test results arrive after the school year ends, teachers may not be able to use the results of assessments effectively to provide supports for their students' learning.

In science education, accountability systems have been shown to shift the primary purpose of education from underlying structure and process of subject matter to the acquisition of skills and facts that are out of context. Wood (1988) has conducted one of the relatively few studies of the effects of high-stakes testing on science instruction in the United States. Junior high school teachers in Mississippi report feeling monitored and constrained by administrators and district officials. Because they are aware that they are being evaluated according to the standardized test results, teachers make sure that they teach the basic skills necessary to get through the standardized tests. Teachers alter their

behavior in the classroom in order to achieve uniformity in the curriculum. All students are taught the same skills regardless of ability. Teachers ask open ended questions, but they are looking for specific “correct” answers. Finally, teachers seem to accept the idea that science can be taught and learned as a series of disjointed facts outlined in a textbook. The experiences of teachers in Wood’s study contradict current research on quality science teaching and learning that stresses the importance of scientific inquiry and the nature of science (Haury, 2001; Wideen et al., 1997), and the national science standards that heavily focus on the scientific inquiry process (AAAS, 1993; NRC, 1996).

In addition to reducing the understanding of quality teaching and learning to a discussion of test scores, educational accountability may divert resources and attention to a narrow group of students. The Brickland, Texas school district has created three distinct categories to describe its students (Booher-Jennings, 2005). Students who score above 70% on the state test are called “passers;” students who score between 60% and 69% are called “bubble kids” because such scores leave the students on the “bubble” below the passing score; and students who score below 60% are called “foundation” or “remedial” kids. The districts’ resources, then, are focused on ensuring that the “bubble kids” pass the next benchmark test. Thus, not only may the curricular focus be reduced, but the target audience may be reduced to a fraction of the class.

An International Perspective

Other countries have implemented accountability systems similar to those in the United States. Therefore, a review of these systems may provide further insights into the U.S. high-stakes testing practices. For example, in England and Wales, the British 1988 Education Reform Act, in addition to mandating summative assessments administered at

the end of the school year, initially included formative assessments for teachers to administer and use to improve their instruction. When the act was implemented, however, publicly shared formative assessment scores were undervalued compared with those on the summative tests. Therefore, much of their meaning was lost. Teachers complained that the assessments took time away from their teaching, rather than improving it (Black, 1994). Although the state maintained the accountability system (Ranson, 1995), teachers, through collective public protest, were able to restructure the accountability policy to reduce the amount of testing (O'Hear, 1994).

School systems in all Australian states also have established testing policies and practices to raise levels of teacher accountability, but it remains doubtful whether the systems have the capacity to deliver on teacher quality assurances or improvement in instruction (Kleinhenz & Ingvarson, 2004). Kleinhenz and Ingvarson argue that the two main purposes of teacher accountability, assuring teacher quality and facilitating improvement in teachers' work, can be achieved when teachers and their professional organizations claim the responsibility for developing and implementing methods for evaluating teacher performance. With such evaluation practices, respect for the complexity and depth of teachers' professional knowledge and practice is overlooked. Their recommendations, however, take for granted that teachers have time to be involved actively in the development of assessments. Unfortunately, research has found that compared to other industrial nations, science teachers in the United States are overburdened and cannot make extra time to use test results appropriately, and certainly not to develop assessments or other evaluation strategies (U.S. National Research Center for TIMMS, 1997).

In British Columbia, Canada, Wideen et al. (1997) find that high-stakes science tests in high school decrease teacher autonomy and de-emphasize instruction that facilitates students' understandings of the nature of science. They also narrow the focus of the science curriculum. Teachers report that testing reduces opportunities for spontaneity, favors breadth over depth of learning, and decreases the number of labs performed. They believe that too much class time is being spent preparing students to pass exit examinations. Further, teachers report that students tend to ignore information that will not be tested. Many teachers complain that their science classes have been reduced to a collection of facts theoretically disconnected from one another (Wideen et al., 1997). In general, teachers feel that they have to abandon efforts at scientific inquiry-based instruction to teach to the test.

High-stakes tests may diminish teachers' expertise as representatives of the scientific community (Mislevy, 2004). Therefore, accountability measures may lead to loss of student and teacher agency and to the de-professionalization of teaching (Apple, 1992). In the top-down accountability structure, teachers are surrounded by those who are considered experts, and they are asked to subscribe to a technical perspective of curriculum which strips the teacher "of the humanness of his being, reducing him to a being-as-thing, a technical being devoid of his own subjectivity" (Aoki, as cited in Pinar, 2005, p. 3). Consequently, rather than being an active facilitator in the classroom, a teacher simply becomes an overseer who must insure compliance and adherence to standards established at the top of the educational bureaucracy. In other words, teachers become little more than test enforcers.

Reducing the Definition of Quality Teaching

Proponents of high-stakes testing, argue that when administered well, tests can serve to clarify goals, unify curricula, and foster a sense of community among teachers and administrators (Scott, 2005). Supporters tend to attribute student achievement and school improvement directly to the presence of accountability systems (Finn, 1993; Ravitch, 1995), without questioning the consequences of accountability measures. For example, in a study of large scale portfolio assessment in Kentucky, researchers found that ongoing assessments promote more collaboration among teachers and administrators, and they provide more opportunity for professional development for teachers (Berryman & Russell, 2001). Professional development has been found to improve scores on high-stakes tests (Browder, Karvonen, Davis, Fallin, & Courtade-Little, 2005). Berryman and Russell (2001) also found that large scale portfolio assessment is more successful when the assessment falls in line with the teacher's own values regarding instruction. However, neither of these studies critically examines the nature of the collaborations, professional development, or the high-stake assessments' definition of academic success.

Unfortunately, claims of accountability's success do not question the logic of accountability's reductionist effects on the curriculum. Most proponents of educational accountability do not assess critically whether high-stakes tests are actually measures of quality teaching. Instead, they define success according to the epistemological precepts of the accountability measures (Finn, 1993; Ravitch, 1995). In other words, proponents equate high test scores with educational quality. Such a construction does not examine what is missing from pedagogy specifically geared toward the improvement of test scores. An institutional environment that defines good teaching as the ability to improve

test scores may shape teachers' and administrators' ideas about what it means to be a "good" or "bad" teacher (Booher-Jennings, 2005). Booher-Jennings finds that high-stakes test results become the sole measure of good teaching and learning. Consequently, non-technical curricular paradigms may be displaced, and the public may be left to subscribe to a monolithic understanding of educational success.

In other words, accountability becomes an "epistemological monster" (Reynolds & Webber, 2004) that silences the cacophonous chorus of teachers, students, and other components of the classroom (Miller, 2004). Reynolds and Webber (2004) suggest that curriculum cannot be viewed through a single or even dual lens. Curriculum evolves through a series of multiplicities or what they call "lines of flight" (p. 2). Curriculum causes disarray, a lack of order or sequence. Being in the disarray induces learning. This reframing of curriculum opens it to examination through multiple directions, with multiple opportunities. As such, it cannot be defined, standardized, or assessed with a single measurement tool. And, it certainly does not fit within the epistemological boundaries of *No Child Left Behind*. In order to incorporate such an understanding of curriculum, an accountability system must be expansionist, not reductive in nature, and open to multiple curriculum discourses.

Even when accountability is considered through a single epistemological lens, different testing systems will influence pedagogy differently. Kentucky's reading portfolio assessments (Berryman & Russell, 2001), which promote teacher collaboration, likely consider reading instruction more holistically than a multiple choice test. Like the former MSPAP in Maryland, the Elementary Schools Performance Assessment (ESPA) in New Jersey is a performance oriented test. In order to prepare their students for the

test, New Jersey elementary school teachers report trying to get students to explain thinking more, and they are more likely to emphasize writing, even in math (Firestone, Monfils, Camilli, Schorr, Hicks, & Mayrowetz, 2002). However, these same findings that reveal at least some focus on the process of learning, may not apply to the Maryland's State Assessment and the High School Assessment, which are outcomes-based, not process-based tests (MSDE, 2004a). As such, Maryland's outcomes-based high-stakes testing system is unlikely to induce teachers to focus on scientific process skills as they prepare their students for the tests.

Science Education Accountability

According to the *No Child Left Behind* legislation, academic proficiency has, thus far, been defined and measured by state mathematics and language arts standardized tests. Teachers in other subject areas have remained partially spared from the reduction and strain posed by the high-stakes testing movement. For science teachers, however, the reprieve is over because *No Child Left Behind* requires states to develop standardized science tests for grade spans 3-5, 6-9, and 10-12 by the 2007-2008 academic year. The results of the tests may count toward AYP, and failure to make AYP can result in serious consequences for schools, teachers, and students (Flynn, 2002).

In order to facilitate a smooth transition to science testing, schools and districts have been revising or rewriting science curricula and implementing professional development programs to prepare teachers for the new high-stakes tests. Most states, including Maryland, are already testing students in science ahead of the 2007-2008 academic year deadline. These major curricular changes and institutional pressures are likely to affect science teachers and their pedagogy significantly. In fact, the explicit

purpose of educational accountability policies is to influence teaching (Finn, 1993; Ravitch, 1995).

Thus far, little research has been conducted on how curricular changes resulting from mandated testing have influenced the science curriculum. Further research needs to focus on how science accountability measures mandated by *No Child Left Behind* and interpreted by the states will change the way that science is taught and learned. In this study, I examine how high-stakes accountability tests coexist with science instruction. I examine in detail the process that science teachers undergo in response to the policy's messages. I examine how science teachers understand and respond to the institutional pressures and consequences outlined in the legislation, and how, if at all, they are resisting the curricular changes forced upon them.

In-cluding Policies in Teaching

Educational change depends on what teachers think and do – it's as simple and as complex as that. (Fullan, 1991, p. 1)

With all the complicated implicit and explicit messages emanating from policy makers and authorities on best practices in teaching, how do individual teachers incorporate all of the information into the curriculum? How do they deal with the tensions, if any, between what they know to be quality teaching and what is mandated by external policies? By understanding how teachers interpret what they have to do as they grapple with threats and mandates, we can understand better why teachers, even when supported by professional development, often translate policy into classroom practice in a manner that is inconsistent with policy's intentions (Olsen & Kirtman, 2002). Examining what is meaningful to teachers, how they act on those meanings, and the relative relationships between school actors can illuminate instantiations of policy

implementation as a situational praxis at the school or classroom levels (Aoki, 2005). In other words, the relative power of actors and their function within a school can be gauged by the roles they assume or prompt others to assume (Malen & Muncey, 2000).

Before exploring *how* teachers implement curriculum change, it is important to consider the meaning of curriculum implementation in order to inform teachers' practice in response to accountability mandates. Accountability, as construed by the standards movement and mandated by *No Child Left Behind*, subscribes to Tyler's (1949) technical notion of curriculum implementation, where teachers produce the policy makers' intentions in the classroom. Tyler's rationale, however, is not the only way to understand curriculum. Rather than being a product of knowledge developed outside the classroom, curriculum is both environment-producing and knowledge-producing (Pinar, 2002). Therefore, that curriculum is not just a product of implementation. Curriculum is about understanding the complexities of what goes on in the classroom (Pinar, Reynolds, Slattery, & Taubman, 2002). Teachers are not "technicians...[who] accept unquestioningly others' priorities" (p. 6). They are professionals whose work is the production of knowledge. Therefore, Tyler's expectations of teachers are unreasonable and impractical. Teachers are not automatons who do other peoples' bidding without placing their own imprint on the curriculum. They are agents in the manifestation of the curriculum. Thus, an examination of what is meaningful to teachers must look beyond instrumental curriculum implementation. It must include curriculum understanding, as situational praxis (Aoki, 2005) and as a form of resistance (Apple, 1996).

Situational praxis is rooted in the human sciences, at the intersection of theoretical and purposeful practice. According to Aoki (2005), praxis is a "way of knowing in which

the subject in a pedagogic situation...reflectively engages the objective world guided by the [theory] of ordering human action” (p. 116). Praxis is a holistic activity that engages the whole person – mind, body, spirit. Aoki outlines three assumptions underlying implementation as situational praxis that distinguish it from instrumental implementation:

- 1) A teacher is a subjective human being interested in his or her own engagements and interactions with students.
- 2) A teacher acts purposefully to transform and to engage in his or her reality.
- 3) Education is not neutral. The activity of curriculum implementation is a political act with social consequences. (pp. 120-121)

Implementation as situational praxis involves penetration of a particular curriculum into the lifeworld of the classroom. The teachers and students, guided by their “personal and group intentionalities,” cocreate the curriculum (Aoki, 2005, p. 121).

At the forefront of our top-down education system, teachers are ultimately those who are held responsible for implementing (from both an instrumental and situational praxis perspective) policies in the classrooms. They are what Lipsky (1980) calls street-level bureaucrats, or “public service workers who interact directly with citizens in the course of their jobs, and who have substantial discretion in the execution of their work” (p. 3). In the case of our education system, citizens are students, and in the case of accountability, the amount of discretion that teachers have may be diminished.

Nevertheless, teachers do interact on a day to day basis with their students. In fact, they often argue that they need to act in their classroom to protect and serve their students in ways that are not completely understood or appreciated by policy makers and other individuals outside the classroom (Rex & Nelson, 2004).

Olsen and Kirtman (2002) have developed a theory that positions the teacher as a mediator of school reform. The authors identify three strands of mediating influences: the

formal implementation process, school-wide influences that shape culture and climate, and individual influences on the teacher. All three of these strands interrelate to mold a teacher's disposition to implement particular reforms. The interaction of these three forces produces tremendous variations among teachers in a school, departments, schools adopting the same policy, and intended and actual outcomes. In effect, each teacher internalizes a set of (competing) messages and develops a product that reflects an interaction between an "interrelated tangle of influences" (Olsen & Kirtman, 2002, p. 303). This theory highlights the fact that teaching cannot be reduced to a simple product. Teaching is a praxis that includes contemplation, assessment, and cognition on the part of the teacher (Grundy, 1987). Ignoring these components of the curriculum misses the nature of teaching, which is a source of major tensions between curriculum-as-plan and curriculum-as-lived-experience (Aoki, 2005).

In addition to the complicated policy digestion that occurs in teachers' own subconscious, teachers are also outwardly responsive and accountable to public discourse. "The imposition of [state] testing locks the [state] curriculum in place as the dominant framework of teachers' work" (Ranson, 1995, p. 438). Although some space exists for resistance, if the public favors an accountability policy, teachers may feel compelled to engage in practices that they would otherwise reject (Apple, 2001). They may feel forced to engage with students in ways they would not consider, if not for the pressures to improve performance on the high-stakes test. As a result, their practice may become compromised, "less authentic, less ethical, less professional, and less pedagogically sound" (Craig, 2004, p. 1243).

In a technical top-down educational policy implementation system, when curriculum is changed from the top, teachers may be left de-skilled by the new policy, and, as a result, are likely to oppose the new curriculum (Grundy, 1987). Sociologist Bourdieu (1977) constructed a theory of teacher action in relation to educational practice. Bourdieu suggests that teachers behave purposefully, but their purpose may or may not conform with policy rules or objectives. His theory may help explain how teachers react to high-stakes tests. On the one hand, they want to help their students score high on high-stakes tests, but, on the other hand, they may be conflicted about “teaching to the test.” In their daily practice and interaction with students, teachers may be faced with a barrage of messages that they may or may not completely understand. While they behave purposefully, their choices may not always be based on rational or consistent calculations. Bourdieu argues that when faced with competing or ambiguous messages, teachers act pedagogically according to their own constructions of what constitutes good teaching in those narrow situations. Teachers constantly make quick decisions based on their own understanding of the context in which they find themselves. As such, they often may feel compelled to make decisions about what practices are best for their students within an educational atmosphere rife with inconsistencies and ambiguities.

Hursh (2005) argues that proponents of educational accountability have been able to overcome resistance by arguing that the new reforms are necessary to improve economic productivity, decrease educational inequality, and provide assessment objectivity. I already have outlined some of the possible tensions that are embedded in the accountability systems. Resistance to educational accountability measures is beginning to emerge from educators as a response to the myriad tensions and problems

embedded in the accountability movement (Hursh, 2005). Grassroots efforts have prompted twenty-five national organizations – representing children, teachers, administrators, and school boards – to issue a joint statement objecting to aspects of *No Child Left Behind*. Currently, the National Education Association (NEA) and eight school districts from across the country are filing a lawsuit against the U.S. Department of Education. They argue that the federal government is not sufficiently funding its own mandates (Dillon, 2005). Fifteen states have considered legislation to “opt-out” of *No Child Left Behind* and to forgo federal education funds, and four states considered bills that would prohibit the use of state money to comply with *No Child Left Behind* (Peterson, 2005).

The legislation’s harsh punishments make non-compliance by schools difficult, but some schools have not surrendered to *No Child Left Behind*’s prescriptions (Dillon 2005; Hursh, 2005; Waldorp, 2003). When schools fail to meet the requirements of *No Child Left Behind*, they are labeled as “failing.” The labels may be completely inaccurate, and may they may unfairly stigmatize students and schools. In Florida, three-quarters of the state’s top-performing schools were labeled as low-performing under *No Child Left Behind* (Hursh, 2005). Some schools have written to their districts and states, appealing these labels, but thus far, no school’s label has been changed (Dillon, 2005). Further, teachers, themselves, are often mislabeled by the legislation (Craig, 2004). Oregon has one of the nation’s most stringent teacher licensing requirements, yet about one out of every five teachers in the state does not meet the federal definition of a highly qualified teacher (Waldorp, 2003). Schools and teachers may take great efforts to avoid the

policy's negative ramifications, and their efforts are likely to be complicated due to policy's mistranslation during implementation.

Defining the Teacher

Schools and school systems are contextualized entities in which specific types of knowledge and understandings are embedded (Clandinin, 1986). Teachers' beliefs, understandings, and constructions of their own role are strongly influenced by their surrounding culture. Their practices and identities are subject to the social control of the institutions in which they work (Bidwell & Yasunmoto, 1999), and their relationships with other teachers and administrators condition the subject matter they teach and their pedagogical practice (Bryk & Schneider, 2002). In other words, the pedagogical paradigm established by a policy's objectives, and the culture called for by its implementation, are likely to influence much more than the predicted outcomes. They may influence the entire process, including how teachers see themselves, understand their practice, and relate to their colleagues.

Bryk and Schneider (2002) maintain that collegiate relationships within schools strongly influence how well schools function, especially in periods of reform. Bryk and Schneider suggest that throughout their everyday practice, school stakeholders play largely independent roles. Teachers, especially, are isolated in the separate domains of their classroom. Thus, any interactions between teachers and parents, administrators, and other teachers are colored by their attempts to maintain the boundaries established by their work functions. In order to maintain the boundaries, teachers must have a clear understanding of their own role and the roles of the other members of the school. Relationships between members of a school stay strong as long as everyone involved

perceives that others are maintaining their individual roles and fulfilling the responsibilities of their domain. Bryk and Schneider identify four key elements for maintaining trust and social cohesion within schools: respect, competence, integrity, and personal regard for others. A policy like educational accountability that individualizes teacher performance, based on the publication of student test scores rather than fomenting a sense of collective responsibility for all students, may change the relationships between teachers and all school staff (Booher-Jennings, 2005).

Much of the research on educational accountability has focused on the effects of policies rather than on the processes that influence their implementation. Specifically, little attention has been given to how and why teachers change as a result of accountability. Some studies attest to the motivational effects of accountability systems, which integrate both incentives and sanctions to promote student achievement (Abelman et al., 1999; Firestone, Mayrowetz, & Fairman, 1998). On the other hand, other researchers contend that negative outcomes result from sanctions. Threats from the state, districts, and administrators force educators to change their pedagogy (Craig, 2004; Jacob, 2002; Horn & Kincheloe, 2001). Further, accountability systems that include penalties for students, such as failure to graduate, may cause profound shifts in the curriculum because most teachers, first and foremost, value the need to help their students (McNeil, 2000). Despite their focus on the curriculum, these studies still concentrate on outcomes of teacher practice, rather than on examining the process through which teachers change their understanding of pedagogy and their own function in the classroom.

Power Struggles

Schools are located within a system of complicated social, cultural, and political contexts that can either reinforce or counteract state policy objectives. In other words, local forces can overwhelm policies (Malen & Muncey, 2000). Thus, school cultures and climates can make or break policy outcomes. Since teachers believe that they are pedagogical experts who know what is best for their students, they expect schools to support them in their work. They consider administrators responsible for buffering teachers and helping them negotiate through policy pressures as they face pedagogical challenges (Rex & Nelson, 2004). In other words, although they are experts, teachers do not see themselves as able to do the entire job of educating their students alone. They need the backing of the administration to support them in their work. There is evidence, however, that accountability policies create tensions between administrators and teachers. Jacob (2002) finds that educational accountability often induces administrators to exert psychological pressure on teachers to bring them into policy compliance, which in effect adds an additional tension for teachers to grapple with in their day to day practice. Rather than providing the foundation that teachers rely upon to do their difficult work, administrators become an added stress on the teachers through which they must negotiate.

Teachers may resent policies that infringe on what they believe to be good pedagogy. Nevertheless, their territorial instincts may be weighed against the interests of their students. Thus, they often decide that teaching according to policy demands and being loyal to a school's mission are more important than their resentments regarding policy mandates with which they do not agree (Rex & Nelson, 2004). In fact, some

teachers have been found to refer students for special education in order to reduce liability to their school's accountability rating (Booher-Jennings, 2005).

This example may be just one instance of the moral backlash that results from accountability policies. Apple (2001) argues that in an educational climate where each teacher is induced to act as a separate unit trying to increase his or her individual ranking as a quality teacher (as outlined by the accountability system), the competitive nature of the individual dominates and social justice is somehow meant to resolve itself. Most schools within the accountability environment engage in "unreflective and at times self congratulatory policies around markets, standards, [and] testing" (p. 192). Therefore, it is not surprising that teachers may compete within the accountability paradigm, and that the competition may come at the expense of their own notions of quality pedagogy (Rex & Nelson, 2004).

When teachers fail to integrate their conception of quality teaching with what is mandated by a policy, they may become anxious. In a study of math teachers (Craig, 2004), when test scores demonstrate that student performance is decreasing, teachers (despite their support for a curriculum that is not influenced by high-stakes testing) express overwhelming sadness. They abhor being tagged with the label of "failure." One teacher describes herself as feeling unglued, as if she were "swirling around in unknown territory" (p. 1248). Thus, Craig finds that teachers may comply with accountability in order to avoid its negative penalties. Such instances of dutifully begrudged compliance allude to the complex inner struggle that teachers have as they implement policies that may violate their understanding of good pedagogy. In essence, they are stuck between

frames of curricular reality: acting like beings-as-things and beings-as-human (Aoki, 2005).

Teachers may have a difficult time integrating a policy's objectives into the curriculum because they often are excluded from the curriculum and policy conversations. They are not allotted the time to influence policy, nor are they challenged to articulate their views on practice (Craig, 2004). In other words, teachers are not given the opportunity to integrate a policy into their own understanding of curriculum practice. As a result, they may be de-skilled by the policy (Apple, 1992). In order for teachers to be true and ethical to their own pedagogy, they must translate the policy into their own teaching methods and styles. Knowledge and understandings are embedded and shaped by specific contexts (Clandinin, 1986). Thus, opportunities to process policy messages would likely allow teachers to incorporate policy objectives more purposefully into their own lived curriculum.

Aligning Policy with Pedagogy

Because high-stakes testing is intended to control and reform educational practice (Glassnap & Poggio, 1991), it is reasonable to assume that pedagogy will be influenced more or less, depending on how closely teachers' practice, knowledge, and beliefs align with or deviate from the philosophies of the accountability system (Cohen & Ball, 1990). In other words, teachers' preconceived notions about teaching and learning are likely to affect the influence of accountability in the classroom. Some teachers can almost wholly bar the influence of testing from their classroom (outside of its actual administration). Others attempt to resist but fail to bar its influence, which results in fragmented instruction that sends mixed signals to students. And other teachers see little need for

resistance because their understanding of pedagogy aligns well with the high-stakes testing model (Zancanella, 1992). The type and magnitude of pedagogical influence appears to be dependent upon the degree to which teachers' conceptions of the subject align with the tests' conception of the subject, and the amount of curricular power – characterized experience, status, and position – that the teacher possesses (Olsen & Kirtman, 2002).

The interaction of curricular alignment and power may be particularly problematic in a field like science where there are two distinct and conflicting paradigms of teaching and learning, positivist traditional instruction and scientific inquiry-based instruction (Hammer, 1997). Because teachers may have very different, even opposite, understandings of how science should be taught, internal conflicts regarding what constitutes the best pedagogy are likely to surface in response to a high-stakes test that represents a single perspective of the science curriculum (Zancanella, 1992). Both curricular power and curricular alignment are likely to shift toward conceptions espoused by the accountability system. Teachers' views of their own pedagogy and outsiders' views of their practice may be influenced by the high-stakes test. If they do not agree with the curricular perspective espoused by accountability, teachers may become disenfranchised by the messages emitted from high-stakes accountability policies because the messages conflict with their own notions of quality teaching (Apple, 1992).

Examined from another perspective, even if teachers are committed to creating a particular set of practices in their classrooms, such a commitment does not necessarily guarantee that such practices will emerge in the curriculum. In fact, many teachers who espouse progressive, democratic, and inquiry-based practices may resort to traditional

teaching methods in their classrooms (Segal, 1998). The reality of the classroom and its real-time pressures can transform progressive beliefs into lecture-style traditional practice (Widlack, 1985). In effect, teachers hide behind their pedagogy rather than emerge from it. Educational accountability policies attempt to elicit high performance and compliance based on threats, which translate into teachers' fears of their own failure (Craig, 2004) and the failure of their students (McNeil, 2000). The notion of "high-stakes" may create a constant presence of uncertainty in the classroom, and teachers are likely to respond in ways that work against their pedagogical beliefs. Thus, stress, anxiety, and fear may shape the pedagogy of teachers in the classroom, located within a system that subscribes to the current conception of educational accountability (Segal, 1998). How teachers adapt and respond to such pressures is a focus of this study.

Summary

Although policy and standards documents may advocate inquiry-based instruction in science, efforts at quality inquiry-based pedagogy are likely to be stifled by high-stakes accountability. In reading and math, high-stakes accountability has been shown to significantly influence pedagogy, causing teachers to teach to the test. High-stakes testing may influence teachers through fear, induce stress and anxiety, and cause teachers to be de-skilled. Therefore, it is important to critically challenge some of the technical assumptions of high-stakes accountability. As a researcher, I hold the following normative assumptions: a curriculum cannot be developed outside of the classroom; pedagogy, standards, and tests cannot be perfectly aligned; and students cannot be evaluated meaningfully by a single measure.

Evidence suggests that accountability policies reduce quality in the classroom. Although a focus on test preparation may raise test scores, it is likely to limit and narrow teaching and learning. Quality teaching is integral to achieving the stated aims of *No Child Left Behind*. If they are to be pedagogically effective, teachers cannot feel mistrusted or undervalued by the system of education. In order to improve educational quality for all students, accountability policies should encourage teachers to reach out to all students, but evidence suggests that they cause many teachers to retreat into a narrow pedagogical focus (Craig, 2004; Booher-Jennings, 2005; Koretz, 2005). Because the current manifestation of accountability may call teachers to act in ways that contradict their own understandings of quality teaching, teachers are likely to find themselves embroiled within a system of competing messages and signals.

The process of negotiating through a field of inconsistent policy messages may be paralyzing; teachers can be stuck in a minefield of all the conflicting messages they receive. Evidence suggests that current high-stakes accountability policies are able to infiltrate into the curriculum more readily and powerfully than most policies that attempt to control teaching and local governance (Malen & Muncey, 2001). Because of the conflicting terrain between the beliefs of what they should be doing and what they are being told to do, teachers also may waiver about what they could do. There may be no doubt in their minds that they should be accomplishing something, but teachers may be less sure about what goals could be accomplished and when (Rex & Nelson, 2004).

In the field of science education, with two diametrically opposed constructions of teaching and learning, teachers particularly may be confused by the barrage of signals they receive. Consequently, teachers may not be exercising all the pedagogical options

that they do have. Yet, they are doing something. They are teaching science, and they do have conceptions about both education and the discipline of science. They are in the classroom. They interact with students. How do they reconcile the mandates of accountability and their own conceptions of best teaching practices? How do they reconcile potentially opposing messages and understandings about teaching and learning? How do they construct the purpose of their interactions with students? What is the context of their interaction with students when uncertainties exist about what and how they should and should not teach? How is the outcome of these internal tensions manifested? The purpose of this study is to examine how teachers internalize these tensions, live in them, and implement them as situational praxis. In the next chapter, I provide the methodological grounding and framework for this study.

CHAPTER III: METHODOLOGY

Introduction to the Study

The aim of this study is to explore how biology teachers understand and construct their practice in a high-stakes accountability environment that is likely to be riddled with tensions. Understanding teachers' experiences and the perspectives they hold regarding their work requires a rigorous examination of the possible conflicting messages, implicit and explicit, present in the policy, and how these messages shape their practice.

Teachers' beliefs, knowledge, and experiences are important factors that inform their interactions with students and the pedagogical and science content choices they make in the classroom. However, the classroom also is influenced by external conditions like the school climate, expectations of the administration, and district and state policies (Craig, 2004). Taken together, the multiple factors that influence what teachers do and how they perceive what they do create a complex context that lends itself to examination through a qualitative research paradigm (Merriam, 1998).

Which qualitative methodology can examine most fruitfully, the effects of policy messages on teacher understandings of their practice? In choosing the methodology, two issues need to be addressed: the epistemological orientation of the researcher, and the type of research questions being examined. In a constructivist paradigm, all research must deal with the social origin of meaning (Crotty, 1998). Crotty contends that "While humans...may be described as engaging with their world and making sense of it, such a description is misleading if it is not set in a genuinely historical and social perspective" (p. 54). Case study methodology allows the researcher to do just that. In a case study, meaning is grounded in a particular bounded system that must be understood and

described in order to illuminate the phenomenon of study (Merriam, 1998). In other words, the knowledge gained from the research is fundamentally and contextually based in the environment from which it is derived.

In terms of the research questions, each research methodology bears its own procedures, concepts, strengths, and limitations. As Shulman (1988) writes about different forms of research, it “is not only the procedures they employ, but the very types of questions they tend to raise” (p. 6). Like Yin, Shulman further argues that “what” questions are generally quantitative in nature, while “how” questions tend to lend themselves to case study work. Case studies can provide a rich examination of an issue or issues in a particular context. They “document or portray the everyday experiences of teachers” (Shulman, 1998, p. 8). In other words, case studies can use a specific “case” to inform a broader phenomenon, which Patton (1990) defines as “what it is you want to say something about at the end of the study” (p. 168). This case study focuses on biology teachers within the context of a Buckley County (pseudonym) public high school in order to examine teachers’ perceptions of and responses to particular educational accountability policy messages. This study lends insight into the implementation of the current high-stakes accountability system.

Purpose of the Study

The purpose of this study is to examine how biology teachers understand and construct their practice in a high-stakes accountability environment that is likely to be riddled with tensions. Research on the implementation of science accountability measures in the classroom is critical to providing insights that may improve the teaching and learning process. As the public school system is constructed presently, teachers are the

street-level bureaucrats, or the implementers, of the accountability system (Lipsky, 1980). As such, their interpretations of the policy's messages and their resulting practice are the face of how the policy becomes embedded in the curriculum. The policy's signals are likely to challenge teachers' pedagogical and content beliefs. Related research has shown that the current high-stakes standardized testing model may not be aligned well with the science pedagogy that is advocated by national, state, and local school districts (Stetcher, 2002). Thus, teachers may find themselves negotiating their way through competing and possibly irreconcilable institutional signals.

Because Maryland's accountability system has high-stakes outcomes attached to them, teachers are likely to succumb to the overt pressures presented as threats to their own careers and to their students' academic success. At a minimum, science teachers must develop mechanisms to negotiate their way through a maze of policy structures and signals. A study that examines how teachers confront these challenges and understand their practice in light of them can provide valuable insights for the implementation of science accountability programs throughout the United States in response to the mandates of *No Child Left Behind*.

Research Questions

The overarching research question that is the focus of this study asks: **How do high school biology teachers negotiate the explicit and implicit messages regarding high school biology accountability policies governing their work?** This question is examined and elucidated through three auxiliary questions and sub-questions:

1. What, if any, tensions are embedded within and across the national science standards, Maryland State science standards and the high-stakes assessment, and the district's biology curriculum?

This question includes three layers of policy messages – the national, state, and district levels that are designed to influence science teaching. I conducted an analysis of documents generated at each level of the system (described later in this chapter) to uncover underlying assumptions regarding science content and pedagogical practices. Although there is some overlap across the three layers of the system, there are also different emphases and assumptions, both between and within them. In particular, the Maryland state science standards, the High School Assessment, and the Buckley County Biology Curricular Frameworks are not perfectly aligned in terms of content and pedagogy. Additionally, the *National Science Standards* and the *Benchmarks for Science Literacy* were written by separate organizations, and although their congruency is reported to be “remarkable” (Leonard & Penick, 2005), they are likely to deviate from one another in some substantive respects. A careful analysis, of these three policy layers provides a major source for the identification of conflicting themes and tensions.

2. What, if any, tensions do biology teachers perceive in these messages or between these messages and their construction of teaching?

I do not assume that teachers perceive tensions simply because I identify them. It is likely that teachers are unfamiliar with all or some of the documents that I examine. Additionally, one policy layer may be more important in their daily lives than the others. Further, teachers may perceive and consider other tensions based upon their pedagogical or content knowledge, beliefs, and experiences, as well as the school’s social and professional environment. Although teachers may perceive multiple conflicting messages, I cannot assume they are necessarily the same as the ones that I uncover in my analysis of the three policy layers outlined in auxiliary question one. In other words,

teachers' perceptions of accountability policies may not necessarily coincide with their conception of quality teaching.

3. How, if at all, do biology teachers' interpretations of the implicit and explicit policy messages shape their teaching practice?
 - a. How, if at all, do biology teachers adjust their teaching in response to the accountability system?
 - b. How, if at all, do biology teachers adjust their understanding of the discipline of science to conform to the messages they receive from the accountability system?
 - c. How, if at all, do biology teachers construct and understand their role in the classroom in response to the accountability system?
 - d. How, if at all, does the accountability system influence biology teachers' passion for their profession?

This final auxiliary question and its sub-questions seek to determine how teachers respond to the messages they perceive. As developed in the review of literature, teachers are likely to respond to curricular policy messages differently, both in magnitude and in kind. In other words, some may react more strongly than others, and their reactions may be quite different. This question seeks to understand how teachers are influenced by the messages that they perceive. This auxiliary question is intended to uncover the most nuanced findings, because each teacher is an individual who processes messages from his or her own perspective. At the same time, teachers' reactions provide some insight into how the science accountability policy is manifested at the classroom level. In other words, because accountability policies seek to influence teaching, teachers' experiences are integral to understanding the effects of the policy.

In summary, the order of my questions proceeds as follows: First, I examine what tensions are embedded within the standards and accountability movement in science. Next, I uncover what tensions teachers perceive. Finally, I explore how these perceived messages manifest themselves in teachers' practice. These questions are intended to

inform the complicated relationship between education policies and teaching practice. Through an interpretive analysis of these questions, I reveal the tensions, conundrums, paradoxes, and compatibilities that exist between how teachers understand and construct their practice and externally imposed accountability policy regulations on their performance.

Methodological Grounding for Case Study Research

In order to develop a specific method for investigating this phenomenon, it is important to develop a proper epistemological understanding of the methodology to be employed in order to contextualize the knowledge gained from the study. All research methodologies have theoretical perspectives that guide the processes and procedures associated with them (Crotty, 1998). In other words, in addition to having its own set of methods and techniques for conducting research, each methodology bears its own set of assumptions about the nature of knowledge, its presuppositions and foundations. When conducting investigations, researchers hold specific assumptions about whether and how knowledge and understanding are socially embedded. A particular methodology allows the researcher to uncover and examine knowledge through a particular paradigmatic lens. Thus, examining the same issue with multiple methodologies bearing different theoretical perspectives may render very dissimilar results.

Case study methodology is one of the research methodologies that can be used for an interpretive study of teachers' experiences. As a combination of research methods, a case study is broadly defined and used in a multitude of social and educational settings. One of the greatest strengths of a case study design is its call for an "intensive description and analysis" of a specific case to gain an "in-depth understanding of the situation and

[its] meaning” (Merriam, 1998, p. 19). Case studies allow for an examination of a phenomenon as it is manifested in an actual context. Yin (2003) suggests that the “case study method allows investigators to retain the holistic and meaningful characteristics of real-life events” (p. 2). In other words, case study methods inherent in its methodology allow researchers to examine the case *in vivo*, like the behavior of a living organism in its natural habitat. Thus, case study methodology affords me the opportunity to study the effects of science accountability measures on teachers within a particular school, district, and state climate.

The significance of an opportunity to study a phenomenon embedded within a real-life case cannot be taken for granted. The researcher’s epistemological perspective is integral to understanding the particular habitat in which a phenomenon resides and how the phenomenon can be uncovered within a particular social context. According to the constructivist paradigm, which holds that knowledge and meaning are “constructed rather than discovered” (Stake, 1995, p. 99), an investigator must seek “culturally derived and historically situated interpretations of the social life-world” (Crotty, 1998, p. 67). As a researcher who holds a constructivist epistemological perspective, I consider all knowledge to be socially embedded. Therefore, an interpretive approach to research allows the investigator to gain a greater understanding of the phenomenon itself *and* its meaning to those who experience it. In other words, the phenomenon cannot be fully understood in and of itself. Its meaning must be filtered through the lens of individuals who experience it. It is the interaction between the researcher and the participants that leads to the generation of data (Guba & Lincoln, 1989).

Constructivist researchers conduct investigations guided by their beliefs in multiple realities and truths that are discovered by the interactions between researchers and respondents (Guba & Lincoln, 1989). Stake (1995) outlines three conceived realities.

One is the external reality capable of stimulating us in simple ways but of which we know nothing other than our interpretations of those stimuli. The second is a reality formed of those interpretations of simple stimulation, an experiential reality representing external reality so persuasively that we seldom realize our inability to verify it. The third is a universe of integrated interpretations, our rational reality. (p. 100)

The aim of research based on the constructivist paradigm is not to discover the first reality because that is impossible. Rather, it is to construct a clearer second reality and a more sophisticated third reality in order to build an integrated understanding.

Max Weber suggests that human sciences strive for understanding that is substantiated by empirical evidence. He argues for the development of appropriate methods devised to uncover understanding with rigorous methods and deduction. He calls this methodology the “ideal type,” which serves as a heuristic tool for studying social phenomena. Rather than being a set of tangible procedures, Weber’s “ideal type” methodology calls on the researcher to develop conceptual constructs framed for studying a specific societal situation or event (Crotty, 1998). The myriad possibilities embedded in case study methodology allow the researcher to use theory to frame an investigation that matches the nature of the phenomenon and the social conditions in which it is embedded.

The constructivist paradigm, coupled with the contextual nature of case study methodology, informs how I approach this study. All teachers have unique and individual experiences. Many perspectives exist about educational accountability and the messages it sends to teachers. In addition, the multiple tensions that teachers perceive are socially

constructed, and filtered through their own consciousness. In light of this epistemological belief, I focus in great detail on how teachers understand curricular policy tensions that inform their pedagogy. By developing an understanding of how teachers experience and formulate their own realities, I seek to gain insight into the process through which teachers adapt to or resist the accountability climate. The specific intent of my research is to gain a deeper understanding of how teachers “make sense” of the science accountability system (Guba & Lincoln, 1989, p. 89).

Bruner (1986) distinguishes between two modes of thinking: the narrative mode and the paradigmatic/logico-scientific mode. The narrative mode uses lived language to convince the readers through vicarious experience. Narrative work strives to make meaning from participants’ experiences. It is concerned with a search for meaning among many possible meanings. The narrative mode allows me to gain an understanding of teachers’ experiences with high-stakes accountability, and to present my understandings clearly and vividly. Narrative language and description allows readers to feel and appreciate the complexities and paradoxes that teachers face in the high-stakes accountability educational climate.

A Critical Perspective

The difference between an interpretive lens and a critical lens is fundamental in intention. According to Crotty (1998), “It is a contrast between a research that seeks merely to understand and a research that challenges...between a research that reads the situation in terms of interaction and community and a research that reads it in terms of conflict and oppression...between a research that accepts *status quo* and a research that seeks to bring about change” (p. 113). This study seeks to go beyond interpretation and

understanding. I seek to challenge some of the structures of high-stakes accountability and its effects on teachers and, as a result, on students. As such, I include a critical lens in my interpretation of biology teachers' experiences with accountability.

Because high-stakes accountability seeks to influence teaching through top-down policy implementation, power over districts, schools, and teachers may be germane to the policy. Saul (1992) argues that power arises from how knowledge is used rather than its effect on people. The power of accountability policies may not depend on how they construct education and knowledge, but on how effectively they control their use in schools and classrooms. The development of a critical perspective that challenges socially embedded structures, such as those legislated by *No Child Left Behind*, requires critical research methods (Comstock, 1982). As such, I examine the current high-stakes accountability system through a critical lens that "strives to unmask hegemony and address oppressive forces" (Crotty, 1998, p. 12).

Critical thinking and theorizing in education are active processes and result in freedom (hooks, 1994). In other words, when researchers and participants critically engage in the educational process, they can gain an understanding of their role in the system, and, as a result, challenge the system's principles and assumptions. The purpose of a critical perspective in interpretive case study research is to "increase the awareness of social actors of the contradictory conditions of actions which are distorted or hidden by everyday understandings" (Comstock, 1982, p. 371). This study goes beyond uncovering how teachers interpret accountability messages. It seeks to trouble teachers' understandings, by developing an alternative paradigm for accountability. Rather than succumbing to the epistemological constructions and social foundations of high-stakes

accountability, I develop alternative frameworks for teachers' understandings of their own practice.

A critical perspective calls the researcher to treat society as a human construction and participants as active subjects of the construction. As such, critical methods call for a dialogue with the participants and their experiences, rather than mere observation or manipulation of subjects (Comstock, 1982). I use teachers' language and experiences to present a critical interpretation of the current manifestation of the high-stakes accountability system. Through a critical understanding of the effects of high-stakes accountability, I seek to help change the underlying presumptions and processes that currently wield power over actors in the educational system. Further, a critical perspective is based on normative dimensions. Thus, throughout this study, I draw on my own renderings of what it means to be a teacher.

The Design of the Study

Accessing teachers' understandings of their own practice requires an interpretive research approach. An interpretive case study allows me to situate teachers' experiences with the high-stakes testing policy within the complicated social context in which teachers find themselves (Crotty, 1998). I can study the process of how teachers navigate their way through a barrage of complicated messages from the multi-level accountability system, vis-à-vis the context in which they perform their work. Merriam (1998) says that in interpretive research, "Education is considered to be a process and the school is the lived experience" (p. 4). Teachers find themselves conducting their work while living in a school's accountability system. Interpretive case study methods can help illuminate this process and its meaning from the teacher's perspective in order to generate propositions

inductively about teacher experiences with accountability. Ultimately, the intention is to further consideration by actors involved in policy implementation (Yin, 2003).

Yin (2003) says that case studies have a distinct advantage for answering “how” and “why” questions about “contemporary events over which the investigator has little or no control” (p. 9). In this study, I examine science teaching in a school system that is quickly raising the accountability stakes. Beginning with the 2007-2008 school year, science testing is mandated by *No Child Left Behind*, and in the 2008-2009 school year passing the HSA (which includes biology) will be a graduation requirement. In order to prepare for these eventualities, Buckley County has enacted a new science curriculum and other professional development activities. In light of the fluid nature of the implementation of science accountability, it is difficult to predict how the policy is manifested in schools. Case study methodology can accommodate the unpredictability and fluctuation of the context in which teachers find themselves (Merriam, 1998).

Yin (1981) outlines three possible types of case studies – explanatory, descriptive, and exploratory. Each of these types of case studies is valuable under different circumstances. Because science accountability as mandated by *No Child Left Behind* is new, and Buckley County is implementing a newly developed strategy to address the legislation, this interpretive case study is in some ways exploratory in nature. In order to develop a meaningful account of teachers’ experiences, an exploratory interpretive case study must contain thick, rich description of the case. The descriptive data then can be used to develop conceptions or to illustrate, support, or challenge currently held theoretical assumptions about teachers’ response to high-stakes accountability initiatives (Merriam, 1998). In the case of science accountability, my research is likely to inform

what is already known about how teachers in other disciplines experience accountability, but it also is likely to discover nuances that are specifically endemic to science or biology education.

Thus, the purpose of data collection becomes to gather as much information about the issue as possible from a variety of sources to analyze, interpret, and theorize about the phenomenon. An interpretive case study goes beyond rich description with an inductive model of analysis. In what Stake (1995) calls an instrumental case study, the purpose of the research is to gain a theoretical understanding of something greater than the case itself. In other words, the purpose of the research is to study a phenomenon that is embedded in the case but exists in a greater context outside of it as well. In other words, how teachers experience accountability in a particular school may inform our understandings of the implementation of *No Child Left Behind* in general. The policy of science accountability is a nationwide reality. It is being implemented in most school districts across the United States. Although the level of abstraction made about the phenomenon may range from suggesting possible relationships or connections between policy mandates and outcomes, to developing a cogent theory about teachers' experiences with high-stakes accountability, the explicit purpose of an interpretive case study is analytical (Merriam, 1998). The data collected must be rich and descriptive, but also useful for generating propositions about teachers' experiences with high-stakes accountability.

Although the research is not structured around a hypothesis, an interpretive case study is based on a set of underlying assumptions that can be used to guide the inquiry. The study still has a purpose and criteria by which it can be evaluated (Yin, 2003). I

conduct this study in order to offer policy makers and implementers insights into how science educational accountability influences classroom practice. After all, the explicit purpose of science accountability is to improve teaching and learning, which ultimately occur at the classroom level. Thus, the research can be judged based upon my ability to use the data to develop cogent and meaningful insights that contribute to a greater understanding of the policy's effects.

As such, my investigation is grounded in three streams of research literature: how quality science teaching based on scientific inquiry is outlined in policy documents in order to lay a foundation for understanding what is cast as quality science instruction in science standards documents; accountability and its effects on the curriculum; and a grounding in research on how teachers process competing messages in the classroom, especially in response to high-stakes accountability policies. Teaching is a lived experience. Because each teacher's construct of good teaching is unique to his or her own experiences, beliefs, and context (Bryk & Schneider, 2002), the data collected are used to understand how teachers interpret science accountability and how they construct their practice around it. My goal in conducting the research is to analyze data inductively in order to make valuable thematic connections between the policy and its effects on teachers: their understandings and constructions of the discipline of science, teaching practices specific to science, and pedagogy in general (Merriam, 1998). Further, I use these connections to develop a set of propositions about the influence of science educational accountability on the curriculum.

The Boundaries of the Case

Stake (1995) defines case study methodology as the study of “the particularity and complexity of a single case, coming to understand its activity within important circumstances” (p. xi). Merriam (1998) distinguishes case studies from other forms of qualitative research by their focus on a “single unit or bounded system” (p. 19). Because the case itself is central to case study methodology, it needs to be explicitly and clearly defined in order to be studied in a meaningful way. I can only make valid general assertions about the policy of science accountability if I am clear on the specific context in which it is studied. In other words, if I am inducing thematic assertions about a policy based in a particular context, I must understand and define comprehensively the context in order to be able to extract general principles about the phenomenon of study from it.

A case is a bounded system. Merriam (1998) sees the “case as a thing, a single entity, a unit around which there are boundaries” (p. 27). She characterizes qualitative case studies as “particularistic” because they focus on something specific (and bounded), “descriptive” because they provide thick description of the phenomenon of study, and “heuristic” because they offer the reader a deep understanding of the phenomenon. Miles and Huberman (1994) illustrate a bounded case as a circle with a heart in the center. The heart of the circle represents the focus of the study, which gives the researcher a reference point – a representation or embodiment of the source of the research questions. The circumference of the circle delineates the boundary of the investigation. This construction underscores the potential of case studies for being too broad by encompassing so much that their focus becomes blurred. A bounded system provides the researcher with structure to begin the research endeavor with conceptual clarity and a set of tools for

investigating the area of the circle. It also keeps the length and scope of the study manageable.

In this case study, the heart of the circle is the teacher. I seek to examine how each teacher understands his or her practice. I focus on the curricular messages they receive from and about the accountability system, what contradictions they see in the messages, and how those contradictions affect their understanding of pedagogy, the discipline of science, and their passion for teaching. In essence, I seek to examine how teachers interpret accountability and how their interpretations inform their practice and the curriculum.

The surrounding area of the circle is the science accountability system. Although I focus on accountability messages at the national and state levels, I understand that policies interact with the district and school context, as well as with teachers' beliefs and constructions of quality teaching and learning. Therefore, in addition to the policy messages regarding accountability, the surrounding area of the circle encompasses the culture of Buckley County Public Schools, and its prominent role as a high achieving school district (as measured by the accountability model). The surrounding area also encompasses the school climate, including the teachers' professional relationships with other teachers and administrators, the level of the school's buy-in to the accountability system, outside professional associations that influence the teachers, teachers' background and education, and the teachers' core beliefs about quality science education.

Choosing the Buckley County school district is, in itself, an important sampling decision. In selecting Buckley County, I use Flyvbjerg's (2001) construction of a critical case. Flyvbjerg (2001) contends that a typical or representative case often may not be the

richest in information. Some cases, on the other hand, may activate more actors and processes in the situation to be studied. In addition, Flyvbjerg (2001) points out that the purpose of interpretive research may be to “clarify deeper causes behind a given problem and its consequences” rather than simply to “describe the symptoms of the problem and how often they occur” (p. 78). Using these definitions, critical cases have strategic importance in terms of revealing valuable insights about the general problem. The data gathered from a critical case can provide the greatest possible insight into the phenomenon being studied. It is this uncovering process that is at the heart of interpretive case study research (Merriam, 1998). Flyvbjerg (2001) asserts that critical cases permit the investigator to make deductions like, “If this is (not) valid for this case, then it applies to all (no) cases” (p. 79) or, at least, to a broader set of cases.

The Buckley County school district fits Flyvbjerg’s definition of a critical case. The county explicitly states that it is revamping the science curriculum to prepare for the new high-stakes testing system. The formal process that it is undergoing is well documented by them, explicitly organized, and purposeful. Clearly, the district has activated “actors” and “mechanisms” in an effort to comply with the accountability policy, and maintain its good standing as measured by the high-stakes test results. Because Buckley County is responding vigorously to state and national science accountability mandates, a study within the county is likely to uncover more nuances about the implementation of science accountability than a district that is less invested in complying with high-stakes accountability mandates. Further, as an affluent district, Buckley County Public Schools is as likely as any school district to have the proper resources to reach the benchmarks established by *No Child Left Behind*. Thus, Buckley

County fits the definition of a critical case well and offers a logical context for exploring a highly controversial testing policy. If science high-stakes accountability can be implemented properly anywhere, then it can be implemented in Buckley County Public Schools. Conversely, if a failure occurs in Buckley County Public Schools, it is likely to occur anywhere.

Another important issue to consider is the time boundaries of the case. *No Child Left Behind* mandates that “stakes” for schools and teachers are added to the standardized tests in the 2007-2008 academic year, and they go into effect for students the following year. The results of this study provide valuable insights into the effects of a high-stakes testing policy before (or very soon after) the actual “stakes” go into effect. Therefore, I conducted this study in the 2006-2007 school year, so policy implementers are able to use this research to make necessary adjustments before test results count toward graduation. In order to get an accurate, holistic understanding of teaching practice in response to high-stakes testing, I studied biology instruction prior to and through the administration of the HSA in May 2007. This timeline allows me to see the effects of the HSA unfold. I study how the impending test influences pedagogy, and what happens after the tests have been administered. Thus, the time boundaries of this study are March 2007 through June of 2007.

Sampling

When considering the question of school and teacher sampling, like all researchers, I must consider my intended audience to shape how I construct my study (Becker, 1990). The audience for my research is comprised of science curriculum and education policy researchers, as well as science accountability policy makers and

implementers – district, state, and federal personnel; administrators; curriculum writers and coordinators; and teachers. The purpose of my study is for researchers and policy implementers to gain insights into how science accountability influences teachers in order to highlight some of the policy's consequences in the science classroom.

In order for this study to be convincing, it must provide readers with vicarious experiences regarding the effects of accountability. They must gain meaningful knowledge that they can apply in their own context (Donmeyer, 1990). Therefore, in order to maximize the quality of the knowledge gained from the case study, it contains thick, rich description of both the context and the phenomenon (Merriam, 1998). A reader can then gain experiential insights akin to those he or she would acquire by being there. Consequently, although each school, district, and state context is different, translations from this study to other contexts can be developed by seeing how each teacher represents a different expression of “generic processes” that are emblematic of science high-stakes testing (Becker, 1990, p. 240). In other words, translations are not made about the case(s), but rather about the mechanisms and processes involved. Because this study focuses on the process of teaching science within a high-stakes accountability system, translating the process of how teachers perceive and negotiate through contradictions to other contexts is the explicit intent of this study. Different school, district, and state conditions may create variations in the outcomes of the teachers' negotiations, but similarities exist in the process of those negotiations as well, because accountability systems do share common principles (Linn, 2005a).

Because the unit of analysis in this study is teachers, I confine the boundaries of the study to teachers in one school. Doing so serves four interrelated purposes. First,

focusing on one school allows me to analyze the phenomenon of the study within the same social context and to interpret the science accountability policy more concisely without having to worry about differences in school context influencing and informing teacher behavior. Second, spending all of my time at one school allows me to develop a deeper understanding of the school's context and culture, which allows for a thicker and richer analysis. Not only is rich description the essence of case study methodology (Merriam, 1998; Stake, 1994), but it also allows the reader to make more meaningful extrapolations about the processes involved in the phenomenon to another context (Becker, 1990). Third, limiting the study to one school is likely to enhance participant acceptance of my study. Because I devote all of my energy to one school, the teachers and administrators become more familiar and comfortable with me which potentially affords me deeper access and a more "natural" context for study (Adler & Adler, 1994). The final purpose for limiting the study to one school is what Patton (1990) refers to as "convenience sampling." Because I am the sole researcher with limited money and time, remaining within one school allows me to maximize my resources for collecting rich contextual data.

Because I am limiting my study to one school, I must choose the "right" school. In order to choose a school, I use similar criteria that I used for choosing the school district. Patton (1990) recommends purposeful sampling in qualitative research that selects "information-rich cases" for in-depth study. He defines information-rich cases as those from "which one can learn a great deal about issues of central importance to the purpose of the research" (p. 169). Because I am investigating the process of teachers' interactions with the high-stakes testing system in Maryland, I must choose a school that

is engaged strategically with compliance efforts regarding accountability. Strategic engagement with policy implementation requires both will and resources. Therefore, I choose a school with capable leadership (as defined by the ability to influence teachers) that believes in accountability, or is driven to implement it for fear of its consequences (Craig, 2004). Further, the school must have the resources required for proper compliance with the policy's mandates. A school with these criteria allows me to gain a greater and more nuanced understanding of the processes involved in the implementation of science high-stakes testing. A school heavily involved and invested in the implementation process fits Flybjerg's (2001) criteria of a critical case, which suggests that if something occurred (or didn't occur) at this school, it is likely to occur (or not occur) at other schools who are implementing science accountability policies.

Another important sampling decision is how many teachers to include in the study. Because teachers are the heart of my case and the unit of analysis (Guba & Lincoln, 1989), I made a sampling decision to maximize both the richness and the significance of the data to be collected. Therefore, I use Patton's (1990) notion of stratified purposeful sampling in the selection of teachers. In order to maximize the richness of the data I must limit my sample size, but I also must capture variations that would occur due to particular teacher characteristics. I consider two important teacher traits that have been shown to influence teacher responses to accountability. They are the achievement level of teachers' students (Booher-Jennings, 2005), and the teachers' curricular power: experience, status, and position (Olsen & Kirtman, 2002). For the purposes of this study, I place teachers into two power categories, novice and veteran, based on their number of years teaching and educational experiences. Novice teachers

may have less pedagogical experience and/or biology content knowledge, while veteran teachers likely have both content knowledge and pedagogical experience.

Herriot and Firestone (1983) highlight the constant tension that researchers face between including enough cases to capture the important variations, and using limited resources to explore individual intricacies effectively. In light of my goals and limitations, I focus my study on the six biology teachers at my school site. The teachers have different levels of teaching experience with both high and low achieving students. Two teachers are novices and four are veterans. (See Appendix A for the Letter of Invitation that was sent to the teachers to elicit their participation.) Studying all six biology teachers at the school allows me to maximize the richness of the data by keeping the sample small, but I still may capture important variations that increase the study's significance.

The final sampling decision in this study deals with the subject area and is, in this instance, quite straight forward. In Maryland, biology is clearly a case of critical import. Of all the science disciplines covered by the Maryland's science content standards – earth/space science, chemistry, physics, and biology – the Maryland State Department of Education only includes biology on the HSA. Because biology is the sole branch of science included in the state's science accountability policy, high school biology classes should be most affected by the new high-stakes test and the curricular restructuring accompanying it. In other words, high school biology teachers should experience the mechanisms and processes of Maryland's science accountability system more significantly than other high school science instructors. Consequently, this study focuses on biology teachers.

Data

Qualitative case study research requires various data gathering techniques. Data collection is a “series of interrelated activities aimed at gathering good information to answer emerging research questions” (Creswell, 1997, p. 110). In order to maximize the strengths and minimize the weaknesses of different techniques, case study researchers often incorporate several techniques into their research (Marshall & Rossman, 1989). According to Merriam (1998), three types of data often are used in qualitative case study methods: interviews, observations, and document review. Interview data can be characterized as “direct quotations from people about their experiences, opinions, feelings, and knowledge” (Patton, 1990, p. 10). Data obtained from interviews consist of “detailed descriptions of people’s activities, behaviors, and actions” (p. 10). Document reviews produce “excerpts, quotations, or entire passages” (p. 10) that shed light on the phenomenon of study.

The notion of data collection in the constructivist paradigm is in and of itself misleading. Data are not out there waiting to be collected. They must be noticed, considered for the purposes of research, and interpreted by the researcher (Merriam, 1989). In effect, data do not exist without the researcher, who is the “agent of new interpretation, new knowledge, [and] new illusion” (Stake, 1995, p. 56). When data are extracted or produced, they result from the interplay between the researcher and the person(s) or manuscript. Therefore, the selection of data and the techniques used in a study are determined by the researcher’s theoretical orientation, the problem and purpose of the study, and the sample selected (Merriam, 1998).

In this study, I seek to make assertions about the effects of high-stakes accountability on science teaching. Therefore, the data that I choose to collect are geared toward illuminating the processes involved in the implementation of science accountability. In other words, my data collection and analysis strategies are logically aligned with the intended purpose of the research and the theoretical foundations on which it is grounded (Yin, 2003). As a researcher, my interpretations and assertions are based on my own understandings and constructions of the phenomenon, which are a mix of personal experience, scholarship, and other research that has informed my knowledge base (Stake, 1995).

For the purposes of this study, I use all three collection techniques – interviews, participant observations, and document review – outlined by Merriam (1998), and I use Merriam's and Stake's (1995) protocols for their implementation. Before outlining the three techniques, it is important to note that the process of gathering and interpreting data is not linear. Creswell (1997) envisions the data collection process as a circle into which a researcher can enter and exit at any time. Although I start the data collection with document review, and then move on to interviews and observations, I continuously interpret the data to discover emerging trends, which I use to guide the research (Creswell, 1997). Thus, I refer to the documents repeatedly after making discoveries through interviews and observations. Also, after noticing something during an observation, I often explore it further through conversations with the teacher after class and during the second interview. (See Appendix B for a data collection timeline.)

Document Analysis

Merriam (1998) defines documents as a “wide range of written, visual, and physical material relevant to the study at hand” (p. 112). For the purposes of this study I limit the definition of documents to written materials on standards, high-stakes testing, and implementation of accountability policies. According to Merriam (1998), data extracted from documents have three advantages over data gathered from observations and interviews. First, documents usually are developed independently of the research at hand. Therefore, they are an authentic product of the context in which they are produced. Second, documents are non-reactive. That is, the interaction between the investigator and the documents does not influence the documents the way the investigator would influence human participants. In other words, although the researcher interprets the documents, they are not influenced by the research like human participants. Finally, documents are not dependent on the “whims of human beings whose cooperation is essential for collecting good data through interviews and observations” (p. 112). For the most part, they are readily available for interpretation and re-interpretation throughout the duration of the study. In this case, I explore documents to discover tensions that arise in the high-stakes accountability system in order to develop an understanding of the climate in which teachers do their work.

In order to explore the tensions (if any) that are embedded in national-, state-, and district-level signals aimed at science teaching, I review the *National Science Education Standards*, the *Benchmarks for Science Literacy*, Maryland State high school science and biology standards, the biology section of the High School Assessment (HSA) from 2001-2006, and Buckley County’s High School Biology Framework. Because most scientists

and science educators consider scientific inquiry-based instruction the most meaningful method for teaching science to all students (AAAS, 1993; NRC, 1996), and numerous studies have shown its effectiveness relative to traditional teacher-centered instruction (Haury, 1993; Kinney, 1989; Light, 1990; Lin, 1998; Lord, 2001; Schmidt, Gillen, Zollo, & Stone, 2002; Watson, 1991; Zion et al., 2005), the document review has a particular focus. My interest is in discovering and interpreting tensions between professional calls for scientific-inquiry based instruction and the content of standards, the biology curriculum aligned with them, and the state's high-stakes assessment. (See Appendix C for a list of questions that guide the document review.)

I carry out the document review prior to the onset of interviews and observations in order to have a grounding in the tensions that are present (whether teachers perceive them or not). Thus, the review generates themes and questions to guide the interviews with teachers. After conducting interviews with teachers, I return to the document examination and interpretation in order to inform the data collected through interviews and observations. Further, throughout the research process, I review documents that teachers brought to my attention. All of these documents were created by teachers and/or administrators in response to accountability mandates.

Interviews

The decision to use interviews as a method for collecting data is based on the kind of information sought by the researcher and whether interviews are the best way to access that data (Merriam, 1998). Stake (1995) considers interviews to be the “main road” to “discovering and portraying multiple views of the case” (p. 64). According to Merriam (1998), interviews are used as a data gathering strategy under three potentially

overlapping circumstances. First, interviewing is necessary when the researcher cannot “observe behavior, feelings, or how people interpret the world around them” (p. 72). Second, interviews are necessary when studying past events that cannot be replicated. Finally, Merriam suggests that interviews are the best technique to use when conducting “intensive case studies of a few selected individuals” (p. 72).

In this study, I interview teachers for all three reasons. First, I seek to gain a deep understanding of a broad phenomenon. It is impossible to observe all behaviors, feelings, and interpretations related to how teachers experience high-stakes accountability. Second, many of the teachers’ experiences may have occurred in the past. Although some of the tensions they perceive and experience may be ongoing and observable in the classroom, I cannot assume that they are all available for me to observe. Finally, this research is, in fact, “an intensive study” of six select teachers and how they think and interpret high-stakes accountability policy messages.

Interview formats range from highly structured to unstructured interviews (Merriam, 1998). In a highly structured interview, the exact words and order of the questions are predetermined prior to the onset of the interview. In qualitative research, such interviews are used only to gain socio-demographic data such as age, income, and education level because rigid adherence to predetermined questions may not allow the researcher to gain access to the *participant’s* perspectives and understandings. On the opposite end of the spectrum, unstructured interviews have no predetermined questions, only potential themes to discuss. According to Merriam (1998), these are particularly useful when the researcher does not know enough about the phenomenon of study to ask

relevant questions. In fact, one of the goals of unstructured interviews is to learn enough about the case to develop questions for subsequent interviews.

For this study, I use semi-structured interviews that lean toward the unstructured side of the spectrum. With semi-structured interviews, the largest part of the interview is guided by a set of issue-oriented questions to be explored. “This format allows [me] to respond to the situation at hand, to the emerging worldview of the respondent, and to new ideas on the topic” (Merriam, 1998, p. 74). I approach the interviews with Stake’s (1995) assumptions about qualitative research. Each interviewee is “expected to have had unique experiences, stories to tell,” so I attempt to access “descriptions of an episode, a linkage, [or] an explanation” (p. 65). Formulating questions and anticipating probes that lead to such responses is a special art. Effective questions can open up the interviewee, while focusing him or her on the research topic (Merriam, 1998). Semi-structured interviews allow me to respond to emerging data and themes. I explore issues brought up by the participants that I did not expect prior to the interview, and I use their responses to develop questions for the second interview session. As such, I use two rounds of interviews. The first occurs at the beginning of my on-site data collection, and at this time teachers were asked to sign the Informed Consent Form (see Appendix D). A second interview occurs toward the end in order to explore further what I witnessed while observing the teacher-participants.

Both Merriam (1998) and Stake (1995) recommend piloting questions and issues prior to the onset of the interview. Therefore, from the themes generated by the document review, I pilot questions with a high-school biology teacher in BCPS, who is not a participant in the study. Beyond getting some practice with interviewing, the pilot

interview allows me to see which questions and ideas are confusing and need rewording, which questions yield irrelevant and useless data, and which questions open the respondent up to exploration or shut him or her down. Additionally, the pilot interview produces important questions that I previously had not considered.

In order to get the richest possible set of data, an interviewer must have a “strong advance plan” (Stake, 1995, p. 64). In order to prepare for my interviews, I develop what Merriam (1998) calls an interview guide. My interview guide contains a few specific background questions relating to issues like educational history, teaching and/or administrative experience, and membership or affiliation with professional organizations. Most of my interview guide contains open ended-questions that can be followed up with probes and a list of issues and topics to explore. (See Appendix E for the Interview Guide I use for this study.)

The interview guide follows the logic of my research questions. The purpose of the interviews is to gain an understanding of the tensions and conflicting messages that teachers perceive in the science accountability system. These tensions may emerge from the policy documents, but they also may stem from professional development, educational background, beliefs about teaching, experiences in the classroom, relationships with colleagues, and affiliations with professional organizations. Therefore, my questions encourage teachers to explore these topics as they consider how they all conflate to inform their teaching. I also attempt to gain an understanding of how teachers might be adjusting to the science accountability system. Specifically, I focus on its effect on teaching, teachers’ understanding of the discipline of science, teachers’ understanding of their role in the classroom, and teachers’ passion for their profession.

Observations

Observations move the researcher to a greater understanding of the case (Stake, 1995). Like interviews, they are a primary source of qualitative data. They can be distinguished from interviews in two important ways. First, observations take place in a natural setting, rather than one designated for the purpose of the research. Second, rather than being filtered through an interviewee, observational data represent a firsthand encounter between the researcher and aspects of the phenomenon of interest (Merriam, 1998). Of course, the presence of the observer interrupts the “natural” character of the setting (Adler & Adler, 1994). “Researchers are rarely total participants or total observers” (Merriam, 1998, p. 102), so the “investigator effects what is being observed” (p. 103). This relationship between researcher and participants aligns with the constructivist paradigm that guides this research, which suggests that all data are the product of interaction between the researcher and the participants (Stake, 1995). Nevertheless, the researcher must be sensitive to the effects he or she is having on the situation being observed (Merriam, 1998).

In this study, I strive to achieve a level of comfort with the study’s participants to establish an “insider’s identity” (Adler & Adler, 1994). Group acceptance is likely to make the classroom experience more authentic, which increases the richness of data collected. To that end, I observe approximately fifteen class sessions for each of the six teacher-participants. I observe all six teachers concurrently, based on their teaching schedule and availability. Normally, I do not notify teachers of my observations in advance, but on a few occasions, such as field trips or extensive group projects, I pre-arrange my observations in order to assure that I will be able to be present. The specific

focus of my observations is guided by my literature review, but more importantly the observations are informed by the documents that I review, along with the teacher interviews.

What to observe is largely a function of how structured and regimented the observer wants to be (Merriam, 1998). In order to maintain an open-minded perspective that is open to exploring the effects of science accountability on teaching, I maintain a less structured format for making observations. Less structured observations are like a “television camera scanning the area. Where to begin looking depends on the research question, but where to focus and/or stop action cannot be determined ahead of time” (Merriam, 1998, p. 97). The focus of the observations must be allowed to emerge and change throughout the course of the observations. In general, my observations in this study focus on insights into teaching practices that are a response to the HSA, behaviors that are indicative of how teachers perceive their role in the classroom and within the accountability system, and indicators that characterize or run counter to scientific inquiry-based instruction. I consider how labs are conducted, how scientific writing is assigned and evaluated, how pedagogical strategies engage students in the scientific process, and any significant deviations from the standard curriculum that may illuminate the teachers’ rationale about their pedagogy.

In order to gain access to these issues, I consider the following elements outlined by Merriam (1998):

1. The physical attributes of the environment.
2. The role of the teacher in the classroom.
3. The activities that the teacher is engaged in with his or her students, and the nature of those interactions, including student responses to the teacher.
4. The content and meaning of conversations between the teacher and students.

5. Subtle factors like symbolic meanings of words, non-verbal communication cues, and the absence of things that should have happened.
6. The content and messages in the curriculum plans for the lessons. (pp. 97-98)

Each participant observation experience has its own rhythm and flow (Merriam, 1998), and using these elements to gain access to the nuances of the classroom allows me to capture some of the effects, both intended and unintended, that science accountability is having on the curriculum.

A purpose of the observations is to provide a rich context for the research.

“Researchers must put themselves into the interpretation, finding meanings that others cannot grasp” (Stake, 1995, p. 62). The observer should tell the reader a story that starts to form during the observation or emerges when the researcher reviews and interprets the data. The purpose of observations is to “develop *vicarious experiences* for the reader, to give them a sense of ‘being there’” (Stake, 1995, p. 63, emphasis in the original).

Because the physical space is fundamental to the meaning, it should be well described to capture the “uniqueness and ordinariness of the place” (p. 63). Thick, rich narrative description is the hallmark of qualitative research and allows the reader to assess how the rationale of the case applies to his or her particular circumstances (Donmeyer, 1990).

Data Collection and Interpretation

In interpretive case studies, Stake (1995) suggests methodological triangulation, or the use of multiple data sources, in order to increase confidence in the study’s findings. Using different methods of data collection properly can maximize the potential benefits of particular methods, while diminishing their drawbacks (Muijs, 2006). Further, multiple sources of data are likely to yield multiple dimensions of the phenomenon being studied (Yin, 2003). Therefore, this study does not present one “correct” perspective on science

teachers' experiences with accountability. Rather, multiple perspectives emerge from my findings depending on participants' backgrounds, teaching experiences, education, and opinions about accountability, school culture, and other socio-cultural factors that influence them. Each perspective should be and is presented on its own. However, triangulation of data also allows me to uncover deeper meanings, refine interpretations, and establish connections between data sources that may otherwise go unnoticed (Stake, 1995).

The interpretation of all the data sources – interviews, observations, and documents – adheres to the logical foundations of the study. Data analysis follows a chain of evidence based on the study's research questions (Yin, 2003), but the process is a matter of induction not reduction. "Inductive analysis...begins not with theories or hypotheses but with the data themselves, from which theoretical categories and relational propositions may be arrived at by inductive reasoning processes" (Lincoln & Guba, 1985, p. 333). This notion of interpretation calls for rigorous inspection of data as they emerge throughout the study. Data analysis and collection occurs simultaneously in order to guide the research. In other words, observations, interviews, and document review provide the researcher grounds to shift or to alter previously held interpretations (Stake, 1995).

Because data collection and interpretation are ongoing processes that could theoretically be extended forever, one inherent challenge to qualitative data collection is deciding when to stop collecting additional data (Merriam, 1998). Lincoln and Guba (1985) outline four criteria relevant to this study for determining when it is appropriate to stop data collection:

1. Exhaustion of sources – each teacher has been interviewed at least twice, each teacher has been observed for at least three weeks, all the aforementioned documents have been reviewed, and the most relevant documents that teachers referred me to have been reviewed.
2. Saturation – new data collection provides few new insights compared to the time expended to collect the data.
3. Emergence of regularities – characterized by integration of the data collected from different sources.
4. Over-extension – characterized by a sense that new data is outside the realm of the research questions, and thus inapplicable to the study. (p. 350)

Once I determine that most or all of these criteria are present, I stop pursuing further observations and searching for documents to review. However, as I continue my interpretation, on several occasions I return to teachers and documents for further clarification of the data.

In terms of interpretation, Merriam (1998) outlines three dimensions of data analysis: descriptive accounts, theme construction, and theory building. In this study, I use the first two dimensions. The first level of data presentation is description. In this case data are synthesized and organized together in a “narrative that conveys the meaning the researcher has derived from studying the phenomenon” (Merriam, 1998, p. 179). The description is produced in rich narrative form, which allows the “story to unfold from the many-sided, complex, and sometimes conflicting stories” that the participants or documents reveal (Flyvbjerg, 2006, p. 238). This form of data allows readers to make naturalistic translations to other contexts arrived at through “personal engagement in life’s affairs or by vicarious experience so well constructed that the person[s] feel as if it happened to themselves” (Stake, 1995, p.85). This way, readers from different disciplines, professions, and backgrounds can draw diverse conclusions regarding the phenomenon of study (Flyvbjerg, 2006).

The other dimension of interpretation that I use is theme construction through the “aggregation of instances until something can be said about them as a class” (Stake, 1995, p. 74). The themes are not developed according to variables or categories brought into the research; rather, they emerge from the data being gathered. The process is intuitive, but informed by the “study’s purpose, the investigator’s orientation and knowledge, and the meaning made explicit by the participants themselves” (Merriam, 1995, p. 179). The themes are developed by continuous comparisons of all the different data as they are being collected. Each theme results from compilations of heuristic data that stimulate reader interest and prompt the development of greater understandings (Lincoln & Guba, 1985). These understandings are then used to develop assumptions and propositions about the phenomenon that emerges from the data. These themes are not categories that explain some external truth. Rather, they are the outcome of my critical interpretations and efforts to make sense and meaning of the case and phenomenon of study (Stake, 1995).

The reason that I do not subscribe to the notion of theory building in this study is because theories, by definition, are non-contextualized, composed of isolated variables linked together in some way (Patton, 1990). Such a construction is evolved from the positivist paradigm that adheres to an understanding of knowledge outside of people’s experiences. In other words, the theory exists beyond the data, outside of the case itself (Merriam, 1998). By removing case study work from the “studied reality,” something essential may be lost (Flyvbjerg, 2006). The reader is robbed of the essential feature of quality case study research that allows him or her to experience the research vicariously

(Stake, 1995). Lessons learned can then be extended to other possible places and contexts.

Strength and Limitations

This study has three general strengths. First, Buckley County offers a critical case to study how science teachers experience high-stakes testing because it employs myriad mechanisms toward implementing science accountability. Thus, teachers in the county are likely to experience many of the mechanisms and processes that are involved in implementing a science accountability policy. Critical cases have strategic importance in terms of revealing invaluable insights about the general problem. The data gathered from a critical case can provide significant insights into the phenomenon being studied. As a top performing district in education and test scores, Buckley County Public Schools is highly motivated to maintain its reputation for high pass rates on standardized tests after the 2007-2008 academic year. BCPS explicitly states that it is revamping the science curriculum to prepare for the new testing system. At least on the surface, the formal process that BCPS is undergoing is purposeful in intent, well documented, explicit, and organized.¹ Clearly, the district is seeking to comply with the high-stakes accountability policies. Further, as an affluent district with adequate funding, it is likely to have the resources necessary to achieve the benchmarks established by *No Child Left Behind* (if any district can). Thus, Buckley County Public Schools offers a logical example of a case with potential for great insights into a highly controversial testing policy. As Flyvbjerg (2001) would suggest, if science high-stakes testing can be consistently implemented anywhere, then it can be implemented in Buckley County Public Schools. Conversely, if

¹ The entire process is documented online on the Buckley County Public Schools website. The fictitious name is used here to protect the identity of the county.

tensions and contradictions occur in Buckley County Public Schools, they are likely to occur in any district implementing high-stakes accountability measures.

Second, the narrow scope of the study is another one of its strengths. Case study research is designed to uncover *how* something happens, rather than *what happens*, *how many times*, or *how well* (Shulman, 1988). Because this study is conducted in a single school, it can focus deeply on the intricacies of the specific case. Rather than broadening it to several districts or even numerous schools within one district, I strive to gain a deep and nuanced understanding of the effects of science accountability on teaching. A narrow sample allows me to gain valuable insights into the school's culture, organization, and policies, all of which inform this study's data. Further, spending time at only one school allows me more intimate access by building trust with participants. Thus, due to the richness and depth of data that I seek to collect, the potential scope of the examination as a whole, which seeks to provide policy implementers with valuable insights into the practice of science teachers, is actually quite broad due to the richness of data that I can collect about the case.

Finally, I have an educational background in biology (B.S. in Biology), two years of work experience as a scientist in a physiology lab, and three years of experience teaching high school science. Thus, I have a solid understanding of science and science pedagogy, which offers me an insider's perspective on science teaching and a solid understanding of the content being taught. In other words, because this study is grounded in constructivist theories of science education based on scientific inquiry, and because I understand multiple perspectives on science instruction and the discipline of biology, I

can collect richer data and provide deeper interpretations of them than an educational researcher without a background in science.

This study also has a few limitations. The narrow, yet deep scope in which this case is bounded is a design limitation as well as a strength. A study that is limited to six teachers in one school risks missing certain key issues that inform the phenomenon in general, but that may not exist at a particular school. Or it risks overstating features of the phenomenon that may be endemic to one context, but are not inherent to the phenomenon itself (Merriam, 1998). In other words, because each school has its own socio-cultural characteristics, teachers' experiences with high-stakes testing are always situated within that context (Bryk & Schneider, 2002). Narrowing the study to one school may risk missing meaningful data and themes that would lead to a greater understanding of the phenomenon.

Related to this point, a deep, but narrow, investigation may, on the surface, hold less power for translation to other contexts, which would be quite relevant in this case. *No Child Left Behind* mandates that every state implement high-stakes science testing by the 2007-2008 school year. Therefore, many school districts and schools across the United States find themselves in similar situations to Buckley County Public Schools. The explicit purpose of this study is to provide as many policy implementers as possible with valuable insights into instruction under this policy. Thus, in order to make a reasonable case for applicability of this study's findings to other contexts, I turn to Flyvbjerg's (2001) and Becker's (1990) criteria of making translations from case study research.

Flyvbjerg (2001) suggests that individual case studies, if chosen and executed properly, in and of themselves, can be useful for developing assertions that may provide valuable insights for other cases with similar characteristics. He argues that if the internal logic of the study is strong, then it can be applied to other contexts that follow similar processes. Although each school, district, and state context is different, this study's findings can be translated to other contexts by seeing how each teacher represents a different expression of "generic processes" that are emblematic of science high-stakes testing (Becker, 1990). In other words, translations are not made about the case(s), but rather about the mechanisms and processes involved in implementing it (them). As this study focuses on the process of teaching science within a high-stakes accountability system, making translations about the process of how teachers perceive and negotiate through contradictions, not about the teachers themselves, is the explicit intent of this study. Different school, district, and state conditions may create variations in the outcomes of the teachers' negotiations, but there are likely to be similarities in the process of those negotiations.

In addition to a strong rationale, Flyvbjerg (2001) calls for thick, rich narrative text that allows the reader to assess how the logics of the case apply to particular circumstances. In order to maximize the quality of the knowledge gained from a case study, it must contain thick, rich description of both the context and the phenomenon (Merriam, 1998). A reader can then gain experiential insights akin to those he or she would acquire by being there (Stake, 1995). In other words, translations to other contexts can be made on the following criteria: the study must be strong and logically developed; the interpretation of the findings should be rigorous and insightful; and the reporting of

the case must be clear, compelling, and entirely understandable. In this case, because Buckley County is activating a variety of actors and processes to implement a successful response to Maryland's accountability mandates, some or many of them may resemble the responses in other districts and states. This study's rich description and interpretations allow a reader to gain an "awareness [that] is central to developing and passing judgment on social and political affairs" (Flyvbjerg, 2001, p. 86). Thus, through thick, rich narrative descriptions, the reader is given vicarious access into the case to determine how the findings can inform a particular context meaningfully (Stake, 1995).

Another limitation to this study is its exclusive focus on biology. The state of Maryland decided to test only biology on the HSA, leaving the other sciences out of the educational accountability system. Consequently, curricula and classrooms devoted to other science disciplines are to some extent exempt from the high-stakes pressures. Although it may be interesting to examine how exclusion from testing affects these subject areas, an investigation of them falls beyond the scope of this study. However, other states do include chemistry, physics, and earth science on their high-stakes exams. Because each subject discipline bears its own set of facts and assumptions, that may influence pedagogy in different ways, the reliability of this study's conclusions to other science disciplines is compromised by the limited scope of Maryland's high-stakes science testing program.

Summary

Accountability measures intend to significantly influence teachers. In fact, a central intention of the *No Child Left Behind Act* is to "hold...[teachers] accountable for improvements in student academic achievement" (United States Congress, 2002a, statute

1620). Therefore, teachers provide a critical perspective through which to study the policy of science accountability as mandated by *No Child Left Behind*. As such, teachers are at the center of this case study. I focus on the curricular messages they receive from and about the accountability system. I seek to understand how they negotiate these messages and how the policy influences their passion for teaching and understandings of pedagogy, the discipline of science, and the curriculum. In effect, teachers' experiences are the lens through which I seek to understand and interpret science educational accountability as mandated by *No Child Left Behind*. In the next chapter, I provide a description of the context in which this study takes place.

CHAPTER IV: THE SETTING

Introduction

Now that I have provided an overview of the study, the relevant literature, and the methodology, I present the context in which my research takes place. Halbert High School (HHS) is different from most other high schools in Buckley County, MD.¹ When I first walked into the school in February 2007, I saw a vibrant, diverse community of students in the halls. A flurry of Latino, African American, White, and Asian teen-agers meandered through the halls between periods. I walked into the main office and asked for Ms. Victoria's classroom, where I was to meet the HHS biology team for the first time. After I signed into the visitor's log, the secretary directed me upstairs toward the science department. In a way, it was unfortunate that I had come between periods because I felt very much like an outsider walking the halls that day, a thirty-one year old, graying white-guy walking through a sea of teen-agers staring at someone who did not belong. I was uncomfortable, seeking to gain access to a community with which I had yet to connect.

The science department is located towards the back of the school. I was uncomfortable in an unfamiliar setting. I had to climb stairs, turn several corners, and avoid making eye contact with the few students and teachers roaming the halls while classes were in session. After what seemed like a very long and lonely walk, I vividly remember standing outside Ms. Victoria's classroom with a bag full of sandwiches, nervous, waiting for her to finish with her students. As if she appeared solely to rescue me from my discomfort, Ms. Calypso who was walking by with Irina, her student intern,

¹In order to protect the anonymity of this study's participants, I use a pseudonym for the school in which this study takes place.

stopped and asked me, “Are you Isaak Aronson?” Her smile and welcoming demeanor immediately put me at ease. She told me to follow her into Room 216 where I would meet with the biology team.

Ms. Victoria, Ms. Clayspo, Dr. Stevens, Ms. Harris, Ms. Khana, and Ms. Lydia welcomed me with open arms.¹ As we ate our sandwiches and discussed my research, I knew that Halbert High School would be a great location for my research. We would have a partnership. I would tell the teachers’ stories, and they would have an opportunity to reflect and gain perspective on their individual and collective practice. All six teachers immediately agreed to take part in this study, and we scheduled my first set of interviews for the following week.

Since that day, I have come to feel at home in the halls and classrooms of Halbert High School. When I walk through the school, I no longer feel like an outsider. I do not have to sign in at the main office to obtain a Visitor’s Pass. I just walk right in and wave to the security guards. Students look at me with recognition. Some call out my name. A few have nicknamed me “Doc,” despite my efforts to explain that I have yet to receive my Ph.D. When I walk into the classrooms that I am observing, I am a familiar and even a welcome face. The entire science department knows my name and what I am doing there. I have developed a personal and professional relationship with members of the HHS community. Students and teachers have come to me for advice. I have been invited to a district science fair, department celebration, baby shower, going-away party, and happy hour. In fact, some of these occasions and impromptu conversations have been a source of rich data for this study. Thus, as I tell the story of the six biology teachers at

¹ I use fictitious names for the teachers in order to protect their identity.

Halbert High School, I am not telling the story of “Other.” I have become part of their story and their story has become a part of me.

Halbert High School

Halbert High School is one of the Buckley County’s most diverse schools. With a total 2006-2007 enrollment of nearly 1700 students, HHS, is 38.4% Latino, 24.3% African American, 24.1% White, and 12.9% Asian. Students with special needs makeup 13.7% of the total population, 12.6% are ESOL, and 32.1% receive Free or Reduced Priced Meals.¹ HHS is part of the Downcounty Consortium (DCC) which includes five high schools and their feeder schools. Like all DCC schools, HHS is divided into academies: Academy of Finance, Business Management and Marketing, International Baccalaureate Program Academy, Renaissance Academy, and the Academy of Visual and Performing Arts. According to BCPS, these academies were formed to “capture students’ interests, explore possible career pathways, and prepare for higher education....[because] extensive research in education demonstrates that student achievement and school climate improve dramatically when students are able to select programs that are related to their interests in the world” (DCC, 2007, p. 1). Ms. Calypso, however, puts it more bluntly and succinctly by saying that the academies were developed to monitor the students more closely. Unfortunately, she suggests that all but the International Baccalaureate (IB) Academy, are largely ignored by both the students and the faculty. Ms. Lydia hypothesizes that reform efforts geared toward the consortium of schools to which HHS belongs are indicative of the belief that the schools do not

¹ These data are from a document produced by the district, but I have excluded it from the references in order to protect the district’s identity.

perform as well as their western neighbors. She believes that this view is held by county educators as well as those outside of the system.

Although HHS may be diverse overall, individual classes within the school are much more homogeneous. Only one of the twelve classes that I observed reflects the school's demographics. All of the IB track classes that I observed are mostly White with a few Asians and African American girls. General biology and environmental science classes are populated entirely with students of color. Honors biology classes most closely reflect the school population and include some students of all ethnicities. Further, all of the ESOL and SPED students are enrolled in environmental science, a pre-biology remediation class designed for lower performing students. Ms. Calypso, who is a special education instructor in science, reports that only three students with IEPs are enrolled in biology; the rest are enrolled in matter and energy and environmental science.

The Biology Team

The biology team is extremely collegial. The word, "team," which they use to describe themselves is absolutely appropriate for the six biology teachers at HHS. I have never observed a more cohesive, warm, and welcoming cohort of teachers. Each teacher that I interviewed reports having strong ties with his or her colleagues. Ms. Lydia articulately describes the team's character while fighting back tears during our second interview, which occurs the day before she takes a leave of absence for the remainder of the school year:

This is probably the most collegial group that I have ever been in. I've been here for over eleven years. People have come and gone, but the core remains... We're comfortable with one another. We are helpful, comfortable and close... Our relationships are based on sharing, trust, and communication. Those are the values on which our department is built. That is who we are.

As I provide an overview of the biology team at HHS, both as a whole and of each teacher individually, I highlight the department's strong ties and foundation, as well as its members' desires to improve themselves and their students' learning environment. As a whole, the department wants to improve the educational opportunities for its students. All six of this study's participants see this research as an opportunity for professional development.

According to Ms. Lydia and Ms. Victoria, the two most veteran biology teachers at HHS, the biology team has been attempting to align itself with Maryland curricular standards since BCPS delivered biology curricular drafts to teachers in 1999. The biology HSA was administered first in Spring 2002, and county-wide biology semester exams were administered first in the 2002-2003 school year. These measures brought HHS and Buckley County Public Schools into compliance with state accountability mandates and *No Child Left Behind*. Further, they placed teachers at HHS under the scrutiny of the accountability system.

Since the exam's inception, the Buckley County biology semester exam scores are examined by the BCPS science department and the science department head at the school. However, according to Ms. Lydia, there "was no in-depth analysis of questions and student answers [at Halbert High School] until recently. The machines necessary to help with this analysis are not school-based, but they were brought in last Spring [2006] and this past winter" [2007]. Ms. Lydia proceeds to describe how the County Semester Exam is analyzed:

This past year, Ms. Victoria had us make data entries of student scores for each curriculum unit county assessment, hopefully, to target low performing students. During the summer, teachers will be assigned to a

weeklong review for the county exam. The county will provide a stipend for this work.

Ms. Lydia's words clearly delineate the implied value of the county semester exam. The accountability structures in biology at HHS are rather narrowly constructed.

Accountability is based on the analysis of numerical test data. According to Ms. Victoria, BCPS will broaden the monitoring system further by requiring teachers to report all of their unit exam scores to the County for analysis.

In order to prepare teachers for the high-stakes testing, BCPS also developed a strategy to ensure that teachers' professional development corresponds to the ethos of educational accountability. HHS teachers received the Teacher Professional Growth Systems (PGS) in 2000. The Science Department Head oversaw its implementation until 2004 when a full-time school-based staff member was brought in to oversee implementation for the entire school. According to Ms. Lydia, this change afforded the Science Department Head more time, but, simultaneously, diminished the impact of the PGS at HHS. According to BCPS, the goal of the PGS is to create an "environment in which teachers are afforded time, support, and opportunities for continuous growth and opportunities."¹ The growth and support, however, are geared specifically toward the mandates and philosophy of *No Child Left Behind*. The PGS seems to imply that professional development is focused almost exclusively on preparation for high-stakes testing. Professional development and support are to be fostered through "professional growth communities of shared beliefs and accountability in which standards-based

¹ This citation is from a document produced by the district, but I have excluded it from the references in order to protect the district's identity.

teaching results in consistently improved student learning.”¹ Despite subscribing to the narrow technical notion of curriculum, PGS evidently has fallen short of its goals at HHS because every teacher reports having too little time for professional growth during the school year. Further, the Halbert High School HSA Biology Improvement Plan for both the 2005-2006 and 2006-2007 school years reports that teachers do not have enough time “to meet; to analyze data; to work with students” (from the 2006-2007 HSA Biology Improvement Plan).

Teachers

The six teachers who are participants in my study reflect a variety of educational backgrounds and professional experiences. In order to set the context for who the teachers are, I describe how they came to be teaching biology at Halbert High School. I also provide some insights into their character, pedagogy, view of teaching, and relationships with students. In particular, I highlight some of their differences in order to demonstrate how each teacher offers unique insights and perspectives on the effects of high-stakes accountability on science teachers and teaching practice.

Ms. Victoria

Fate brought Ms. Victoria to teaching. After earning a Bachelor’s Degree in Animal and Veterinary Science, Ms. Victoria enrolled in Veterinary School overseas. Two years into her program, Ms. Victoria returned to the U.S. to take care of a personal matter. She tried to enroll in Veterinary School but was unable to get accepted. Professionally at a loss, Ms. Victoria decided to earn a Bachelor’s Degree in Secondary Education and a certification in biology until she could “go back to vet school or find

¹ This citation is from a document produced by the district, but I have excluded it from the references in order to protect the district’s identity.

something else to do.” It was at the end of her program that Ms. Victoria realized her calling.

I came [to HHS] for my observations. I went to Anatomy because the old Department Head was an Anatomy teacher. I went to do my observation, but he left the class, and told me to do something with [the students]. That’s when I knew that this is what I was supposed to do for the rest of my life...I could not imagine doing anything else.

Ms. Victoria’s discovery in the anatomy classroom that day was both transformative and professionally powerful. It has fomented a true passion and love for teaching. Upon graduation, she returned to Halbert High School where she has been a teacher for ten years. Ms. Victoria has continued to develop herself professionally in education. She has earned a Master’s Degree in Administration and Supervision and a second Master’s Degree in Education and Technology. Additionally, she earned a Certificate in Professional Teaching Standards for National Board Certification in biology, which she currently is pursuing.

Ms. Victoria’s professional ambition has translated into being an extraordinarily involved and dedicated teacher. She is the Halbert High School’s Head of the Science Department. She teaches IB biology, which is a two period course. After a pre-IB biology teacher switched subjects in January and the department could not find a “highly-qualified teacher,” as mandated by *No Child Left Behind*, Ms. Victoria took on three periods of pre-IB biology without any extra compensation. Ms. Victoria also is the science department’s AP Coordinator. After school, she teaches High School Plus, an after-school course for students who are deemed to be at risk of failing the HSA. In the past, Ms. Victoria has taught matter and energy, earth/space systems, physical science, environmental science, and oceanography.

Ms. Victoria is devoted to her students and is involved in extracurricular activities. She attends all school plays and many sporting events, chaperones dances and the prom, and buys t-shirts to help classes and clubs earn money. It is quite common to see her sporting an HHS shirt. Ms. Victoria also has been the varsity soccer coach, varsity basketball coach, varsity track coach, indoor track coach, summer basketball coach, senior class sponsor, and the National Honors Society Sponsor. When I ask Ms. Victoria why she does not coach anymore, she responds that the school administration does not allow her to coach and be the Science Department Head at the same time.

Her dedication and commitment result in close, informal connections with students. Students are always hanging around Ms. Victoria's classroom. Being the pre-IB and IB biology teacher, Ms. Victoria teaches Halbert High School's top students, which certainly informs her communications with students and their responses to her. Nevertheless, she has the uncommon ability to make students feel comfortable in her presence while maintaining her professional status as the teacher. She gives all students her cell phone number and encourages phone calls any time. She is available after school and during lunch. Students not only come in for help, but also to share their life stories. It is not at all uncommon for Ms. Victoria to sit with one or a group of students during lunch to listen, share, or impart advice. Her classroom is the place where students go to practice for the play, plan an event, or just spend some quiet time away from the bustle of the dining room or school hallways.

Ms. Lydia

Ms. Lydia has had two teaching careers. She graduated with a Bachelor's Degree in Science Education. Upon graduation, she taught life sciences for three years. Then she married, moved, and became a real-estate manager for fifteen years. Although she left teaching, she always maintained current certification in biology. After her divorce, Ms. Lydia decided to go back to teaching because it provided her more security and consistency. She started as a substitute teacher in Buckley County, MD and another county in the area. She decided on Buckley County because it was closer to home. Ms. Lydia came to Halbert High School because it was the first school to accept her as a full-time teacher. Originally, Ms. Lydia intended to come to HHS for a few years, but she has stayed eleven years because she likes "the diverse group of students, her colleagues, and its proximity to home." Over the years, Ms. Lydia has taught honors biology, AP Biology, pre-IB biology, environmental science, and earth science. She has chaired the science department when the full-time chairs were on leave. She currently teaches Honors biology, pre-IB biology, and environmental science, and continues to moonlight as a real-estate broker on weekends.

Like Ms. Victoria, Ms. Lydia is ambitious and passionate about professional development. After a fifteen year hiatus from teaching, she felt like she needed to update her understanding of education, so she enrolled in a Master's Program in Administration and Supervision. In addition to the degree, Ms. Lydia earned an Administrator I Certification. Soon after graduating, Ms. Lydia began a Doctoral Program in Education. She has completed her coursework and currently is working on a dissertation exploring the experiences of new science teachers in Buckley County. Since she began her doctoral studies, Ms. Lydia has developed a critical perspective of education. She says:

Unfortunately, I started to see the entire picture, which is pretty bad. It's toxic, and not just in the classroom, but also in the BCPS central office and the MTA.

In order to facilitate a better working relationship between teachers, students, and the BCPS central office, Ms. Lydia joined the Maryland Teachers Association (MTA) Board of Directors. Ms. Lydia's membership on the MTA Board has changed her perspective on her work in the classroom.

I see the struggles and the vision of the Superintendent trying to bring initiatives down to the teaching level, to the school level. Sometimes I agree, sometimes I disagree, but I can see where he is coming from. I have more tolerance for it in my own classroom. I don't have the same kind of backlash and resistance to it.

Ironically, Ms. Lydia's critical perspective has resulted in her being more complacent toward district policies. As she gains a broader perspective and takes a more active role in policymaking, she feels invested in policy mandates whether she agrees with them or not. Nevertheless, in our conversations, she is highly critical of centrally enforced policies, especially those associated with accountability. As such, she brings both a unique and critical perspective to this research.

In the classroom, Ms. Lydia likes order and becomes stressed easily. She often gets frustrated with students for talking in class or not engaging in class activities. In nearly every class that I observed, she would yell at her students at least once for non-compliance with her directions and threaten them with busywork if they continued misbehaving. She rarely follows through with her threats, but usually reiterates them two or three times during each period. Her students take advantage of her temper by acting out to manipulate Ms. Lydia's attitude and pedagogy. As such, teaching wore on Ms. Lydia, especially toward the end of the year. In fact, she took a leave of absence from

school a month before graduation in order to recover from a medical condition in a stress free environment, but she plans to return to HHS next year.

Ms. Khana

Ms. Khana wanted to be a biology teacher from the time she was a senior in high school. She received a Bachelor's and Master's degree in biology, as well as Teacher Certification in biology. She taught high school biology in a southern state for seventeen years before moving to Buckley County where she has taught for two years at two different schools. Over the years, in addition to teaching at the high school level, Ms. Khana has taught biology, anatomy, and physiology to nursing students at a community college. Although she taught in BCPS last year, this is her first year at Halbert High School. Ms. Khana currently is teaching general biology, honors biology, anatomy, and physiology. Ms. Khana is not involved in extracurricular activities. She says that she is too busy with the demands of school, and she devotes the remaining time to her family.

Ms. Khana says that she is fulfilling her professional dream. She loves being a biology teacher.

I've always wanted to be a biology teacher. I love the subject matter. I love biology. I love science. In all these years, I have never been bored with it. I am fascinated with it. I love how the human body and all living things fit together. It makes sense. It is one of the more difficult courses, but one of the more interesting ones.

Interestingly, she never mentions students in her description of why she likes teaching.

When I ask her why she chose teaching over being a practitioner of biology, she tells me that being a teacher leaves her more time to be a mother, which is her first priority. Ms. Khana's subject matter-centered approach to teaching is obvious in her practice. Her pedagogy is almost entirely teacher-centered, and she speaks very fast, describing herself

as “hyper.” Ironically, I witnessed Ms. Khana make frequent errors related to biology content in her classroom.

Although, she believes that all students are capable of learning, Ms. Khana subscribes to the belief that a teacher’s job ends once she has gone over the material.

It drives me crazy when kids are not on task and they are not doing what they are supposed to be doing. They are not getting it. They will not be able to perform the way that they should be able to perform. I don’t have kids that can’t do. I have kids who are choosing not to do. I am more convinced than ever that all children can be successful, but they have to want it. These kids don’t want it. Even my best students are just good listeners, they do not study.

Ms. Khana does not see it as her responsibility to foment the desire to study in her students. Her classroom is a reflection of her approach. Students are afraid of her. They are quiet and obedient. They often do not pay attention, but they do so silently. They are very respectful of Ms. Khana, but I never witnessed anyone come to her voluntarily for extra help or guidance.

Ms. Calypso

Ms. Calypso appears to care deeply about her students. She is a special education teacher with a background in biology. Her journey to teaching is a slow evolution toward her calling. Ms. Calypso has a B.S. in biology. Upon graduation, she had no “designs on becoming a teacher.” She worked full-time in a laboratory doing microbiology research. As a part-time job on the weekends, Ms. Calypso began working at a psychiatric hospital for children.

I worked with children ages five to eighteen. The psychiatric hospital was on a farm. It was a medication free model. We used a lot of anger management techniques. When I got in there, I loved it. I worked there for seven years. I worked my way through play therapy, social work, sociology-based, psychology-based techniques. I took extra classes to supplement what I didn’t know.

Ms. Calypso's passion for children with special needs is obvious almost immediately upon meeting her. She moved into teaching because one of her responsibilities at the psychiatric hospital was to follow the children to their schools, and she worked with their classroom teachers.

I began to see a very big need for special education services for students. Their teachers did not understand developmental models and developmental delays. They didn't know appropriate practices and interventions for students. So when I worked in the schools, I knew that this is where I want to be. I went to get a Master's Degree in Human Development and a teacher's certification in special education with a focus on working with students who have emotional disturbances.

Ms. Calypso's journey toward teaching has left a visible imprint on her approach to teaching. She loves her students and considers herself responsible for both their triumphs and failures.

In terms of content, Ms. Calypso has stayed with math and science. She is currently on her fourteenth year of teaching, having taught elementary, middle school, and high school students. She has been at Halbert High School for five years, but for the first three years, she taught only self-contained courses in math and science. After lobbying for inclusion courses, she now teaches only matter and energy in a self-contained setting. The rest of her courses are inclusion environmental science classes that she co-teaches with Ms. Harris or Dr. Stevens.

Although she considers herself to be a passionate expert in special education, Ms. Calypso fears that she is weak in her science content knowledge. She reports that being in self-contained classrooms for most of her teaching career has offered her less exposure to science (and math) content. In her inclusion courses, she has learned more biology content from her co-teacher who has subject area expertise. Interestingly, her two co-

teachers this year are certified in physics and physical education. Thus, Ms. Calypso is largely responsible for developing the environmental science curriculum by herself. Nevertheless, in order to overcome her perceived deficiency in biology expertise, like Ms. Victoria, Ms. Calypso is pursuing her National Board Certification in biology. Ironically, she officially is not teaching any biology courses. On numerous occasions she expresses self-doubt about her ability to become nationally certified. She is confident in her pedagogical skills and practices, but she considers her biology subject-matter knowledge to be weak.

Like Ms. Victoria, Ms. Calypso takes on myriad responsibilities at school. She has a student intern, Irina, who is studying to be a paraprofessional in special education. I witnessed Ms. Calypso spend a lot of time helping and guiding Irina. Additionally, although she is not supposed to be the lead teacher in her inclusion classes, because Ms. Harris has no background in science, Ms. Calypso is the lead teacher for two environmental science courses. During all of my observations of these courses, Ms. Harris mostly sat in the back or simply walked around the room while Ms. Calypso taught the class. Further, Ms. Calypso has been in charge of developing the entire environmental science curriculum because Dr. Stevens and Ms. Harris are in their first year of teaching, and Ms. Lydia segregated herself from the rest of the environmental science teachers at the beginning of the school year.

Ms. Calypso also is involved in student activities. In addition to being responsible for a cohort of emotionally disturbed students, she chaperones dances and proms, takes students to the Buckley County Science Inquiry Fair, and directs the Student of the Quarter Program. Ms. Calypso is often in a rush. She seems never to slow down. In fact,

she often has to miss classes that she co-teaches due to IEP meetings and other pressing responsibilities. Nevertheless, she made time for this research and often sought me out to talk. Her myriad experiences and background in special education offer significant insights and a unique perspective to this research.

Dr. Stevens

Dr. Stevens is a second career teacher. He says that he always wanted to be a teacher but it took him longer than most others. He retired from chiropractic practice in 1998. After seven years of traditional retirement, Dr. Stevens enrolled in a one year Master's Degree program for second career teachers. He became certified in physics. Dr. Stevens is currently in his first year of teaching. He came to Halbert High School as a physics teacher, but he was asked to teach environmental science due to a teacher shortage, so he currently teaches honors physics and environmental science.

Dr. Steven's passion is clearly physics. Whenever I asked him general questions about science or science teaching, he used examples from his physics courses. Nevertheless, he seems willing to teach wherever he is needed. He considers himself to be a novice who is ready and willing to learn from his colleagues. Everyday, he observes the other physics teacher at HHS, who Dr. Stevens credits for the development of his lesson plans. In environmental science, he turns to Ms. Calypso for guidance. Although he is the lead teacher in the class and he does run most of the classes, he credits Ms. Calypso with the formulation of lesson plans and almost all class activities. Because of his inexperience with teaching and his conservative nature, Dr. Stevens often is hesitant to offer his opinion on educational matters. Consequently, I may use fewer quotes from him than some of the other teachers that took part in this study.

Dr. Stevens is extremely patient and easy going. Being a first year teacher in a large class with many special education students, he has a very difficult time with student behavior, especially when Ms. Calypso is unable to be present. Students blatantly disregard him and often curse to and at him. Nevertheless, I never witnessed Dr. Stevens lose his temper. He keeps his cool and tries to move on with the lesson, oftentimes getting very little accomplished in the midst of uncontrolled chaos. He explains his patience with pragmatic optimism.

Why should I be frustrated? Most of my students are on the lower level of science. They ask me, “Why do I need to pay attention? What will I use this stuff for?” In the back of my mind, I am thinking some of these guys probably will never use it again....So when they get out and get a job or they want to go to college and change what they are doing. Maybe they will remember how to learn, that they can learn, and then they will be able to do whatever they want.

Despite Dr. Stevens’ patient and pragmatic approach to teaching some of Halbert High School’s most challenging students, he is an optimist. He is willing to roll with the punches and teach any courses that are assigned to him. He believes that his students can learn despite having to deal with serious behavioral problems. His optimism seems to help him love teaching, and he does not see himself “retiring again anytime soon.”

Ms. Harris

Ms. Harris is a novice teacher. She graduated in December of 2005 with a degree in Physical Education and Health, and is certified in Physical Education. Since her graduation, she has been working as a substitute teacher. This year she is a long-term substitute teaching environmental science. She co-teaches two classes with Ms. Calypso that I described above, and she teaches two ESOL environmental science courses by herself.

Ms. Harris came to teaching science completely unexpectedly, and her science teaching career is likely to be short lived.

Ms. Calypso stopped me one day when I was in the parking lot. I was here that day teaching special ed art history. She's like, "I need you." I said, "Well, I really don't know science. It's not my background." I mean, I feel uncomfortable teaching it, so I've been doing it ever since then. Of course, I took all of the books home. I teach myself...I am getting more comfortable. It depends on what the subject is, if I feel more comfortable with it. Like right now we are doing biomes. That's fine. It's plain and simple to me. But when we are going through the blue bloods and all that, science isn't really my thing.

Ms. Harris' willingness to take on a subject that is foreign to her is commendable. When she was needed, she was willing to help. Unfortunately, her inexperience with teaching and poor knowledge of science are obvious in her practice. Despite the guidance and support of Ms. Calypso, which included the opportunity to observe her as a lead teacher in two classes, Ms. Harris' pedagogy is almost exclusively limited to worksheets that students do in groups and movies that are unrelated to the subject matter, such as *Freedom Writers*. When the class reviews the worksheets together, Ms. Harris often does not know the answers to problems herself. Unsurprisingly, Ms. Harris's students are bored and generally aware of her limitations. This uncomfortable situation wears on Ms. Harris. During her second interview, Ms. Harris told me that she "cannot wait for this year to end."

Despite all of her pedagogical inadequacies, Ms. Harris cares a lot about her students and has a good rapport with them. She often jokes with students in class and seems to care about their well being. The students respond well to her overtures. They even organized a surprise birthday party with a large cake, snacks that they brought to

class, and a movie. Her passion for students is also manifested outside the classroom. Ms. Harris coaches girls' soccer, basketball, and softball at a neighboring high school.

Ms. Harris' mere presence in the science department at HHS informs this study. She is indicative of an increased demand for biology instruction that cannot be met. However, due to her inexperience and lack of educational background in science, she sees herself as an outsider in the department and offers few insights into the effects of *No Child Left Behind* on science teaching. As such, I will be using fewer quotes from her than from the other teachers who participated in this study.

Overview of Themes

Having provided an overview of the context in which this case study takes place, I proceed to develop the themes that inform the study's auxiliary research questions. This study's research questions examine accountability in science at three levels. First, I uncover the tensions embedded in policy documents at the national, state, and district level. Then, I seek to understand which tensions teachers perceive, and whether accountability structures contradict or comport with teachers' own understandings of quality teaching based on prior professional and educational experiences. Finally, I examine how teachers navigate through the tensions in their practice. The first two themes focus on uncovering tensions, while providing insights into teachers' experiences. The third theme shifts the focus toward teachers' coping mechanisms and practice within the structures of accountability.

In the next chapter, Re-defining Science Teaching, I explore how the science standards and high-stakes assessment vary in the messages they send, and how deeply some of the signals and structures have penetrated the science curriculum. In Chapter VI,

The Pressure Mounts, I examine how high-stakes accountability in science has increased the stress placed on teachers. In Chapter VII, I turn to an exploration of how teachers navigate through the different, and often contradictory, signals that emanate from the accountability system. In the final chapter, I discuss how the three themes answer the research questions, explore the implications of this study, and offer some propositions for the future of accountability in education.

CHAPTER V: RE-DEFINING SCIENCE TEACHING

Biology Reigns Supreme

High-stakes accountability has redefined the biology curriculum at Halbert High School. From the very moment that I contacted Ms. Victoria to seek permission for this study, she began to outline how profoundly the biology curriculum has changed as a result of the HSA. The changes have occurred at the county level, at the classroom level, and everywhere in between. Although I expected to see significant influences, I had no idea how profound an effect high-stakes accountability is having on biology instruction. Biology has encapsulated nearly all the science department's focus and resources because it is the only science tested on the HSA. In this chapter, I describe how biology has become the central focus of the science program in Halbert High School. Then, through an analysis of Standards' documents and published copies of the HSA, I uncover some of the mixed messages between biology standards and the HSA, and within the assessment itself. Finally, I return to the teachers' voices to delineate the consequences, both intended and unintended, of biology's presence at the pinnacle of the science curriculum.

Environmental Science Is Not Biology

The first thing that Ms. Victoria told me is that one of the changes to the biology curriculum this year, and in the foreseeable future, is that environmental science is now pre-biology for students who are not ready for biology in 10th grade. Until the 2006-2007 school year, International Baccalaureate (IB) track students took pre-IB biology in ninth grade. Other ninth graders took honors physics ("on-grade level" students) or matter and energy ("below-grade level" students). All students who did not take biology in the 9th grade took general or honors biology in the 10th grade. This sequence has been restructured because during the 2005-2006 school year, only 61.5% of all students passed

the biology section of the HSA.¹ These statistics spurred the school's administration to pressure Ms. Victoria and the science department to raise the test scores because the 2005-2006 10th graders were the last cohort of biology students who would not need to pass the HSA in order to graduate.

In order to meet administrative demands to improve HSA scores, Ms. Victoria single-handedly decided to place students who were on "grade-level or below" into environmental science in the 10th grade and biology in the 11th grade. Ms. Victoria explains her rationale as follows:

It was my idea, and this is where it could bite me in the butt, to move environmental science to 10th grade and move biology to 11th grade for on-level students...I wanted to prepare [the students]. Most of the kids coming out of matter and energy were failing biology. Just straight failing, straight E's, reading on a 4th or 5th grade level. Environmental science is supposed to teach them reading and writing in every class period, build vocabulary, teach them how to write. It's very focused on literacy. It should teach them vocabulary, content, overview general biology concepts. I am hoping that they will carry some of that with them from 10th grade to 11th grade.

Even at the structural level, the new course sequence is a profound curricular change. In most U.S. schools, biology has been taught in 10th grade for at least one hundred years. The Report of the Committee of Ten in 1892 recommended that chemistry and physics be taught in the 11th and 12th grade (NEA, 1893). At that time, biology was disaggregated into zoology, botany, physiology, and anatomy. General biology was created as a course between 1900 and 1920 (Krug, 1964). In 1920, the Committee on Reorganization of Science in Secondary Schools recommended that biology be offered in 10th grade. As pressure to offer more specialized science courses mounted during the second half of the 20th Century, more talented students began to take biology in the 9th grade, chemistry in

¹ This data are from a document produced by the district, but I have excluded it from the references in order to protect the district's identity.

10th grade, and physics in the 11th grade. As of 2005, despite some recent efforts to change this sequence, most notably Physics First, biology was taught in 9th or 10th grade in over 99% of U.S. high schools (Sheppard & Robbins, 2006). Fear of student failure on the HSA has caused Halbert High School to change a national curricular staple.

In addition to mere chronology, the content of the environmental science course has changed to pre-biology. Dr. Stevens describes the change as follows:

No Child Left Behind has changed the school because Environmental [Science] used to just be environmental science. Now it's pre-biology. So we are teaching kids who haven't done well in matter and energy, and we don't really feel that they are ready for biology yet. Now we have a course for them to get biology concepts more or less two times in a row to see if they can pick it up.

Essentially students get double biology, but miss out on an entire science class. Ms. Calypso, who was responsible for developing the environmental science curriculum for four classes and three teachers, powerfully captures the significance of double Biology:

We've incorporated this environmental science class, as a class for students, who when you look at the data, would struggle to pass the HSA. You now have a lost science for these students, and if this were my child, that would not be okay, at all. So yes, it's going to help them achieve on the test, but are they missing out on an entire field of science!?! Yes, because technically what's happening is that the biology curriculum is being taught twice. Very similar concepts, same ideas. It's not fair. It's not fair.

Ms. Calypso's words capture the duplicity of this curricular change. Students are not told that they are signing up for pre-biology. They think that they are taking environmental science, but when they come to class they get biology. The demands of accountability have caused the HHS science department to trick struggling students into taking biology twice.

This year, in particular, the environmental science curriculum is complicated because in previous years, 11th graders and 12th graders who needed another year of science after 10th grade biology took environmental science. Because the course change was not announced to students, juniors and seniors who had taken biology and the HSA already, enrolled in environmental science, thinking that they would be learning a new branch of science. Thus, the environmental science classes were composed of sophomores who were being forced to take pre-biology, and juniors and seniors who elected to take environmental science. Ms. Lydia points out the obvious tensions in having these two populations in one class:

When I teach biology concepts, students say, “I know that. We learned that last year.” They are bored. It’s just not fair. How can we do that to them?

Dr. Stevens, a physics teacher who was notified that he was teaching environmental science at the start of the school year, describes how he handles this predicament:

When I came here, I found out that some of the kids who had taken environmental science had already had biology. They’ve already taken biology and now they are looking for their third science, so they signed up for environmental science. They came in thinking that environmental science would be a little more environmental science oriented, but it’s biology of environmental science. What we are going over, most of them had already seen it before.

The issue of juniors and seniors signing up for environmental science as an elective is likely to pass next year as word gets out to students, who will no longer enroll in environmental science because they know it is a repeat of a course they have already taken. Moreover, by the 2008-2009 school year all students attending HHS will have been through the new environmental science selection process. Nevertheless, it is important to consider what this curricular change says about science at HHS.

The National Science Education Standards, which are one of two documents to which most states align their standards, suggests that students of science should “experience the richness and excitement of knowing and understanding about the natural world” (NRC, 1996, p. 13). Students who are learning the same material a second time without an additional focus on depth would find it difficult to be excited about learning. At a time when scientific literacy is considered to be paramount in the science education community (AAAS, 1993; NRC, 1996), it seems counterproductive to eliminate an entire subject from the curriculum. Further, as environmental issues have been prominently featured in the media and movies like *An Inconvenient Truth* and *Happy Feet*, it seems inopportune to eliminate the subject as a science offering at HHS, especially considering its importance for the future.

In addition to considering the implications of the new course, it is important to consider its viability. Is this change legitimate within the state’s and district’s science education system? Although the change had not been announced to students, it is not a secret. The 2006-2007 Halbert High School Biology HSA Improvement/Intervention Plan includes the following intervention:

Students who were not successful in matter and energy last year were not recommended for biology this year. Students were recommended for environmental science so the teachers could strengthen writing skills and introduce biology concepts, skills, and processes.

Although they may not know to what extent biology has overtaken the environmental science curriculum, the administration and BCPS are aware of what the HHS Science Department is doing. On what grounds besides fear of the test could the change be structurally justified by BCPS and the school administration? It seems that the environmental science course has been lost between the state and district levels.

In 2000, Maryland developed science content standards for five subject areas: earth/space systems, physics, chemistry, biology, and environmental science (MSDE, 2000b). Maryland also has Core Learning Goals for all five science subject areas (MSDE, 2004b). Interestingly, BCPS does not have any curricular guidelines for environmental science. Buckley County Public Schools developed curriculum blueprints and frameworks for earth/space systems, physics, chemistry, biology, and matter and energy. In fact, Ms. Calypso has “combed through all of Buckley County for an environmental science curriculum to no avail.” It seems that environmental science has conveniently slipped through the cracks between the state and BCPS. This omission has allowed the HHS Science Department to develop its own curriculum for environmental science, without explicitly violating any district guidelines. At the state level, however, HHS is not teaching the environmental science standards to its environmental science students. However, in the current context of the high-stakes biology exam, it is unlikely that someone from the state or district will interfere with the school’s plan.

It seems, however, that fear of the HSA has caused some teachers to disregard a course that *does* have a county curriculum. Ms. Harris who recently took over an ESOL matter and energy class said that the lesson plans that she received were for environmental science and biology, not physical science. She says:

[Their previous teacher] was actually teaching them environmental [science] and biology, so I just stuck with the plan that she told me. I’ve been doing the same thing. I make up the lessons every week. The students ask me, “Why are you doing this? How does this all flow?” Pretty much, matter and energy is out the door. She told me this is matter and energy, but teach them environmental [science] and biology. I just go along with it.

Biology has infiltrated the science curriculum at HHS beyond just environmental science. Ms. Harris describes her own tacit compliance to the previous teacher's curriculum. Further, in our conversations, she implies that Ms. Victoria is aware and in favor of what is happening. Although this change may not be a formal policy, the matter and energy example shows just how deep biology has penetrated the entire science curriculum. If biology's seizure of environmental science could be justified by the fact that no district curriculum for environmental science exists and both subjects are natural sciences, both of these explanations are invalid for matter and energy. Now, ESOL students, and perhaps others, will have at least three years of biology in a row!

Student Differences

Despite being a pre-biology course, environmental science differs from biology in several ways. First, the students in the two courses are different. ESOL students were the only group for whom the biology HSA passing rate dropped from the 2005-2006 school year to the 2006-2007 school year, and the drop was a very significant 19.8%.¹ In order to improve those figures, the 2006-2007 Halbert High School Biology HSA Improvement/Intervention Plan prescribes "not offering ESOL biology this year. ESOL students are [all placed] in ESOL environmental science where they can build their skills for the biology course next year." Therefore, no ESOL students took the biology HSA in May. Essentially, the exclusion of ESOL students from Biology bought the Science Department a one year reprieve from dropping ESOL test scores, but Ms. Victoria highlights that there will be a lot of pressure to improve the ESOL HSA passing rate for the 2007-2008 school year:

¹ This data are from a document produced by the district, but I have excluded it from the references in order to protect the district's identity.

It's going to be tough next year. The ESOL scores should go up. We are working on their literacy skills, writing skills in Environmental [Science]. We are taking two years to prepare them....Their scores need to go up, or I'm not sure what we are going to do.

I can feel the pressure in Ms. Victoria's words. She understands that she is making a two year investment into raising test scores, and she knows that the stakes are high.

Students with special needs are in a similar predicament. According to Ms.

Calypso:

This year very few kids with special needs are taking the HSA. Out of the whole group, there are only three students with IEPs this year. There will be ninety plus kids next year. It will be a major problem.

Ms. Calypso is highlighting a problem with special education that actually holds true for all students taking environmental science. The course has become a weigh station for students who are deemed to be unprepared for biology and, consequently, for the biology HSA. Thus, environmental science creates a short term HSA bubble for the 2006-2007 school year. This year's biology HSA passing rate is likely to go up significantly because all of the students at risk of failing have been removed from the pool of test takers. Ms. Victoria is hoping for over a 70% pass rate, up from 61.5%. The increase, however, will be short lived because next year all of the lower performing students will be returned into the pool of students taking the HSA. Consequently, if the new environmental science course is unable to prepare many students significantly for the HSA, HHS will see a decrease in passing rates in the 2007-2008 school year.

Further, the redefinition of environmental science has added a new layer of tracking into the biology program at HHS. Ms. Calypso describes a process that actually seems to run contrary to the ethos of *No Child Left Behind*:

The other thing that I see happening that leaves whole groups of kids behind is almost a reinstitution of tracking. While I know that it is not supposed to be happening, it is happening. It's very ironic. It is what's happening. Kids are being ability grouped because leaving no kid behind also means that you can't put a lower performing kid in with a higher performing kid if the level of instruction is going to be high. If you want everyone to meet standards, you have to find different ways to present the material. Many times that is slower.

Until this year, HHS had pre-IB biology for freshman, honors biology for sophomores, and general biology for juniors and seniors who previously failed biology. Now, there will be pre-IB biology for freshman, honors biology for sophomores, a post environmental science biology for juniors, and general biology for seniors who failed biology or the biology HSA. Therefore, rather than closing the achievement gap, a stated aim of *No Child Left Behind*, environmental science actually adds a new tracking layer of inequality.

Pedagogical Differences

Ironically, the absence of county mandated guidelines, frameworks, or curriculum for environmental science (which makes the course's adaptation to pre-biology less problematic), presents a major dilemma for environmental science teachers who are forced to develop an ad hoc course. Ms. Victoria gives the environmental science teachers two rules. First, environmental science has to introduce students to concepts that are tested on the biology HSA, including writing and literacy skills. Second, biology can not simply be a repetition of environmental science, so teachers can not use any of the available biology materials. In other words, they cannot use any materials that biology teachers already use. Dr. Stevens describes what the environmental science teachers are told to do:

It's a different format. We are not supposed to use what biology uses. We are not supposed to give them the exact same labs, the exact same tests, the exact same questions, and stuff like that. We are supposed to put biology in a different format. It's like giving kids the same thing, but in a different way in order to give kids a little more time to pick it up.

Environmental science teachers essentially have to make up a new curriculum without using any of the available resources that would be helpful to them in developing it. This problem, of course, is confounded by the fact that Dr. Stevens is a first year teacher certified in physics, and Ms. Harris is a permanent substitute teacher certified in Physical Education who has almost no background in science.

Unsurprisingly, this mandate presents a major challenge for environmental science teachers. Ms. Calypso, who is the unofficial curricular guide for Ms. Harris and Dr. Stevens, describes how she navigates the challenge of developing a biology curriculum without using any of the district's or school's materials:

Should I tell you the truth or give you the "right" answer? I decide the curriculum by: What would be interesting to the kids? What do we feel comfortable teaching? We used the ecology curriculum and expanded its time frame. We used the systemics and classification unit. We took the [biology] guides and pulled related material from them.

Ms. Calypso underscores some of the complexities associated with her assignment. She has to prepare students for biology without using the materials, an impossible mission, but she errs on the side of what she thinks is best for the students. She makes what she considers to be the moral decision. She does not draw on the exact materials, but she generally follows the biology guides. The obvious question is: Will her curriculum be different enough? Next year, will environmental science students see biology as a different course? Ms. Lydia addresses this particular tension:

I am a biology teacher. I prepare for one course, but teach two. I try to make the materials different, but the concepts are the same. The students

will recognize them next year. They will know that that they are learning the same thing twice. It's really a shame.

Teachers are being placed into an impractical situation. If the goal is to prepare students for the biology HSA, the implicit purpose in changing environmental science to pre-biology is for students to learn biology twice. Nevertheless, it seems that the department's efforts at differentiating between the two courses are secondary to improving test scores on the biology HSA. Therefore, teaching the material twice may be the default option taken when teachers are choosing between differentiating the curriculum and improving HSA scores.

Despite some serious substantive similarities, there are significant differences in pedagogy between the two courses that illuminate the influence of high-stakes testing on the curriculum. First, the HSA is not administered after environmental science, and although the environmental science students will need to take the HSA next year, its influence on the curriculum is greatly diminished. Unlike biology teachers, environmental science teachers do not feel pressured to cover *all* of the material tested on the HSA. Ms. Khana succinctly describes the difference when I ask her how often she shows a film in class:

Film is wonderful. I encourage students to be curious about the world, to watch the nature channel, documentaries. *Lorenzo's Oil* is maybe the first film I've shown all year. I never show movies. I don't have time. That's for other classes. Environmental science, matter and energy they have time.

Ms. Khana is correct. In fifteen periods of observation, I observed Ms. Harris show four movies in environmental science, and Dr. Stevens showed three in environmental science and one in physics. Besides Ms. Khana showing *Lorenzo's Oil*, I never witnessed a biology teacher show a film in class.

Showing films is a proxy indicator for two major differences between environmental science and biology. First, teachers are less pressured to “cover all of the material.” Second, as a result, they are likely to choose from a wide range of pedagogical options. Another major indicator is the county-wide science fair. Ms. Calypso, Dr. Stevens, and Ms. Nichols took forty-four environmental science students to the 3rd Annual BCPS High School Student Inquiry Conference. The teachers chose these students based on an assigned class project. Eight high-schools from BCPS participated in the conference, which included a guest speaker and an opportunity for students to present their work and interact with their peers about science. Not one biology student from HHS attended. In fact, only two biology students from the entire county participated in the conference. None of the biology teachers at HHS were surprised. As Ms. Victoria says, “In biology, we don’t have time for that.”

In terms of coverage, all three biology teachers always appear rushed and pressed for time. Ms. Lydia appears less rushed, but Ms. Victoria is very concerned about her pace because Ms. Lydia skips around and does “not cover all of the material that is on the HSA.” By her own admission, Ms. Lydia says that she is unable “to keep up with all of the material in the county curriculum,” and she does not know how the other teachers do it. She says:

It’s not fair to the students. They need to learn. We cannot throw the material at them and expect them to pick it up. I rush. I try to get through everything I am supposed to, but I have to make sure that students learn what I am trying to teach them.

Ms. Khana, on the other hand, deals with the whole notion of covering the material quite differently. She pushes right through the curriculum. Ms. Khana describes her teaching as follows:

The fact is, I am strong headed, highly motivated, and I know what I want to do and what I want to teach. When I give a test, I feel good about it. I have prepared the kids. I have done my part, but I sometimes don't feel that the kids do their part. I would love for them to meet me half way, but they are not doing it.

Ms. Khana and Ms. Lydia have very different approaches to their teaching. Their disparate approaches to their practice exemplify a pedagogical tension created by the HSA. Both teachers feel the pressure to prepare their students for the test. On the surface, it appears that the degree to which they succumb to the pressure may be dissimilar. However, it is also possible that the way they deal with the pressure is actually what is different. Ms. Khana is confident, perhaps falsely, but she believes that she fulfills her teaching obligations completely. If students fail, it is their fault. Ms. Lydia, on the other hand, seems to share the responsibility for failure with her students and buckles under that pressure. On the one hand, she cannot deal with all of the material in a way that she deems to be meaningful to her and her students. On the other hand, she feels tremendous pressure to prepare her students for a test that they will need to graduate. Perhaps it is this very powerful, and in her mind irreconcilable, tension that causes her so much stress.

Ms. Victoria also feels the pressure to cover more material. She expects her students to stay on task constantly, waste no time, and do a lot of homework. She teaches pre-IB biology, so her students are more likely to rise to her expectations of them. Nevertheless, I often feel the stress pouring out of her during my observations and our conversations. She always is rushed and pulled in many different directions. In fact, just before our second interview, Ms. Victoria confesses that she was going on a job interview in a different district where she would have fewer responsibilities and pressures. In her

words, Ms. Victoria is looking to go “where she can just teach, instead of worrying about all this other stuff.”

The pressure to cover all of the material for the HSA and the County Exam has a tremendous impact on pedagogy. All of the biology teachers, whom I observed, essentially practice a teacher-centered pedagogy. All three teachers report doing fewer labs, especially as the HSA approached. Ms. Khana’s tone is defensive as she describes how the HSA influences her labs:

The pressure is on to cover the material, but I somehow managed to fit them in. You have to know how to do them. It may not have always been a hands-on lab, like for evolution, but they did a paper lab. I did the labs, but I was more inclined to forgo labs because of the time and pressure to try and cover all of the information. You just have to be extremely on top of things.

Ms. Khana, like all the other teachers, conducts cookbook style labs where students simply follow the instructions. With nearly every lab that I observed in biology, teachers, for the most part, were unable to convey the connection between the biology content associated with the lab and the activity. Interestingly, if the biology teachers think that labs get in the way of their ability to cover all of the material needed for the HSA, why do they conduct the labs at all? It is almost as if the teachers conduct the labs because they are science teachers, and that is what science teachers do. Perhaps, most of the labs were disconnected from the lessons because the high-stakes induce teachers, at least subconsciously, to believe that they are wasting time or doing something less important than their main task of preparing students for the test. Consequently, the teachers may be acting like someone else. The high-stakes may cause a split personality between science teacher and test prep coach. It seems, especially as the HSA draws closer, the test prep coach supersedes the science teacher.

Environmental science teaching practice, on the other hand, is influenced less by the HSA. Environmental science classes include significantly more pedagogical variation in their curriculum. All three environmental science teachers rarely lecture for more than a few minutes at a time. Students do projects, worksheets, group work, games, labs, and presentations. While these activities are only rarely inquiry-based, the teachers do attempt to vary their pedagogy, and on several occasions, Ms. Calypso and Dr. Stevens endeavor to employ scientific inquiry-based instruction methods. When asked about the compatibility of *No Child Left Behind* and inquiry-based instruction, Ms. Calypso says:

It has to be [compatible]. In order for no child to be left behind, we must teach to all learning styles. If not, we leave many students behind who need experiential learning.

Ms. Calypso's response is interesting and illuminating in two ways. First, she clearly is alluding to her students, who are mainly minorities with special needs who have been shown to learn better with scientific inquiry-based instruction than traditional teaching methods (Kinney, 1989; Nesbit & Rogers, 1997; Schmidt et al., 2002). As an experienced special education teacher, she understands that her students learn differently. Therefore, they must engage with science from many different angles.

On the other hand, Ms. Calypso's students will not be taking the HSA at the end of this year. She has the luxury to talk about what *No Child Left Behind* should be rather than what it is. The biology teachers may agree with her in principle, but they argue that they do not have time to present the material through scientific inquiry-based instruction. For the time being, however, the biology teachers do not have to worry about students with special needs, who have all been placed in environmental science. Perhaps next year

they will be more compelled to diversify their pedagogy in order to engage a more diverse population of students in biology.

Nevertheless, it is very likely that biology teachers will continue to experience a tension between quality teaching and the demands of the HSA. When describing her pedagogy, Ms. Victoria highlights a paradox between how *No Child Left Behind* should influence teaching versus how it does influence teaching:

Good teaching engages students and ensures that they “get” what the teacher tells them. It presents students the material in different ways. But I don’t have time for that. Lecture gives teachers an opportunity to tell students what they need to know in a short amount of time.

Ms. Victoria seems to know what she thinks is a better way to teach, but she cannot teach that way because she needs to tell her students “what they need to know” for the test. But teaching is not telling. Ms. Victoria may be choosing her role as a test preparation guru, over her role as a teacher. It appears that the other biology teachers are pressured to follow suite. On the other hand, perhaps environmental science teachers, who are not faced with a high-stakes exam at the end of their course, are less likely to feel like they have to make this choice.

The Changing Curricular Landscape

Another way of looking at the effect of high-stakes testing on pedagogy is through the perspective of curriculum. *No Child Left Behind* aligns with Ralph Tyler’s (1949) technical paradigm of curriculum-as-plan, where teachers implement, with as little interference as possible, what the state outlines as knowledge all students at a particular grade level must possess. In order to mandate compliance with the technical notion of curriculum, students are administered the HSA at the end of the year to determine whether teachers passed on the planned curriculum to their students. Ms Calypso

illuminates some of the challenges with *No Child Left Behind*'s input/output paradigm of teaching and learning:

NCLB presumes that it can test teachers through students. We give students the HSA and assume that their scores reflect teaching, but it's not fair. There is so much else going on in the classroom and at home. We label the teachers according to the scores, but all students are different.

Although *No Child Left Behind*'s accountability logic operates on the false premise that teachers are sieves through which standards can simply pass to students, severe stakes are attached to students' success on the HSA, and teachers feel responsible for their students.

Ms. Lydia describes her greatest concern about the HSA:

It bothers me. I know some of my students will fail. Some of these kids have more issues than any of us. They need support. Instead, some of my students will never get their diploma. They have never faced this. They will be the first students to finish school without getting their high school diploma. How does that look? It bothers me much more than I thought it would.

Teachers hold themselves accountable for their students, and they are concerned about their students' futures, which is likely a reason why the high-stakes environment significantly influences the biology curriculum.

An Emphasis on Inquiry

Most scientists and science educators consider scientific inquiry-based instruction the most meaningful method for teaching science. Research, as follows, has demonstrated that scientific inquiry-based instruction improves science learning, compared with traditional teacher-centered instruction. Through scientific inquiry, students gain a deeper understanding of scientific content and processes (Branford & Donovan, 2005; Schmidt, Gillen, Zollo, & Stone, 2002; Zion, Michalsky, & Maverich, 2005). Consequently, both

national and Maryland state science standards highlight the importance of inquiry-based instruction.

Since inquiry-based instruction subscribes to the constructivist notion of learning, it is not about giving students the “right” answer, but about developing ways for them to discover their own answer much like scientists do in their practice. According to the *Benchmarks for Science Literacy*, “The teacher’s job is not to provide students with the ‘right’ answers...but to see to it that students know what questions to ask” (AAAS, 1993). In other words, the national standards

emphasize a new way of teaching and learning about science that reflects how science itself is done, emphasizing inquiry as a way of achieving knowledge and understanding about the world. They also invoke changes in what students are taught, in how their performance is assessed, in how teachers are educated and keep pace, and in the relationship between schools and the rest of the community – including the nation’s scientists and engineers. The *Standards* make acquiring scientific knowledge, understanding, and abilities a central aspect of education, just as science has become a central aspect of our society. (NRC, 1996, p. ix)

Scientific inquiry-based instruction, utilizes student differences to help students learn from one another, which is precisely why the *Benchmarks for Science Literacy* emphasize that common goals do not require a common curriculum (AAAS, 1993).

The national standards have high expectations of students and suggest that all students can achieve the knowledge and skills outlined in the standards. However, “Different students will achieve understanding in different ways, and different students will achieve different degrees of depth and breadth of understanding depending on interest, ability, and context” (NRC, 1996, p. 2). Thus, the goal of science instruction is to teach students to

use scientific principles and processes in making personal decisions and to participate in discussions of scientific issues that affect society. A sound

grounding in science strengthens many of the skills that people use every day, like solving problems creatively, thinking critically, working cooperatively in teams, using technology effectively, and valuing life-long learning. (NRC, 1996, p. ix)

Although the national standards may be idealistic, they do stress the importance of inquiry and pedagogical differentiation for optimal success in science teaching and learning.

Although they are not a compilation of teaching strategies, the national science standards intended to provide “educators in every state and school district with a powerful tool to use in fashioning their own curricula” (AAAS, 1993, p. vii). At the time of their creation, it may not have been clear whether states and districts would use the national standards. Most states did indeed craft their standards according to the two national documents (Leonard & Penick, 2005; Marx & Harris, 2006). Specifically, Maryland used the *National Science Education Standards* and the *Benchmarks for Science Literacy* to develop its own Science Content Standards in 2000. The first paragraph of the document reads:

In 1992, the science community was challenged to create a vision of the scientifically literate person and standards for science education that, when established would allow the vision to become reality (National Research Council, 1995). The Maryland Team has accepted this challenge. The document that they produced reflects *their* vision of science education in the twenty-first century (emphasis in the original). This vision is based on the national reform movement, the work of the National Science Teacher’s Association, Project 2061, and the National Research Council. (MSDE, 2000b, p. 1)

Like the national standards, Maryland Science Content Standards emphasize scientific inquiry-based instruction. MSDE underscores that its science

standards communicate a definition of science that goes beyond that of a course of study or the name of a textbook. Science is a body of knowledge developed through the process of investigating that is combined with

thoughtful reflections guided by critical thinking skills. This body of knowledge is dynamic and has a dramatic impact on every aspect of social life. (MSDE, 2000b, p.1)

The Maryland State Science Core Learning Goals put it more bluntly by simply stating, “All students should participate in an inquiry-based science program” (MSDE, 1999, p. 8). Maryland could not possibly be any more explicit in its endorsement of scientific inquiry-based instruction. In order to be clear about the meaning of scientific inquiry, Maryland quotes the *National Science Education Standards*’ definition. It, then, goes on to say:

As an outcome of inquiry-based learning, students will recognize that science is more than a body of knowledge. It is also a way of thinking and a way of investigating. (MSDE, 1999, p. 2)

Maryland wants to foster science education that encourages students to think critically, ask questions, and do science. In other words, MSDE wants students to construct their own understanding of science, which certainly operates on the premise that all students learn differently. Ms. Calypso highlights students’ different learning styles, using the lexicon of *No Child Left Behind*:

In order for no child to be left behind, we must teach to all learning styles. If not, we leave many students behind who need experiential learning.

It appears that both national and state science teaching policy documents are in agreement with Ms. Calypso.

Buckley County Public Schools also concurs with the nation and state. Like the nation and state, BCPS highlights the importance of scientific inquiry-based instruction and the myriad benefits to students that accompany it. BCPS often quotes the national and state standards in its science policy documents. Right on its website’s science homepage, BCPS quotes the *National Science Education Standards*:

Science is a way of knowing that is characterized by empirical criteria, logical argument and skeptical review. Students should develop an understanding of what science is, what science is not, what science can and cannot do, and how science contributes to culture. (NRC, 1996, p. 21)

BCPS considers itself a leader in innovative science instruction. According to the BCPS Science Teacher's Handbook, the goal of a science program is "for all students to develop and utilize curiosity and achieve scientific literacy by...engaging rigorously in the process skills of science through real-world applications."¹

The handbook goes on to explain to teachers what kind of instruction helps students achieve scientific literacy:

Science instruction recognizes the natural wonder that students bring to the classroom and weaves this curiosity into learning experiences. Students should do science, not hear or read about it. Science is best experienced through open-ended, hands-on inquiry that promotes student-generated questions that result in children thinking and acting like scientists.²

Specifically, the Handbook prescribes that teachers use the 5 Es Inquiry Learning Cycle, which is a "method of structuring a science lesson that is based upon the constructivist learning theory, research-based best practices in science pedagogy, and cognitive psychology."³ The 5 Es are engage, explore, explain, extend, and evaluate. According to the BCPS science website, a variety of teaching strategies are used, "including laboratory experiments, demonstrations, direct instruction, current events, visual presentations and cooperative learning, as appropriate to the lesson and prior knowledge of the students."

According to the Handbook, a science lesson should take several days or a week to

¹ This citation is from a document produced by the district, but I have excluded it from the references in order to protect the district's identity.

² This citation is from curriculum modules produced by the district, but I have excluded them from the references in order to protect the district's identity.

³ This citation is from curriculum modules produced by the district, but I have excluded them from the references in order to protect the district's identity.

complete. BCPS seems to understand that all students learn differently and, as specified in the *Benchmarks for Science Literacy*, that common goals do not necessarily require a common curriculum. In fact, using reverse reasoning, a common curriculum is actually unlikely to result in the attainment of common goals because all students learn differently.

In addition to promoting scientific inquiry-based instruction in their policy documents, the actual standards and curriculum frameworks that Maryland and BCPS have developed respectively contain many science skills and process indicators. Maryland has developed an entire standard for scientific skills and process (in addition to the biology, physics, chemistry, earth/space systems, and environmental science standards). Although the process standard is separate from the content standards, Maryland explicitly states that the content standards should be taught through the skills and process standard, which will help students develop “a richer, deeper understanding of the facts and principles” (MSDE, 2000b, p.1). The skills and process indicators include inquiry-based instruction components like formulating “questions that lead to a testable hypothesis, which demonstrates the logical connections between scientific concepts and the design of an investigation” (p. 3), and recognizing that “real problems have more than one solution and decisions to accept one solution over another are made on the basis of many issues” (p. 13).

Like MSDE, BCPS has skills and processes as the first part of its Biology Scope and Sequence (Townsend, Lynch, Gagnon, Kraemer, Szesze, & Morse, 2003). Also like MSDE, BCPS considers biology content, skills, and processes to be pedagogically inextricable. According to the biology curriculum revision team, “All concept

indicators...are addressed by using [skills and process] indicators as the primary vehicles of instruction” (p. 3). At BCPS, skills and process indicators address such concepts as recognizing “that real problems have more than one solution and decisions to accept one over another are made on the basis of many issues” (p. 3, aligned with MSDE Standards, 2000, p. 13), and that “connections exist both within the various fields of science and among science and other disciplines including mathematics, social studies, language arts, fine arts, and technology” (p. 5).

The curriculum modules that are used across the district and, at least to some extent, by all three biology teachers at HHS, follow this biology framework. They align with content, skills, and process goals. They follow the 5 Es Inquiry Learning Cycle. They include such activities as a pre-assessment based on constructivist principles. For example, for the pre-assessment on evolution, students receive a picture of a giraffe and another picture of two horses. The caption reads, “It is believed that the giraffe evolved from the horse. Describe how you think this could happen.”¹ In this activity, students are supposed to consider the process of evolution based on prior knowledge and understanding. Another example of an activity that requires critical thinking skills and other analytical processes has students compare transport in plants and in animals using a Venn diagram. Students are given a word bank and told to “place the terms on the proper space on the graphic organizer. Then add additional items.”² Although this activity is more prescriptive than the pre-assessment on evolution, it requires organizational skills and the ability to compare across the Animal and Plant Kingdoms. It has an open ended

¹ This citation is from curriculum modules produced by the district, but I have excluded them from the references in order to protect the district’s identity.

² This citation is from curriculum modules produced by the district, but I have excluded them from the references in order to protect the district’s identity.

component which asks students to add their own items, which calls them to make associations between the lesson and prior knowledge. In summary, although the biology curriculum guides may not be a perfect facilitator of scientific inquiry-based instruction, including some prescriptive labs and memorization-based activities, they do offer opportunities for inquiry and a variety of pedagogical options for teachers.

As I have demonstrated in this section, BCPS, on the surface, seems to follow the national and state recommendations for scientific inquiry-based instruction. A call for this type of instruction is explicit in written policy documents, and it is present as skills and process indicators in state standards and district frameworks. Further, these standards do appear in the BCPS biology curriculum guides. However, despite the rhetorical presence of inquiry throughout the written science and biology curriculum in Maryland and BCPS, I observed almost no inquiry-based instruction in the classroom. In fact, all six teachers on the biology team at HHS suggest that scientific inquiry-based instruction is incompatible with the mandates of *No Child Left Behind*. Dr. Stevens describes why:

Classes that have a state exam at the end of the year are taught where the [teachers] know what they have to get through and if there is too much material they have to get through it anyhow. That may be an instance where inquiry gets cut. Where [teachers] can't put the inquiry in because they feel like there is so much they have to get in. They think that if we don't do inquiry, which takes more time, we have more time to get through everything.

The disparity between the stated aims of national, state, and district documents and the lived curriculum at the school level is profound. Thus, it is important to examine how *No Child Left Behind* excludes the principles of scientific inquiry-based instruction from the biology curriculum.

Mixed Messages

Science content standards may be an epistemologically incorrect promulgator of the constructivist approach to learning. By definition, standards are objective in that they are developed and exist outside the classroom experience. Educational standards align with a positivist epistemology in that they have a pre-determined truth and meaning that must be discovered in the classroom (Crotty, 1998). By its own admission, BCPS wants to align its science curriculum with the constructivist learning paradigm, where students would create truth and meaning through engagement in the biology classroom.¹ Because standards and inquiry seem to subscribe to opposing epistemological constructions, the development of science standards may not be a reasonable method for promoting scientific inquiry-based instruction in the classroom. In other words, standards and inquiry seem paradigmatically incongruent.

However, the main reason for the absence of inquiry in biology classrooms at Halbert High School is more straightforward than an epistemological contradiction. As Dr. Stevens points out when he explains why *No Child Left Behind* is incompatible with inquiry-based instruction, the high-stakes summative test, not the science standards, drives the curriculum. In our conversations, all six biology teachers at Halbert High School discuss the effect of the HSA on the curriculum. None of them state that they are influenced by policy documents or standards. Teachers want their students to graduate. They want their own reputation and that of their school to be solid. Currently, both of these are measured by students' success on the HSA. Therefore, when examining

¹ This information is taken from curriculum modules produced by the district, but I have excluded them from the references in order to protect the district's identity.

pedagogy, it is important to consider the influence of the HSA, in addition to the standards.

The biology HSA is a 2.5 hour test that includes fifty-three questions. Despite evidence that multiple measures are needed to evaluate the depth and breadth of science learning (Herman, 1997), the questions come in two types: multiple choice and brief constructed response (BCR). During my document review, I analyzed all five publicly released biology HSAs from 2002-2006. The breakdown of multiple choice questions and BCRs is as follows:

Types of Questions that Appear on the Biology High School Assessment

Year	Multiple Choice	BCR	Total
2002	47	6	53
2003	46	7	53
2004	46	7	53
2005	46	7	52
2006	47	6	53

It is unclear why 2002 and 2006 only have six BCRs. The tests for those years do not appear to be longer than the other years.

In terms of HSA content, MSDE publishes a chart of topics that are covered on the HSA.¹ In my content analysis, I use the categories established by the state, despite their imperfect alignment with the MSDE Standards and Core Learning Goals. In order to inform the relationship between the test and the BCPS biology curriculum, I categorize the questions to align with the indicators in the BCPS biology framework (Townsend et al., 2003). Below is a list of the six categories developed by the state in which I include subheadings based on the biology framework.

¹Maryland State Department of Education website at <http://hsaexam.org/img/CD%20Content/HSA%20Topic%20Areas.pdf>.

1. Skills and Processes of Biology
2. Structure and Function of Biological Molecules (Biochemistry)
 - a. Molecular Properties
3. Structure and Function of Cells and Organisms
 - a. Systems of the Body
 - b. Cell Structure and Function
 - c. Plants
4. Inheritance of Traits
 - a. Genetics
5. Mechanisms of Evolutionary Change
 - a. Evolution
6. Interdependence of Organisms in the Biosphere
 - a. Relationships between Organisms
 - b. Environment

Not all questions fit into completely discreet categories, which is another reason that I chose to go with the broader categories developed by MSDE. Thus, when assigning a particular question to a category, I chose where its main focus seems to fit. For the most part, each question is aligned with a particular category. In a very few instances where a question seemed to span more than one topic, I assigned the question to more than one category.

The content data for each year are compiled in the table below:

Biology Categories Tested on the High School Assessment

Category	2002	2003	2004	2005	2006	Total	Percent
Skills & Processes of Biology	10	10	12	12	9	53	19.6%
Structure & Function of Biological Molecules	4	4	4	3	4	19	7.0%
Structure & Function of Cells and Organisms	12	11	10	11	12	56	20.7%
Inheritance of Traits	13	10	11	11	13	58	21.5%
Mechanisms of Evolutionary Change	5	4	4	2	6	21	7.8%
Interdependence of Organisms in the Biosphere	10	14	12	15	12	63	23.3%

Clearly, not all categories are tested equally, but the consistency with which each category is tested is remarkable, if not surprising. For example, Biochemistry does not get tested frequently. However, the HSA includes four questions on Biochemistry every year except 2005, when only three questions on Biochemistry appeared. Structure and Function of Cells and Organisms, Inheritance of Traits, and Interdependence of Organisms in the Biosphere are the most popular categories. Within each topic, however, not all subcategories were tested equally. Within Structure and Function of Cells and Organisms, Cell Structure and Function was tested much more frequently than Systems of the Body and Plants. In terms of Interdependence of Organisms in the Biosphere, only a couple questions on each HSA were devoted to Environment, while the remaining ones pertained to Relationships between Organisms. All of the biology content questions on the HSA correspond to state and district indicators. However, only approximately half of the content indicators were tested on the HSA. Therefore, when teachers make curricular decisions based on the high-stakes test, they are likely to focus on some indicators over others based on the frequency of their appearance on the HSA.

Further, the depth of knowledge consistency between standards and the HSA is extremely poor. The Maryland Science Content Standards (MSDE, 2000b) calls students to “develop a richer, deeper understanding of facts and principles” (p. 1). Maryland’s one content goal for biology is for a student to

...demonstrate the ability to use scientific skills and processes...and major biological concepts to explain the uniqueness and interdependence of living organisms, their interactions with the environment, and the continuation of life on earth. (MSDE, 2002b, p. 7)

Similarly, in its biology framework, BCPS mandates teachers to “set high standards with challenging and rigorous expectations for all students” (Townsend et al., 2003, p. 2). Its biology content goal is

for students to understand the complexity of the living world, including the functions and processes of organisms, their interactions with one another and the environment, and to respect the living world. (p. 1)

Despite state and district standards that prescribe in-depth knowledge of complex relationships and systems, few questions on the HSA demand more than rote memorization of unrelated concepts.

The HSA does not cover the range of knowledge that is covered in the Maryland biology standards. The test covers a relatively broad range of concepts, especially as compared to the depth of knowledge, but not nearly as extensive as the collection of concept indicators in the biology framework and biology core learning goals. A fifty-three question exam cannot possibly assess over 150 content indicators, which would be okay if HSA test questions were evenly distributed across indicators, making each one as likely to be tested as any other. My review, however, has demonstrated otherwise. Genetics, ecology, and cells are, by far, the most tested topics on the HSA.

Thus, teachers who are pressed for time and driven to prepare their students for the high-stakes exam, can and do focus on the most commonly tested topics. In fact, the biology team at HHS has reduced the relevant biology content into a seven page HSA Review for biology that they give to all their students as a study guide for the exam. As she hands out the review sheet to her students, Ms. Lydia says:

This review sheet has everything you need to know for the HSA. Take it home. Study it. Know it, and you'll pass.

As expected, the review sheet devotes one and a half pages to cell structure and function, two pages to genetics, and two pages to cells. The remaining one and a half pages are split between biochemistry, animal organ systems, and taxonomy. The biology teachers are well aware of which topics most frequently appear on the HSA, and they emphasize them in the curriculum.

Disposing of Scientific Inquiry

In terms of evaluating scientific inquiry, the HSA is lacking. Despite the state's and district's emphasis on scientific inquiry in all of its policy documents, the HSA is nearly void of meaningful scientific inquiry skills. Although Skills and Process questions make up nearly 20% of the HSA, most of them deal with reading simple tables and charts or basic lab procedures that require no critical thinking or analysis. The following is a typical skills and process question from the 2005 biology HSA that deals with reading a chart:

3. The table below shows the number of species of different types of simple land plants. (p. 3)

NUMBER OF SIMPLE PLANT SPECIES

Simple Plants	Number of Species
Brophytes	20,000
Club Mosses, spike mosses, and horse tails	1,000
Ferns	12,000
Total	33,000

According to the table, approximately what proportion of all simple plant species are brophytes?

- A. $\frac{1}{3}$
- B. $\frac{1}{2}$
- C. $\frac{2}{3}$
- D. $\frac{3}{4}$

The indicator that this question tests is “1.6.1: The student will use ratios and proportions in appropriate situations to solve problems” (MSDE, 2002b, p. 5). Although question three may require a skill that would be necessary in biology, it certainly does not examine scientific inquiry because the information in the question is not integrated into a scientific process. The Maryland Science Standards state that skills, process, facts, and principles “cannot be taught in isolation of one another” (p.1). That is exactly what is happening in this question. The rudimentary chart in reading and math skills required for the question are not connected to the material in the chart. A student does not need to know anything about simple plant species to answer this question. In fact, a student does not even need to know the meaning of the term “plant” or “species.” A scientist would never be in a situation where she was extracting and manipulating meaningless information from a chart. This question both devalues the facts and isolates scientific skills from the context in which they would be performed.

Similarly, questions on scientific investigations that appear on the HSA do not require students to engage with an activity conceptually. They ask students to regurgitate lab procedures without engaging in a scientific process. The following is a typical lab procedure question from the 2005 biology HSA:

8. A scientist is performing an investigation funded by a company. Which of these would be least likely to produce biased data? (p. 6)

- F. making the results please the company paying for the research
- G. being open minded and honest throughout the research project
- H. using only data that supports the hypothesis
- I. using personal opinions to decide the results of the research

The indicator that this question tests is “1.1.5: The student will explain factors that produce biased data” (MSDE, 2002b, p. 3). Besides being a stretched attempt to test

ethics in conducting scientific research, this question is disconnected from scientific reality. Kulm, Dager-Wilson, and Kitchen (2005) reveal that student thinking in answering test questions often does not correspond to the intended standard. Their findings are likely to apply in this case. At least question three mentions plants. This question does not refer at all to what research is being conducted. Further, the indicator requires students to *explain* bias. In this question, students do not need to explain anything. They simply are asked to *choose* a prescribed answer. In other words, the question does not place ethics into a real context for the student, and it certainly does not require any inquiry or critical reasoning from the student.

Questions three and eight are typical multiple choice questions on scientific skills and processes that are intended to test scientific inquiry. Instead, they ask students to regurgitate isolated skills and process knowledge that are disconnected from any reasonable scientific reality. Other skills and process questions require students to read graphs, convert units, identify independent and dependent variables, and denote safe laboratory procedures. These types of questions do not require students to analyze, investigate, compare and contrast, evaluate, hypothesize, critique, understand, or do any of the other inquiry skills that are espoused by national standards, state standards, and the county's curriculum guides. The mere fact that each question is aligned with a single indicator suggests that biology knowledge is not integrated, and students are not evaluated on a comprehensive understanding of biological principles.

When preparing her students for the HSA, Ms. Khana says:

This question asks you to do math. These kinds of questions are uncommon on the HSA. They never want you to do two things. They want to see if you know the biology concept. It would be unfair if you got a problem wrong because you didn't know the math.

Although it may be unfair, this approach is antithetical to national and state calls for an integrated and interdisciplinary approach to teaching science (AAAS, 1993; MSDE, 2000b; NRC, 1996). In order for students to understand science, instruction should be comprehensive and informed by technology, math, reading, and history (NRC, 1996). Such an approach to teaching science not only is a bedrock principle of scientific inquiry-based instruction, but also it has been shown to improve student learning in both science and other disciplines (Harwood, 2002; Nesbit & Rogers, 1997; Ryan & Walking-Woman, 2000). Students with special needs, especially, benefit from an interdisciplinary approach to science instruction (Schmidt et al., 2002). Testing students on individual indicators promotes teaching and learning that reduce science into discreet bits of knowledge, rather than a “way of knowing that is characterized by empirical criteria, logical argument and skeptical review” (AAAS, 1993, p. 21).

Perhaps multiple choice questions are, by their very nature, inappropriate for evaluating students’ understanding of the scientific inquiry process. All six teachers much prefer BCRs to multiple choice questions as a more meaningful evaluation method. On the surface, BCRs appear to ask students to analyze and synthesize information. Occasionally, they even ask students to develop a set of procedures or a hypothesis. Because each HSA only has six or seven BCRs, such questions are very much the minority on the exam. Nevertheless, it is important to consider them in this analysis because they are a regular part of the curriculum.¹

¹ A single BCR comes with several parts. Thus, it often takes more time than a single multiple choice question. Some multiple choice questions, however, require students to read a passage, which may occupy students as long as a BCR.

The following question from the 2006 HSA is an example of a BCR that attempts to evaluate a student's understanding of a scientific investigation:

49. Students studied a species of fish. They wanted to find out if these fish grow faster in warm water. The students designed an experiment to determine how different water temperatures affect the growth of the fish.

They place one fish in a tank at 26°C and another fish in a tank at 22°C. The fish were fed the same amount of food during the experiment. The mass of each fish was recorded at the beginning and end of each experiment. The data that the students collected are shown in the table below.

EFFECTS OF WATER TEMPERATURE ON THE GROWTH OF FISH

Temperature (°C)	Starting Mass (g)	Final Mass (g)	Percent Change
26	3.68	7.84	113%
22	6.8	9.09	34%

Analyze the procedure and the data from the experiment. In your response be sure to:

- include the hypothesis that the students were most likely investigating
- explain whether their data supports this hypothesis
- describe how other variables would affect the outcome of the results
- explain how the experiment could be designed to gather more reliable data

Write your answers in your Answer Book. (p. 35)

On the surface, this question seems to evaluate some of the processes of scientific investigations, like making predictions, connecting data to hypotheses, analyzing variables, and validity and reliability. Below the surface of the BCR, however, lies a flaw that may contradict the principles of scientific inquiry, skills, and processes. Scientific inquiry is based on constructivism, where students or scientists make their own meaning with what they observe and create (MSDE, 2004c; NRC, 1996). Meanings depend on science students and the context in which they are doing their work. Therefore, scientific

endeavors can take multiple paths; problems have more than one correct answer; and observations can have many explanations.

BCRs, however, create little context for the student. Students are graded using a rubric by an outsider who does not know the student and grades hundreds of BCRs. When test questions are written, the author(s) develop a specific rubric for evaluating answers based on a total of four points. Consequently, BCRs nearly always are accompanied by four bullet points for students to cover in order to earn maximum credit. In effect, the scoring rubrics, and the bullet points that accompany them, reduce an open ended question where the student can truly think critically and creatively, to one where the student needs to regurgitate the *right* answer. In number forty-nine, the question, “Analyze the procedure and the data from the experiment,” can send the student into multiple directions in no particular order. The author(s) of the question, however, intended for students to come up with a particular response which they would lay out in a scoring rubric. Thus, when responding to the question, students must make the author(s)’ meaning instead of their own if they want to receive credit and graduate from high-school.

Further, the bullet points reduce knowledge that should be integrated into four separate knowledge bits. Ms. Calypso describes how she learned about BCRs:

I first heard [about BCRs] in BCPS, 5 years ago. I was told that a well written BCR had to have structure, format, a topic sentence, a closing sentence, and three sentences in between. Then recently, I was told that a science BCR is different. You can use graphs, data, tables, and drawings. BCRs follow a prescriptive formula. Two years ago I was told that students need to use five vocabulary words related to the prompt in order to receive, at least, partial credit.

Dr. Stevens ironically describes how BCRs are graded:

A BCR can be open-ended, but there are certain ways you have to grade it. There is a certain way you have to look at it. There are certain things that they are looking for.

It is evident that Ms. Calypso and Dr. Stevens are not describing something that evaluates open-ended scientific inquiry. Scientists do not follow someone else's "structure" and "format" based on a "prescriptive formula," and they certainly do not strictly adhere to the use of five vocabulary words related to the topic of their research. Consequently, when examined below the surface, BCRs may not be significantly superior at assessing scientific inquiry than multiple choice items. Interestingly, MSDE recently announced that it will eliminate BCRs from the HSA beginning in May 2009 (Kinter & Schwadron, 2007).

Perhaps a summative high-stakes assessment is, in general, epistemologically incompatible with scientific inquiry-based instruction (Zion, Michalsky, & Maverich, 2005). The HSA may be appropriate for measuring a student's knowledge of scientific facts, but insufficient for determining how well a student understands scientific inquiry. Whatever the reason, scientific skills and processes are poorly tested on the HSA. The questions pertaining to them are superficial and narrow in focus. Creativity is stifled by the test, rather than encouraged. Scientific knowledge is reduced and disaggregated into unrelated pieces of knowledge that correspond with specific indicators. Although teachers may receive a barrage of messages about science and science instruction from myriad sources, the HSA is the loudest because of the high-stakes attached to it. Thus, when teachers navigate their way through all of their pedagogical and curricular possibilities, they likely will make choices based on the messages they receive from the test.

Footprints on the Curriculum

The mixed policy messages from the standards and the HSA have left their mark on biology teaching and learning at Halbert High School. Teachers have been along for a circuitous and bumpy ride. Ms. Lydia describes the policy messages as follows:

BCPS has been attempting to align itself with Maryland curricular and teacher standards since the late 90s. First, we got biocurricular drafts. Then, came the county biology exam and the pilot of the HSA. Then, we got the Professional Growth for Teachers' Handbook. Then, they produced a new draft of the biology curriculum, and there was so much mandatory training. They told us one thing, then another. It is hard to keep up. I think they aren't sure what they want us to do.

Ms. Lydia clearly is frustrated with the implementation of accountability and all of its accoutrements. Starting in the late 1990s, teachers at BCPS have been bombarded with the latest round of accountability policy messages. It may be impossible to implement them all in their entirety due to sheer volume, but especially because many of them are contradictory. Ms. Lydia is left with the feeling that BCPS does not know what it wants. Perhaps, BCPS's policies responding to national and state calls for accountability are irrational, but it is impossible for teachers to inoculate their classrooms from some of them because of the high-stakes attached to the HSA. Thus, the policy messages have left some heavy footprints on the biology curriculum.

Accountability Lexicon in the Classroom

As I started to observe the teachers, I became astounded by the infiltration of accountability messages into the classroom. I felt like I had entered a parallel educational universe in HHS biology classrooms. They closely resemble the science classrooms where I taught, but something is eerily different. "HSA," "county exam," "standards," "objective," "Core Learning Goal," and "BCR" are everywhere. During my first week of

observations, I was shaken by a question from one of Ms. Victoria's ninth grade students. Ms. Victoria was administering a test in class. As part of the test, students had a take-home essay on the ethics of genetic engineering. The title of the essay was "Biotechnology Take Home Essay Response." Ms. Victoria had her students read the one page description of the essay topic. One of the students immediately raised her hand and asked, "Essay, is that like a BCR?" No student in the classroom batted an eye. Ms. Victoria was slightly stunned and quickly tried to explain the difference between an essay and a BCR:

Kind of, but no. There is no specific structure [with an essay]. You need to have an intro and a conclusion, but there are no specific points that you need to cover. It's just an essay. You know. You have done one in English class, right? You know what an essay is!

Ms. Victoria did not pass up the opportunity to explain how to write a proper BCR:

With a biology BCR, you don't need to write in complete sentences. You can use graphs, charts, or pictures. Do whatever you have to do to get all the information on there that *they* are looking for. (emphasis added)

That moment was the beginning of my journey down the rabbit hole of accountability. I began to see how strange things have become. I still am not sure that I can completely comprehend the significance of this exchange.

A student felt uncomfortable enough with the meaning of the term "essay" that she felt obliged to ask if it was similar to a term she knew, "BCR." Further, other students did not seem to think it was, at all, an odd question. They engaged with Ms. Victoria's answer and asked follow-up questions. Essays have been a part of the curriculum since before biology existed as a subject. BCRs are a more recent construct of accountability, spreading across state high-stakes tests faster than an epidemic. The accountability lexicon seems to be annihilating everything in its path. An essay is "a short literary

composition on a single subject, usually presenting the *author's* (emphasis added) viewpoint” (Berube, 1995, p. 384). With a BCR, a student is trying to present the viewpoint of the question’s author. That is a both a fundamental and phenomenal difference. If the goal of science education is to promote critical thinking, essays, not BCRs, achieve that goal.

Science standards call for students to write. What they often write may not be called an “essay,” but they write “reports,” “articles,” “memos,” and “manuscripts.” The purpose of scientific writing is to “impart [the scientist’s] thoughts or ideas and their bases and implications” (Carraway, 2006, p. 383). Thus, scientific writing is similar to essay writing, at least in terms of purpose. In both cases, the author is producing and analyzing his or her own work. One form of writing that scientists do not practice is Brief Constructed Response, or writing that superficially outlines someone else’s perspective. By fostering an educational environment where students expect writing to be in the form of BCRs, we are nurturing reproduction and regurgitation rather than independent thought, creativity, and analysis which are supposedly the values embedded in the national and state standards.

Accountability terminology began to gain supremacy in the educational lexicon during the 1990s. Dr. Steven, who has children attending schools in Buckley County, describes when he first heard the terms HSA and BCR:

My kids told [BCR] to me when they were in second or third grade, ten or eleven years ago. I first heard about the HSA from my kids when they were in middle school for Algebra. They started talking about them a lot...if the kids don’t pass, they don’t graduate. The kids have to have them. It was hard for me to fathom that at first.

Dr. Stevens' children started to be indoctrinated into accountability education at the age of seven. The HSA and its requirement for graduation have filtered down to the seventh grade for students taking Algebra. According to Dr. Stevens, his children understood the significance of the words from the first day, but it took him longer to grasp their impact on his children's and students' lives. When students are exposed to the pressures of accountability from the early grades of elementary school, it is no wonder that they understand accountability words better than traditional language associated with education.

According to Tannen (1995), jargon is used when normative words are not sufficient to connote subtle technical meanings, or when people want to establish themselves as an authority. In the case of accountability, the new language deals with establishing two separate, yet related paradigms: objectives and testing. As I describe in the previous section, although these two major components of accountability are interconnected in the sense that assessments are supposed to evaluate whether students have learned the standard, a reasonable alignment between the two has not been achieved. The current manifestation of accountability is all about authority, which is established by the high-stakes associated with the assessment. Thus, of the two components, the assessments hold more authority than the standards, and as expected, testing language is more pervasive than objectives language.

The standards-based lexicon has another significant effect on the curriculum. It may place a wedge between families and schools (Henderson & Berla, 1996). Like Dr. Stevens with his two children, many parents are unable to keep up with the new educational dialect. Ironically, reducing a term to an acronym, like BCR, further

obfuscates the meaning from both laypeople and even members of the educational establishment (Tannen, 1995). A cultural disconnect is created when children speak to parents about school using incomprehensible language. According to Tannen, poor and less educated families, who are the prime targets of *No Child Left Behind*, have the most difficult time adjusting to new language. Although part of accountability is public transparency as spelled out by *No Child Left Behind* (U.S. Congress, 2002b), the language of accountability likely makes the reports on school performance more confusing for the general public.

Pulling the Curriculum Apart

The Third International Mathematics and Science Study (TIMSS) found that U.S. science curricula are unfocused compared with other countries in terms of topics included, repetition, and emphasis (Schmidt, McKnight, & Raizen, 1997). The study found that U.S. science textbooks include far more topics than the 75th percentile internationally. The average textbook covers fifty to sixty-five topics, while Japan covers five to fifteen, and Germany includes seven. Interestingly, the study found that reforms in the United States generally add to existing content without removing anything. Ironically, in addition to being broad, U.S. science curricula were also found to be repetitive and lacked a “strategic concept of focusing on a few goals, linking content together, and setting high demands on students” (p. 4). Unsurprisingly, the diversity of topics and repetition limits the emphasis and depth of the curriculum.

Accountability has mirrored or reinforced the findings of the TIMSS study. The Maryland biology standards cover at least seventy-two topics ranging from biochemistry to ecology, which only includes biology facts and principles. Additionally, biology skills

and processes are covered by forty-two indicators (MSDE, 2002b). Most of the topics covered may be included on the HSA.¹ The BCPS biology curriculum is aligned with the standards and covers all of the topics in the standards. As a result, teachers feel compelled to cover all of the material. Citing a benefit of accountability, Ms. Victoria says:

The HSA holds teachers accountable. Now they have to teach the whole curriculum. If you like ecology or biotechnology, it's too bad. You cannot spend an entire month on one topic. You have to teach the kids all of biology now.

In the extreme case, Ms. Victoria is clearly correct. For example, if a teacher really likes insects, it may be a disservice to students to cover entomology only in a general biology course, because if students want to enroll in AP or college level biology, they need to have exposure to a wider range of topics. On the other hand, students may gain more insights and greater analytical capabilities from a deep knowledge of a few topics.

Nevertheless, how common is this extreme example? And, is a high-stakes exam the only way to address such a concern? How many professional high school teachers only teach one or two topics? If a curriculum is already too broad, is there a reason to broaden it further? Since other countries like Germany and Japan teach only a few science topics, would it not benefit the curriculum to reduce the number of biology topics? The standards and HSA confound some of the problems highlighted by the TIMSS. As I demonstrated in my analysis of the standards and the HSA, the biology curriculum has broadened, and as a result, become increasingly more shallow.

The effects on the curriculum, however, are further nuanced by some unintended consequences of accountability. Overall, the biology curriculum has broadened, but it has

¹ The standards indicate that a few of the topics are not covered on the HSA.

done so while becoming reduced to discreet bits of information. In an effort to prepare students for a comprehensive high-stakes examination, the biology team at HHS tries to make biology more manageable for its students. They have developed a biology HSA review. The teachers looked at the Maryland biology indicators and examined previous HSAs for relevance. The result is a seven page list of topics and diagrams that are supposed to encompass everything covered on the biology HSA. Major topics, like cell biology are reduced to one and a half pages. In effect, these seven pages become the focus of the biology curriculum. However, because the HSA does not test profundity of knowledge and analysis, redundancy supplants depth in the curriculum.

In some cases, biology gets reduced to a set of interrelated vocabulary words. Ms. Calypso says that she was trained to tell students to use at least five related vocabulary words for each BCR. During one of her many discussions of the HSA in class, Ms. Khana tells her students:

You need to know vocabulary words for the County Exam and the HSA. You have to use biology words, at least to get partial credit. Mutation offspring, genetics. If you use these words, the graders will at least think that you are on the right track.

Ms. Khana always compliments students for using biology words, even if they are loosely related to the general topic. Entire concepts may be reduced to a single word or phrase. The following is a vignette from a lesson on evolution:

Ms. Khana: How does evolution occur?
Oscar: Genes.
Jill: Natural selection.
Sam: Mutations.
Ms. Khana: Great! You guys got it.

Yes, genetic mutations occur within a species, and some of them naturally are selected because they are advantageous in particular environments, but the three words cannot

stand alone. They do not give the complete picture. Ms. Khana's students do not necessarily comprehend evolution. They just know some of the terms that are related to it. Ms. Khana does not help them place the words into a context. The students may not understand the meaning of the terms they use, because their knowledge base is built on a framework of words. In effect, the focus on disaggregated, frequently tested concepts has stretched and torn the biology curriculum apart.

Standardized Test Prep

Although the *Benchmarks for Science Literacy* state that common goals do not require a common curriculum (AAAS, 1993), the biology curriculum at Halbert High School has become quite standardized. Ms. Victoria insists that all biology teachers should cover the same material in essentially the same time frame. She says:

All the biology teachers should be on the same page. They should have the same data in the computer. I don't have time to do it, but if I walked around from classroom to classroom the teachers should be covering the same material.

In order to monitor teachers and facilitate homogeneity, all assignments – homework, class work, quizzes, and tests – must be reported into a central database which is monitored within the school. Exam scores from the county biology curriculum also are monitored by the BCPS central office. Even in environmental science classes, which do not have a pre-planned curriculum, three of the four environmental science teachers follow the same lesson plans because Ms. Calypso is the unofficial lead teacher for the new pre-biology program. Thus, a class that was essentially ungoverned by Buckley County accountability structures in previous years, almost instantly has become a standardized pre-biology program.

A rigid, standardized structure can cause serious problems when a teacher veers off course. Alluding to Ms. Lydia, Ms. Victoria describes what happens when a teacher does not follow the standard curriculum:

We all have to be on the same page. We have to be a cohesive team. If there is one member of the team who isn't a team player, it ruins it for everyone. The curriculum falls apart.

It is interesting that Ms. Victoria discusses curricular homogeneity in terms of a team. As the Science Department Chair and a proponent of accountability, it makes sense that she would discuss standardization in terms of cohesion and unity. Nevertheless, her uncritical approach to the curriculum evades the crucial issue of differentiation. She expects the same high standards of all her students, which likely works better in a pre-IB class than a special education class. High expectations should be set for students with special needs, but as Ms. Calypso says:

My students need scientific inquiry. They need differentiation. They cannot just sit there and listen to me speak for an hour or even twenty minutes. It doesn't matter that I tell them [the material]. They need to be engaged at their own pace and in different ways.

All students learn better through a variable science curriculum, but it is especially important for students with special needs (Nesbit & Rogers, 1997; Schmidt et al., 2002).

Further, standardization also has created a rift in what is otherwise a closely knit science department. At the beginning of the year, Ms. Victoria told the environmental science teachers that their class would now be pre-biology. As the veterans, Ms. Calypso and Ms. Lydia agreed on a syllabus for the course. Ms. Victoria vetoed the syllabus because it was not aligned adequately with the biology curriculum. Ms. Lydia protested saying that it was not fair to teach her juniors and seniors biology twice. She threatened to quit, but Ms. Victoria conceded and allowed her "space" to teach her own class.

Consequently, on the surface, the HHS science department has two different environmental science programs: one taught by Ms. Lydia and one taught by Ms. Harris, Ms. Calypso, and Dr. Stevens. Interestingly, by her own admission, Ms. Lydia follows the pre-biology program. When developing her lesson plans for biology, she puts an environmental science cover over the lesson, and teaches the same material to her biology students and environmental science students. For example, while teaching genetics to her biology students, Ms. Lydia uses plant reproduction to teach genetics to her environmental science students.

Ms. Lydia's non-compliance has resulted in her being a partial outsider in the department. She is, at least in part, not seen as a team player. Ms. Victoria is always worried that Ms. Lydia is not teaching the students the material that they need for the HSA and county final exam. For example, Ms. Lydia did not teach her students biotechnology at the same time as the other biology teachers, so she had to, as she calls it, "squeeze it in" after genetics. Thus, Ms. Lydia fell behind the other teachers. Ms. Victoria often would express concern to me about Ms. Lydia's students' level of preparedness for the HSA, especially when Ms. Lydia announced that she would be taking a leave of absence in early May for the remainder of the school year.

Interestingly, because not all topics are tested equally on the HSA, the standardized biology curriculum at HHS does not weight all topics equally. In other words, some topics have gained arbitrary supremacy in the standard curriculum. Because the biology team studied the indicators and the HSA to prepare the HSA Review for Biology, they know that cells, genetics, and ecology are the most frequently tested topics on the HSA. They also know that biochemistry, systems, and evolution are less

frequently tested. In fact, biology teachers' knowledge of the HSA is significantly more nuanced than a general familiarity with coverage. For example, they have broken down DNA replication into four easy steps for students to remember. They know which types of questions are asked in BCRs. They know that kingdoms, math, and technology are almost never tested.

Teachers' knowledge of the HSA is seen almost daily in the curriculum. Biology teachers always point out if a topic that they are covering is frequently tested or not. Topics that are frequently tested are repeated throughout the year as "warm-ups" at the beginning of class, or through causal connections to other topics. Teachers often go on tangents to highlight something about the HSA. For example, Ms. Lydia highlights gel electrophoresis:

There will definitely be a BCR on gel electrophoresis. You must know it. If you will remember anything, at least remember that DNA has a negative charge.

In this example, Ms. Lydia reduces a relatively large concept into a small rule of thumb, remembering that DNA is negative. She does not explain why students need to know that it has a negative charge, or even the significance of the charge. She simply is preparing students for the basics of the test. Further, Ms. Lydia does not have to specify why students need to know gel electrophoresis. Whenever a teacher says that students must know something, it is always for the HSA. Thus, the test, rather than disciplinary or societal importance, becomes the authority on or justification for the subject. Similarly, Ms. Khana, who includes a warm-up on independent and dependent variables during a class that otherwise focuses on gene therapy, warns her students:

This is important. It is on the county exam and the HSA. Some of you missed it on the test in the fall. If you don't do [the warm-up], you will miss it again.

This lesson is disconnected from the rest of the class, but Ms. Khana wants her students to understand that types of variables may be the most important topic they will learn this day. She sends them the message that the test determines what is and what is not important in biology.

These two examples illuminate how certain components of biology gain supremacy due to their presence on the HSA. The effect of this hierarchy of biology knowledge is twofold. First and foremost, certain biology topics are emphasized for no valid scientific reason. Second, an artificial integration rather than logical connection of topics occurs. For example, students may start a class with an activity on food webs during a unit on DNA replication. Such peculiar misconnections create a disjunction in the curriculum. The teacher often abruptly switches topics, disrupts the flow of the class and interrupts the students' processing and integration of the material.

Environmental science teachers, on the other hand, pride themselves on their lack of knowledge of the HSA. Ms. Harris says:

I don't really know much about the biology HSA. We're not allowed to use the biology curriculum. We are told not to look at the HSA.

Dr. Stevens also highlights that he has never seen the biology curriculum or the HSA:

I have never seen the HSA. We are discouraged from looking at the published tests, because we shouldn't "teach to the test."

Interestingly, Dr. Stevens and Ms. Harris are expected to prepare students for the test, without teaching to the test or even seeing it. Perhaps, the environmental science teachers

are not as inoculated from the biology curriculum as they profess to be. Dr. Stevens describes how environmental science teachers have to test students:

I know that when we are going to test students, it's always going to be multiple choice and BCR, instead of asking them any open-ended questions...As a matter of fact, it was suggested to us that we make the final really close to how the [biology] final is set up. Fifty multiple choice questions and three BCR questions, so it's set up the way the HSA is set up. That way the students can get used to it.

Dr. Stevens is expected to write and grade a test that prepares students for the HSA without being exposed to an actual HSA. He has never seen an officially written BCR. When I ask him how he knows how to write or grade a BCR, he shrugs his shoulders and says, "They are everywhere. I just pick it up."

As a result of the focus on the HSA, much of the standardized biology curriculum, at least in part, centers on test preparation. In class, teachers give their students warm-up questions from the HSA. They focus on and repeat topics that are tested most frequently. They go off on tangents toward commonly tested topics that are disconnected from lesson plans. They discuss test taking strategies with students and write tests that mimic the HSA in order to prepare students for its format. For two weeks leading up to the HSA in May, biology teachers spend the entire class period reviewing for the test. The amount of time that is allocated to the HSA is likely to squeeze out valuable learning from the biology curriculum. The biology teachers worry that there is not enough time for scientific inquiry-based instruction. Perhaps they can replace some test prep with it.

Summary

The biology curriculum at Halbert High School has been influenced profoundly by accountability, especially by high-stakes testing. As teachers negotiate their way through the different policy messages, they are compelled to listen to the loudest voice,

the one that threatens their standing and their students' graduation. Consequently, biology has captured the nearly undivided attention of the science department at HHS. Further, the biology classroom has shifted its focus toward HSA preparation, limiting pedagogical diversity, restructuring ongoing evaluations to mimic the HSA, and infiltrating the language of the school and classroom. While compelling teachers to expose students to a wide variety of topics, the high-stakes have caused biology to be reduced to a superficial survey of discreet facts that most frequently appear on the high-stakes test. The influence on the curriculum may be so great in magnitude that it has, at least in part, transformed the kind of course that students receive from biology to biology HSA test preparation. In the following two chapters, I examine how these curricular changes have influenced teachers' experiences and perceptions of their role in the classroom.

CHAPTER VI: THE PRESSURE MOUNTS

Monitoring Teachers

According to the United States Department of Education (2004), *No Child Left Behind* is broken down into four pillars: stronger accountability for results, more freedom for states and communities, proven education methods, and more choices for parents. Although not explicitly stated, the first and fourth pillar combine to form the underlying ethos of the current accountability system, one in which teachers find themselves to be under administrative scrutiny and tremendous pressure. The second pillar states, “Annual state and school district report cards inform parents and communities about state and school progress” (§ 2). If schools fail to make the grade on their report (fail to make AYP), then parents have certain choices, while schools do not. Parents may send their children to a safe school or one that makes AYP, and they can have their children qualify for free tutoring services. The money for these parent choice programs, which would have otherwise gone directly to the “failing” school, instead comes from the district’s education budget (Irons & Harris, 2007). Further, if schools continue to fail to make AYP for five years in a row, “they must make dramatic changes to the way they are run” (USDE, 2004a, § 2).

What is not explicitly stated is that teachers ultimately are responsible for ensuring that schools will make AYP. Therefore, as Ms. Victoria points out, “They are the ones who are being closely watched.” In other words, teachers are held responsible for implementing the school’s strategy to make AYP and remain off the list of failing schools that are “in need of improvement.” In this chapter, I describe how accountability creates a high pressure environment for HHS biology teachers. First, I examine how and

why they are monitored and scrutinized by the high-stakes accountability structures. Then, I explore how they internalize the pressure and how it is manifested in their practice.

No Dentist Left Behind

During our first conversation, Ms. Calypso kept referring to the “dentist analogy” whenever she discussed the effect of *No Child Left Behind* on teachers. I was not clear on what she meant during our interview, but as soon as I got home that day, I received an email from Ms. Calypso that explained what she meant by the “dentist analogy.” (See Appendix F for the full text of this email.)

The analogy highlights some of the major tensions embedded in *No Child Left Behind*. The email contains a conversation between a patient and a dentist. The patient informs the dentist about a new legislation enacted to measure the effectiveness of dentists. The email poignantly underscores three major features of high-stakes accountability. First, dentists’ effectiveness is measured by a single quantitative indicator, the number of calories each patient has at different ages. Second, the power, and, as a result, some degree of expertise lies with the patient and other non-dentists. The patient informs the dentist of the new policy and considers the policy fair and just. Dentists are monitored by a committee composed of laypersons. The patient does not understand why the dentist is dismayed by the policy. The dentist’s efforts to convince the patient of the policy’s flaws are fruitless, and in the end, the dentist is frustrated and the patient begins to think less of the dentist. Second, the email points out that the effectiveness of dentists is, at least partly, correlated to the behavior of patients. In other words, patients with the

same dentist who brush their teeth daily will likely have fewer cavities than those who do not brush regularly.

After reading this email, I began to think differently about why teachers, in particular, perceive that they are under the scrutiny within the structures of *No Child Left Behind*. The analogy highlights three themes that explain how teachers are targeted by the policy, and as a result, why they are under enormous pressure from it. Teachers are stripped of their authority; teaching and learning are very narrowly defined; and the specific context in which teachers work is ignored by the policy. In the following sections, I explore how each of these themes affects the practice of biology teachers at Halbert High School.

Powerless

In reality, dentists do have oversight and controls over them. They are educated by accredited institutions. They have to pass a licensing exam. If they make errors, they can be sued. Patients can choose their dentist, but their choice is generally made on the basis of convenience, cost, professionalism, personality, and expertise. They do not make their choice based on a single measure that is chosen by a board of non-professionals. In other words, dentists retain their dental expertise. They choose how to run their practice based upon what they consider to be their best practices. They choose how to treat patients and what is considered to be proper dental care. Quality is not dictated to them by outsiders.

The dentist in the analogy, however, finds himself to be powerless. The new policy is thrust upon him without any of his own input. In fact, a patient is telling him about it. As the dentist learns more, he becomes increasingly anxious, but he cannot do

anything about it. The dentist's desperation grows as he realizes that the patient does not understand why the new policy is fundamentally flawed. The patient and the state both have power over the dentist, but the dentist is the expert. He is the one who comprehensively understands the nuances of his job. Consequently, no one with authority can appreciate how the policy will impact the dentist and the care he gives to his patients, and the only person who does understand does not have the authority to make any changes.

In reality, of course, it is teachers who find themselves in such a predicament. When I ask the biology teachers at HHS whether they feel included in school policy decisions, all five of the non-administrative teachers say that they feel uninvolved, but not excluded. Only Ms. Victoria feels involved in policy decisions, but her example deals with her role as an administrator, rather than as a teacher. The following example of how Ms. Victoria is influential in setting school policy illuminates how profoundly accountability has reframed the focus of educational planning and goals. Also, Ms. Victoria's solitary involvement in writing the plans suggests that teachers may have a say in developing school policies, but for some reason, they elect not to engage in the conversation.

Ms. Victoria, with little input from teachers or administrators, developed the 2005-2006 HSA Intervention Chart and the 2006-2007 Maryland HSA Biology Improvement/Intervention Plan. As the author and Science Department Head, she considers herself to be in charge of the plan's implementation. Ms. Victoria describes her role in developing the plans and how she feels about the new policy:

Oh, yes, in science I am definitely involved in policy development at the school...I came up with all of the ideas for the plan and [the

administration] supported me. Next year, we'll have the majority of biology kids be on-level, special ed, inclusion, ESOL students. I'm going to have fifteen sections of biology next year. I'm going to have four full-time biology teachers that are going to have to pass this HSA, which means that if they fail in May, they will only have August, October, and January to pass it in order to graduate. So I'm very nervous about that.

Ms. Victoria is alluding to the central feature of her plan, the shift from environmental science to pre-biology. Unsurprisingly, as the policy's architect, she appears to be invested in the policy and wants it to succeed.

Before examining Ms. Victoria's agency and the degree to which she can meaningfully influence school policy decisions, I turn to her Freudian slip. She says that she will have "four full-time biology teachers that are going to have to pass this HSA." Biology students, not teachers, will have to take and pass the HSA. Ms. Victoria's misspeak, may have been an innocent mistake, but it is quite telling in terms of how teachers view the policy, and why they report sensing a lot of pressure emanating from the high-stakes testing policy. As Ms. Victoria insinuates, teachers ultimately are held responsible and/or hold themselves responsible for their students' test scores. It is as if they are taking the test themselves, and they pay the consequences if they fail. In actuality, of course, the students take the test, and teachers mostly are evaluated as if there is no difference between what and how they teach and how students perform on the test.

Besides Ms. Khana, who is an advocate of student responsibility and claims to see herself as a transmitter of information whose job is over after the biology lesson leaves her mouth, each biology teacher outwardly bears responsibility for student failure. Ms. Calypso describes her greatest fear about student failure on the HSA:

It's self perception. Tests don't define who we are or what we do. Many bright students don't pass exams, but have skills that will help them be successful in many areas. But these four tests really dictate how they perceive themselves and their future. I can't let them fail. I can't let that happen.

Like all the teachers, Ms. Calypso is aware of what it means if students fail. She looks beyond the surface and understands the psychological ramifications of failure, and she holds herself responsible for her students' welfare. McNeil (2000) found that most teachers, first and foremost, value the need to help their students. Teachers are held accountable by both Ms. Victoria, as a proxy for the school's administration, and by their own genuine feelings of responsibility for the welfare of their students.

The only apparent exception is Ms. Khana, who seems to hold her students entirely responsible for their own successes and failures. Ironically, in class Ms. Khana focuses more on the HSA than any other biology teacher, which indicates that she understands the significance of student failure, and wants to help her students pass. Ms. Khana often tells her classes, "I'm going to tell you exactly what you need to know for the county exam and the HSA...Every one of you can do well on the HSA. It's only a question of whether you will put in the effort to do well." Perhaps she attempts to inoculate herself from the tremendous moral burden, by claiming to hold students entirely responsible for their own failure. When I ask her if some of her students will fail the HSA, Ms. Khana provides two reasons for students' failure:

Some will fail simply because of attendance. Attendance is a great issue. That is the greatest problem that I have, and, of course, second to that is the motivation. Not because of a lack of ability, it is because of choice that they may not succeed in biology.

Bowe, Ball, and Gold (1992) find that when teachers feel overwhelmed by a policy that is out of their control, they silently may subvert it without explicitly disagreeing with it. Ms.

Khana subconsciously may be compelled to deny the very premise of accountability, that teachers are responsible for their students' performance, by laying all of the responsibility squarely in her students' laps. Interestingly, the view that Ms. Khana espouses about her students' performance on the HSA seems to contradict her own beliefs about teaching:

You can be the best teacher-planner in the world, but kids still will not get the material that you teach. You have to find a way that they would get it. It is a great challenge that can not be taken lightly.

Here, Ms. Khana uses the language of Ted Aoki. Aoki (2005) discusses the distinction between curriculum-as-plan and curriculum-as-lived-experience. A curriculum-as-plan considers teachers to be purveyors of a prescribed curriculum that is perfectly aligned with particular goals or outcomes. The curriculum-as-lived-experience entails the dynamic interactions that actually occur in the process of teaching. Perhaps Ms. Khana sees a disjunction between herself as a planner-teacher and experiential-teacher. She cannot reconcile the contradictory responsibilities of the two roles, and she feels powerless to influence the policy. Thus, her role in the accountability structure is that of an "autonomous resister or subverter of the status quo" (Bowe, Ball, & Gold, 1992, p. 101).

Undervalued

In reality, each teacher experiences powerlessness, to some degree, when it comes to accountability policies. In many ways, accountability structures dictate many facets of teachers' practice. Ironically, their perceived powerlessness may be the cause for their lack of involvement in the development of the HSA Intervention Chart and the Maryland HSA Biology Improvement/Intervention Plan. The teachers, for some reason, may not think like they are able to contribute meaningfully, despite Ms. Victoria's request for

their input. Reay (1998) argues that “The teacher is increasingly an absent presence in the discourses of education policy, an object rather than a subject of discourse” (p. 194). Like the dentist, the teacher is required to modify practice by others who are less familiar with the context in which they work (Sikes, 1992). In other words, in the *zeitgeist* of accountability, teachers are expected to implement policies despite being uninvolved with their formulation. In the case of HHS, each teacher has her or his own reason for being uninvolved, but they all share a connection to feeling powerless and undervalued.

Both Ms. Harris and Dr. Stevens are uninvolved in policy decisions because of their lack of experience. As a substitute teacher, Ms. Harris feels like an outsider in the department and does not feel that she is “qualified to make policy decisions.” Dr. Stevens thinks that as a first-year teacher, he can only learn without providing input. He does not see his own expertise as a teacher:

I do not have an opinion about [*No Child Left Behind*]. I came to teaching when this program was already begun. I did not work before this program was implemented. I do not have any view on this legislation and its challenges...Since it is my first year teaching, my experience is still forming. I’m just putting everything together right now, and my colleagues are the biggest influence on me. They shape what I am doing, and how I am doing it.

How is this possible? How could Dr. Stevens have no opinion about the most prominent legislation governing his profession? His uncritical perspective might be attributed to feeling undervalued and unqualified by the accountability system, like a “being-as-thing, a technical being devoid of his own subjectivity” (Aoki, as cited in Pinar, 2005, p. 3). During our conversations, Dr. Stevens often used his novice teacher status to deflect questions that required him to pass judgment or make suggestions. Ultimately, it is quite possible that the accountability structures amplified Ms. Harris’ and Dr. Steven’s

inexperience and lack of expertise. During our conversations, they never referred to their own agency in the classroom. They always deferred to other teachers, administrators, and the requirements of high-stakes testing.

The accountability climate appears not to value teachers, and it seems to create an educational façade in which teachers have prescribed roles to which they may not, themselves, subscribe. Dr. Steven's expectations of what it would be like to teach were shattered when he walked into the accountability environment at the HHS Science Department. When describing his job as a teacher, he says, "When I came to this school, I thought that I would be teaching students how to learn, not facts." Perhaps, Dr. Steven's realization sheds light on his perceptions of himself as a teacher. Apple (1992) argues that if teachers feel excluded from the policy, they may become disenfranchised by it if the policy's messages conflict with their own notions of quality teaching. Dr. Stevens acts as if *No Child Left Behind* is imposed upon him, as evidenced by his unwillingness to discuss or consider its merits and drawbacks. Like an omnipresent deity, it seems to him to be a given, a central part of today's educational climate.

Ms. Lydia, on the other hand, is an outspoken critic of *No Child Left Behind*.

When I ask her opinion of the legislation, she replies bluntly:

It sucks! It has no value! It's touted as a panacea to solve societal problems...but you are taking away the arts. You are taking away a piece that is invaluable. You are killing the holistic approach to teaching. You are putting in something that [students] cannot use and destroying the critical inquiry. It should be cherished. *NCLB* is beneath the standards. It's beneath our government. We are spending millions and billions on one test that is supposedly monitoring our children's proficiency. It's crazy! It's a big travesty, and I regret that I am in teaching when this happened.

It would be difficult for Ms. Lydia to be more critical of a policy. She truly is frustrated with accountability and its effects on the curriculum. Further, on several occasions she

describes how she wishes that educators would object and stand up to the policy. She provides an account of how she would like to protest:

As a citizen, I am outraged. I want to go down and picket, and chain myself to the White House fence with a big sign that says “This sucks.” The people that we put into office are supposed to be working for us. I really resent...the philosophy that they have to help the citizen as though we have no sense at all, that we are not rational and can’t think on our own, that they have to protect us or tell us what to do. That’s Big Brother. No, I don’t go for that. I have more degrees and know more than anyone there, and I probably have more morals.

It is interesting that Ms. Lydia talks about her civil disobedience as a citizen, rather than as an educator. The disenfranchisement of teachers runs deep. Her statement suggests that as a citizen she has more power to influence change than as a teacher. Like the dentist who is governed by the DOC, “a group made up of mostly lay persons to make sure dentistry in this state gets improved,” teachers are governed by the U.S. Congress and an educational bureaucracy above them who are unaware of how the policy manifests itself at the classroom level (Sikes, 1992). Ms. Lydia claims to be an educational expert, but still considers herself to be unable to stand up to the accountability structures. In other words, even as a veteran teacher, she may feel overwhelmed by her standing at the bottom of *No Child Left Behind*’s monolithic bureaucracy.

Ms. Calypso appears to hold a more optimistic view of the potential of *No Child Left Behind*. As a special education teacher, she sees value in the attention that the legislation brings to children with special needs:

A positive aspect [of *No Child Left Behind*] is more inclusive practices for students with disabilities. It helps teachers see that no child should be left behind. As a teacher, you have a set of standards that you need to be teaching to. You should teach to every student.

Ms. Calypso's rationale resembles that of Ms. Victoria, who argues that *No Child Left Behind* ensures that teachers are teaching the entire biology curriculum, rather than "wasting time." However, the subtle difference between Ms. Calypso's and Ms. Victoria's argument may actually be quite profound. Ms. Victoria focuses on content, while Ms. Calypso focuses on students. As a special education teacher, Ms. Calypso has a vantage point that Ms. Victoria does not share. While Ms. Victoria teaches HHS's pre-IB and IB students, Ms. Calypso teaches students with emotional and learning disabilities who have traditionally been ignored by our system of education. Although Ms. Calypso concedes that the attention her students get with *No Child Left Behind* may be less than perfect and even may stigmatize them further, she contends that it "is better than nothing."

Despite her optimism, she sees many pitfalls to *No Child Left Behind*, especially how it is being implemented and how it subjugates the work of teachers. Ms. Calypso sees a major flaw in holding teachers entirely and solely accountable for their students' performance. Referring to the dentist analogy, she describes her perception of this injustice:

So you go to the dentist regularly, but you get a cavity. Is it the dentist's fault? He flossed your teeth, and he cleaned them every six months. The dentist did what he was supposed to do, but to hold him accountable for your cavity. I think it's a little ridiculous. There is accountability in teaching, but I think that holding all teachers to the same standards, regardless of who their students are, is ridiculous.

Ms. Calypso highlights how unreasonable it is to expect the same of every classroom and to hold the teacher accountable while ignoring the specific context in which the teaching takes place. In reality, schools and classrooms are contextualized entities that are informed by the individuals present (Clandinin, 1986). *No Child Left Behind* strips

teachers of their autonomy and character. What Ms. Calypso describes as “ridiculous” actually may de-professionalize teachers. What is a teacher without a context and students? Why is it necessary to have a teacher when we can just have a tape recorder play the standards to the students? Giroux (1988) argues that teaching is much more complex than simply mastering a body of knowledge or a curriculum. Teaching is an intellectual endeavor that is both responsive and reflective of the context in which it takes place. Teachers’ practice should be defined by the specific context and their ability to mediate between different groups and persons. Ignoring the context in which teaching occurs completely disregards the teacher’s agency and role in the curriculum, reducing him to a being-as-thing (Pinar, 2005).

Unsurprisingly, Ms. Calypso feels powerless to influence the school’s accountability policies. More so than any of the other biology teachers, Ms. Calypso has Ms. Victoria’s respect and ear. They are friends, but more importantly, Ms. Victoria trusts Ms. Calypso’s professionalism, knowledge, and pedagogy. Nevertheless, when Ms. Calypso developed the first environmental science curriculum, it was vetoed by Ms. Victoria for not being aligned closely enough with the biology curriculum, but, in reality, it was not really Ms. Victoria’s judgment. She, too, is following the rules. It was the call of accountability, one that all teachers and administrators must heed. In this case, despite being in charge of the environmental science curriculum, Ms. Calypso is teaching someone else’s course. She often says, “I’m supposed to be including more literacy skills” or “they tell me to teach them graphs, so I am putting them in.” Here, Ms. Calypso is referring to what Ms. Lydia calls “Big Brother.” Words like “they” or passive phrases

like “I’m supposed to” may demonstrate how teachers perceive their own position next to the almighty policy. There is no escape from it or anyone to turn to for help.

Ms. Khana claims to be uninterested in influencing school or departmental policy. As a teacher, she plays the role of the silent rebel. She claims to subscribe to accountability, while holding that she is not responsible for her students’ behavior and is not as affected by the policy as other teachers. In fact, describing the legislation’s effects, Ms. Khana says, “I imagine that it has made some teachers’ lives much harder.” She, of course, denies any effect on her work despite clear evidence of its influence in her classroom. Her stand on intra-school policy is also one of denial. When I ask her whether she is involved in Halbert High School’s policy decisions, she responds, “Yes, to the extent that I would want to be involved.” Ms. Khana may not want to be involved because she feels devalued by the system, powerless to change it, or hopeless about its effect on her students’ success. In the current high-stakes accountability climate, rather than being more involved in policy formulation, teachers like Ms. Khana may be induced to inoculate themselves and their students from school policies that they deem to be inappropriate.

Ms. Victoria’s case is the most revealing of the role of power in this context because on the surface, she has significant influence over policy developments that relate to the science curriculum. She, nearly single-handedly, developed the 2005-2006 HSA Intervention Chart and the 2006-2007 Maryland HSA Biology Improvement/Intervention Plan. The schools’ administration has granted her the authority to change the very nature of the environmental science course in order to include another step in the biology program at HHS. However, although this change may be profound, Ms. Victoria is not

necessarily the agent of change. In other words, she may usher in the change, but she is not necessarily its creator.

As a science department head, Ms. Victoria is a cog in the wheel of accountability. The pressure of accountability comes down hard, especially in BCPS. As Ms. Claypso reminds me, “There is lots of pressure for schools to look good in Buckley County.” What it means to look good has been redefined significantly by *No Child Left Behind*. In the accountability climate, district and school level discussions of success are limited to analyses of scores on standardized tests (Booher-Jennings, 2005). As such, although the administration does not explicitly tell Ms. Victoria what to do, she feels compelled to tow the line of accountability. Ms. Victoria describes the pressure from the HHS administration:

Yes, I feel pressure. It’s just expected. I know [that] I expect it. They don’t say anything, but what are they going to say? I know they are there. I know they expect the kids to succeed.

Jacob (2002) finds that educational accountability often induces administrators to exert psychological pressure on teachers to bring them into policy compliance. The pressure is omnipresent because it does not originate necessarily from school administrators. They, too, are simply part of the accountability system. The district feels the pressure. The school feels it. Ms. Victoria feels it, and the teachers in the science department feel it. Accountability is like a gas that has spread throughout the science department, Halbert High School, and BCPS.

Teaching in the Accountability Vice

The word “vice” originates from the Old French *vis*, meaning “device like a screw or winch for bending a crossbow or catapult” and from the Latin *viere*, meaning “to bind,

and twist” (Harper, 2001, ¶ 1). The word may be appropriate to describe the position of teachers in the climate of accountability. High-stakes testing bends, binds, and twists teachers and their practice to conform to particular protocols determined outside the classroom and school. The accountability movement has been organized around a concern for external oversight, regulation, and judgment of performance due to an institutional mistrust of educational producers, especially teachers (Apple, 2001; Mentor, Muschamp, Nicholls, Ozga, & Pollard, 1997). Oversight in accountability comes by strict management, both active and passive.

Control over biology teachers at Halbert High School can be seen throughout the curriculum. Ms. Lydia offers an overview of how state accountability measures translate into control over teachers:

Maryland was prescient, in the sense that [the State Board of Education] already knew [high-stakes accountability] was coming. BCPS was always being pressured by the state to develop a curriculum, core learning goals, indicators. The County has always been trying to align itself with the state. It has been trying to develop a curriculum and do some teacher-proofing on its own. Even though it’s been saying that it hasn’t, it’s been trying to make things more prescriptive.

The County’s efforts at “teacher-proofing” the curriculum are disguised under the auspices of standardization. Ms. Lydia goes on to describe how the new biology standardized curriculum was presented to the teachers.

They say that it’s altruistic. If they left this school and went to Riverside,¹ per se, we would be on the same page. We’re not being a disservice to the child if they move from one school to the other. We are all teaching the same thing. I am not doing woodworking and someone else is cutting up a frog or something. And that’s not bad. I can agree with that.

Ms. Lydia’s description is telling for two reasons. First, it displays how accountability is synonymous with control, or as she calls it, “teacher-proofing.” She offers an overview of

¹ “riverside” is a pseudonym for another high school in Buckley County.

how the district responded to the state's accountability mandates, by developing a "prescriptive" curriculum. Her words demonstrate how intertwined accountability is with mistrust of teachers. In her mind, the two are inextricably linked. Accountability reflects a technical input/output paradigm of teaching where teachers are the medium for transmitting standards to students. The accountability paradigm sees the curriculum solely as environment-producing (Pinar, 2002). Any knowledge production that goes on is useless, at best, because it is not part of the standards and, counterproductive, at worst, if it somehow obfuscates or contradicts the standards. Thus, accountability messages suggest that the teacher should be transparent, producing as little knowledge interference as possible.

Ms. Lydia's account also demonstrates how "teacher-proofing" can be camouflaged under the auspices of curricular standardization. This way, teachers can be complicit in their own de-skilling (Apple, 1992). As Ms. Lydia says, a standard curriculum ensures that all students in the county are getting the same material, which appears fair on the surface, so teachers buy into it. Even Ms. Lydia, an experienced and knowledgeable educator who is extremely critical of accountability, "can agree with" giving students equal access to knowledge. Ms. Lydia internalizes and perpetuates the myth that without being controlled, teachers will waste their students' time. The equal access to knowledge argument gets Ms. Lydia to subscribe to a process that she would otherwise condemn, demonstrating the pervasive psychological and sociological authority of *No Child Left Behind*.

The “Highly-Qualified” Teacher

Accountability uses its authority to rigidly define or redefine teaching and learning, and, sardonically, the new educational structure that emerges may add further controls on teachers. Thus, continuing with the vice metaphor, teachers find themselves bound and twisted in a continuously narrowing work environment. When student learning becomes “subservient to testing, curriculum becomes bounded by what is testable, instruction wrapped around in bubble-in worksheets, and human worth (that of teachers, students, and institutions) determined by the rankings and ratings produced by the accountability marketplace” (Craig, 2004, p. 1240). The biology curriculum at HHS is surreptitiously enclosing around the teachers.

First and foremost, the very meaning of a qualified teacher has been defined by *No Child Left Behind* in terms of credit counts. The legislation has based its definition of “high-quality” on teachers’ subject matter knowledge as measured by college credits completed in a particular subject (U.S. Congress, 2002a). The underlying ethos of this definition is that teachers must have subject-matter competency for the subjects that they teach. Intuitively, it is commonsense that teachers ought to be proficient in what they teach. Perhaps the number of credit hours completed in biology is a reasonable proxy for proficiency, or the Praxis II exam may be another indicator of biology content knowledge. In addition to the content knowledge requirement for the “high-quality” label, *No Child Left Behind* requires that teachers hold a valid teaching certificate. Teachers who held such a certificate prior to the enactment of the legislation were grandfathered into this requirement. But, beginning in the 2006-2007 school year, all teachers in public

schools are required to meet the highly-qualified criteria in the classrooms in which they teach. In other words, all biology teachers must be certified in biology.

Although every student certainly deserves to have a “high-quality” teacher, *No Child Left Behind*’s definition of quality has had several major ramifications for the HHS science department. First, it has created a teacher shortage by unnecessarily excluding teachers like Dr. Stevens. In the words of Dr. Stevens discussing what he will be teaching next year:

I’ll teach whatever they need me to teach. We just don’t have enough qualified teachers to teach all the classes. I am certified in physics, but I know science and math. I can teach whatever, but I am not allowed to. Like I am only a few credits short in biology, but I’m not allowed to teach biology.

Dr. Stevens has a degree in chiropractic medicine and worked as a chiropractor for years, yet according to *No Child Left Behind*, he does not have the content knowledge to teach high school biology. The most obvious effect of the “high quality” label is that HHS has had to eliminate Earth/Space systems from its curricular offerings because they do not have a “high-quality” teacher for the course. On the one hand, in a context where biology is swallowing the entire science curriculum, it seems particularly tragic to lose another science course. On the other hand, through the lens of accountability, teaching Earth/Space Systems may just be a waste of time when the departmental focus is to prepare students for a high-stakes exit exam that does not include Earth/Space Systems content.

The “high-quality” label could, of course, undermine the goal of preparing students for the biology HSA if there were a shortage of “highly-qualified” biology teachers. In 2000 the National Survey of the State of Science and Mathematics Education

revealed that only about fifty percent of middle school and high school science teachers consider themselves to be very well qualified to teach biology (Weiss, Banilower, McMahon, & Smith, 2001). Therefore, the new teacher requirements may stress an already strained infrastructure of available biology teachers. I asked Ms. Victoria if a biology teacher shortage was a concern, and her response is:

It's a major issue. I don't know where I'll get biology teachers for next year. The County Science Department is recruiting all over the country. They are going to conferences, colleges, and offering signing bonuses. I don't know whether they'll get them.

Although it may be difficult to recruit “high-quality” biology teachers, it seems that BCPS is willing to do everything it takes to ensure that its schools will have enough biology teachers. It does not appear, however, to have the same attitude toward recruiting teachers for Earth/Space Systems. BCPS's recruitment efforts are another demonstration of biology's supremacy in the science curriculum.

Interestingly, the shift from environmental science to pre-biology allows HHS to alleviate some of the stress of hiring “highly-qualified” biology teachers. In the current biology-centered climate, HHS cannot allow itself to be out of compliance with *No Child Left Behind* in biology instruction. There is too much at stake. Environmental science, however, is not a tested subject. In fact, it does not even have a district mandated curriculum. Therefore, it is more reasonable to be out of compliance with *No Child Left Behind* mandates in environmental science than biology. Conveniently, at HHS environmental science is biology. Besides Ms. Lydia who only teaches one section of environmental science, the environmental science teachers are not certified in biology. Ms. Harris is not even certified in science. In the words of Ms. Calypso, “There is no way we could have a substitute teaching biology.” Using the cloak of environmental science,

HHS only employs three “highly-qualified” biology teachers instead of six, but remains in compliance with *No Child Left Behind*.

Next year, the new requirements may pose a major problem for HHS. Ms. Calypso, who is certified in special education, is seeking national certification in biology. Despite holding a B.S. in biology, Ms. Calypso fears that she may fail to get certified because her biology content knowledge is inadequate compared with teachers like Ms. Victoria who teach pre-IB and honors biology. Without certification in biology, she cannot teach biology next year to students with special needs. Research has found that students with special needs benefit the most from hands-on, inquiry-based instruction (Schmidt et al., 2002). If Ms. Calypso fails to become certified, the HHS biology team will have to make due without the teacher who is very well versed in scientific inquiry and is highly-qualified in teaching students with special needs, who also happen to be at a greater risk of failing the HSA.

The Breakdown of Teaching and Learning

The image [of education] involves ideas of breaking things down, fragmentation, isolation, and the consequent dispensing, manipulation, and control of the smallest, simplest, most meaningless bits and pieces of the living inheritances that are entrusted to teachers and learners in school...Imagine if we treated *these* things as the basics of teaching and learning: relation, ancestry, commitment, participation, interdependence, belonging, desire, conversation, memory, place, topography, tradition, inheritance, experience, identity, difference, generativity, intergenerationality, discipline, care, strengthening, attention, devotion, transformation, character. (Jardine, Clifford, & Friesen, 2003, p. 113)

Like other professionals, teachers want to be successful and have the institution that they represent achieve or maintain high status. Under the umbrella of accountability, success in Maryland is narrowly defined by attendance data and test scores. When I ask

Ms. Victoria about students' experiences with the HSA, she immediately starts to run off her students' stats:

This year, I am predicting a seventy percent or higher passing rate on the HSA. That's what I'm hoping because I think we have the right kids and the right classes that are taking it. So I'm predicting that we go from a fifty-seven percent passing rate last year to increase to seventy. I am hoping. The kids that aren't passing are the ones in the bottom area of our intervention. That mainly means that they are not going to pass because they did not come to class. There were twenty-four students that LC'd (lost credit). I don't think that those kids are going to pass [the HSA]. Out of 288 kids who were supposed to take the High School Assessment, by the third make-up only three kids had not taken the test. There are only going to be three kids who are going to get a zero because they didn't show up for the test. Three out of 288. That's good.

Ms. Victoria appears trapped within this rigid, numerical definition of success. The accountability system may have reduced how Ms. Victoria understands teaching, learning, and interactions between students and teachers. As head of the science department, she may make decisions and even enact policies, but when considering her choices and making judgments, Ms. Victoria thinks of scores on the HSA, attendance data, student statistics, and passing rates. She does not seem to question whether these numbers capture the essence of teaching and learning.

The meaning of education gets reframed by both standards and assessments.

According to the legislation itself, *No Child Left Behind* incorporates principles and strategies that

include increased accountability for States, school districts, and schools; greater choice for parents and students, particularly those attending low-performing schools; more flexibility for States and local educational agencies (LEA's) in the use of Federal education dollars; and a stronger emphasis on reading, especially for our youngest children. (U.S. Congress, 2002b, p. 1)

The paramount education bill of the current administration highlights some of what it deems to be the most important educational goals. It is unsurprising, yet interesting to note, of course, that the only mention of learning pertains to reading. The remaining goals deal with accountability, choice, and flexibility. The standards movement has reduced the meaning of teaching and learning and codified it with using colonizers of the educational lexicon.

Interestingly, these new words, which may have not been familiar to everyone involved in education, are defined in the *No Child Left Behind* Glossary (USDE, 2004b):

Accountability System - Each state sets academic standards for what every child should know and learn. Student academic achievement is measured for every child, every year. The results of these annual tests are reported to the public. (§ 2)

Public School Choice - Students in schools identified as in need of improvement will have the option to transfer to better public schools in their districts. The school districts will be required to provide transportation to the students. Priority will be given to low-income students. (§ 20)

Flexibility - Refers to a new way of funding public education. The *No Child Left Behind Act* gives states and school districts unprecedented authority in the use of federal education dollars in exchange for strong accountability for results. (§ 14)

Along with more emphasis on reading, these are some of the pillars of *No Child Left Behind*. It may not be immediately clear how they can meaningfully improve teaching and learning. Even reading is actually Reading First, “a bold new national initiative aimed at helping every child in every state become a successful reader,” which dismembers reading into five components – phonemic awareness, phonics, fluency, vocabulary, and comprehension (USDE, 2004b, § 21). This is a consequence of

accountability. It compartmentalizes knowledge and information into bits, phrases, and gimmicks.

Accountability also creates labels. The passage from the *Executive Summary* discusses low-performing schools. Although the glossary does not include the term “low-performing,” it does include “Adequate Yearly Progress, an individual state's measure of yearly progress toward achieving state academic standards” (§ 4) and “distinguished school,” a label given to schools that “make major gains in achievement” (§ 11). By deduction, a “low-performing school must be one who fails to make AYP. According to the legislation, if a school fails to make AYP it is “in need of improvement” (U.S. Congress, 2002a). Therefore, a “low-performing school” must be “in need of improvement.” Further, a “low-performing” school must be the opposite of “distinguished school,” which suggests that is making major steps toward failure (to make AYP).

These terms obfuscate the truth. First, they steal attention away from real educational issues. Then, they shift the curriculum conversation toward a technocratic discourse that has dissociated education from its traditional emphases (Ferneding, 2004). In effect, there are two simultaneously occurring educational discourses. One deals with the traditional aims of education, while the other one focuses on management and manipulation of education in order to achieve predetermined goals efficiently (Reynolds & Webber, 2004). Further the goals, themselves, are either vague or distorted and limited constructions of achievement that ultimately devalue education. The result is not a “dialectic, but rather a dichotomy of where the possibility of negotiation between the two spheres of discourse has broken down under the weight of the discourse of inevitability”

(Ferneding, 2004, p. 59). In the case of high-stakes accountability, if a contradiction occurs, the accountability discourse supersedes traditional educational discourse because of the penalties of failure.

Due to the extent of the policy's penetration, *No Child Left Behind's* reduction of teaching and learning may have serious ramifications both inside and outside the classroom. Interpersonal relationships between members of a classroom fundamentally shape the environment of the classroom (Hirshy & Wilson, 2002). In the previous chapter, I demonstrate how accountability has infiltrated the biology curriculum and the HHS biology classroom, and how it has reduced biology into isolated and discreet pieces of knowledge that are sometimes as narrow as a single word or phrase. At this point, it is integral to consider the ramifications of this infiltration on biology teachers' practice.

First, the teachers' focus is centered on content, rather than scientific process. All the teachers, with the possible exception of Ms. Calypso, primarily see biology and science in terms of content, rather than skills. Although every teacher includes the phrase "a way of viewing the natural world" in their definition of science, my observations suggest that the biology teachers primarily consider biology to be a body of knowledge. Ms. Khana suggests:

With this curriculum we have to teach them the facts. The BCRs [Brief constructed Response] make them think and analyze, but they have to have the [biology] content. They have to have the fundamentals to do well on the test, to do well in this course. If they don't learn what I teach them, I don't know what will happen.

Ms. Khana's description is very revealing. She clearly describes the biology course as a particular body of knowledge that the students must know in order to be successful.

Interestingly, the body of knowledge actually is disaggregated into what Ms. Khana calls

“facts,” and the State of Maryland and BCPS call “indicators.” Also, Ms. Khana conflates doing well in the course and doing well on the HSA. Since she does not write the HSA, Ms. Khana may be acknowledging that an external assessment that includes no input from her has the greatest influence on the curriculum in her classroom.

If everything that students need to succeed is preset outside the classroom, and students’ graduation is predicated on their performance on the HSA, teachers have an enormous amount of pressure to ensure that students do well. As knowledge transmitters, teachers need to be experts on what is required for students’ success, the HSA. As Ms. Lydia tells me, teachers “study the indicators, but more importantly, we know the tests. We know what BCRs they ask. We know what is tested most frequently, and we give it to our students.” Teachers also must relay all of the information to students. They, alone, are held responsible for imparting the predetermined set of biology fragments to students.

Additionally, such a construction of the biology curriculum and the pressure that stems from it is likely to curb creativity, which contradicts the recommendations of the national standards (AAAS, 1993; NRC, 1996). Unsurprisingly, I witnessed little creativity in the biology classes. None of the three biology teachers engaged their students in long term projects or activities that included technology. Ms. Khana and Ms. Lydia assigned a “paper lab” where students made DNA bracelets, which included drawing and coloring of different base-pair interactions. Unfortunately, the “creative” piece was disconnected from the focus of the assignment, which was getting students to be familiar with base pairs. The decorating was done by students as “down-time” at the end of class. Even when teachers played games with their students, like Jeopardy, the questions were all based on the HSA or County Exam. The message to students seems to

be that creativity is superfluous to science, rather than being an integral part of scientists' work. By contrast, environmental science teachers did include research projects, experimentation, presentations, and model design in the curriculum. In this regard, the difference between the two courses was profound, indicating the curricular significance of a high-stakes exam that looms at the end of a course.

Finally, biology skills like reading, writing, and comprehension, which are part of the scientific inquiry process, are now rooted in test preparatory pedagogy. When Ms. Victoria describes her vision for the environmental science curriculum, in addition to teaching biology content, she wants teachers to focus on "reading and writing in every class period - building vocabulary, teaching [students] how to read and write, developing literacy skills." Although literacy is an essential component of scientific inquiry, Ms. Victoria is motivated by students' performance on the biology HSA, not their understanding of doing science. Even skills that are more directly related to scientific inquiry, like reading charts and graphs, are included in the curriculum for the purpose of the HSA, not inquiry-based instruction. Thus, Ms. Calypso confesses that she sometimes "throws charts or graphs into lesson plans just because students need them for the test." As such, her class work with charts and graphs is often choppy and disconnected from the content that students are exploring. Further, Ms. Calypso reports that she does not enjoy going over these "skills" because her students just cannot follow her instructions. She is fitting her curriculum to the structures of accountability, which has shifted teachers' understandings of which skills are necessary for biology and created pressure to include more test preparation in the curriculum.

Time

Time, one of a teacher's most valued commodities, is also severely reduced by accountability measures. In fact, when I asked the teachers how they would change school policy to improve their work, all six teachers told me that they need more time. Although teachers are likely to request more time to prepare for class, the HHS biology teachers allot a tremendous amount of time to do work that can be attributed to accountability measures. As several of the teachers pointed out, the biology team does not get compensated for the extra pressure and workload involved in preparing students for the biology HSA.

Ms. Victoria describes how accountability has influenced the work of the biology team:

Just in the amount of work we have to do now. Just that we have to meet several times a week to talk about summative assessments; we have to input scores into the database; then, we have to talk about how we are going to re-teach the kids; who's staying after school [for High School Plus]; who's doing lunch remediation; who's going to call these kids... We do Connect Ed calls; we write letters; we make these kids come for remediation. We just didn't have the energy and the time to do it right this year.

The structures of accountability at HHS and BCPS have created an inordinate amount of work for teachers. Ms. Victoria is not describing work that is related to pedagogy or science. All the work that she is describing pertains to two facets of accountability, monitoring and test preparation. For example, the HHS biology team has instituted a lunch program and an after-school program that are specifically designed for students at risk of failing the HSA. According to Ms. Victoria, after spring break, which is about seven weeks before the administration of the HSA, she goes "into kid's classrooms, pull[s] them out of seventh period class right at the end of school, and make[s] them

come to [her] room for remediation.” This year, because of Ms. Victoria’s HSA administrative duties, she was not able to do “it until the beginning of May.”

The context of the remediation is test preparation. Ms. Khana describes what she teaches after school:

We prepare them. We give them practice BCRs. We give them passages to read. We go over the questions. If they come, they will pass biology.

Ms. Khana’s description is very telling. She does not explicitly refer to test preparation or the HSA. She does talk about practice BCRs and going over questions. Such test preparatory teaching practices have become so embedded into the biology curriculum that I am not sure whether Ms. Khana would consider this test preparation or simply “extra help” with class work. Passing biology has become synonymous with passing the biology HSA and the biology county final exam.

The other type of accountability work that consumes teachers’ time and energy pertains to monitoring teachers through their students. The HHS biology teachers are now expected to contact parents using multiple methods, report students’ attendance and test scores, and meet to discuss assessments. Ms. Victoria maintains a central database of test scores, attendance, homework, and worksheets. Essentially, teachers must track all of their students’ work. Ms. Victoria oversees this process within HHS. In terms of BCPS, teachers are required to administer pre-assessments and post-assessments for every chapter in the curriculum modules. Starting in 2007-2008, BCPS will monitor scores for every assessment that is administered to students.

The manifestation of accountability seems to translate into teachers being monitored in everything they do with students. Sardonicly, they are asked to facilitate, even manage, their own oversight by providing information on the students. Although

these prescriptions are mandated under the guise of student oversight to ensure progress toward particular goals, teachers understand that they are being monitored and compared to other teachers in the district. This type of oversight is not limited to biology. A teacher of chemistry, which also has a county midterm and final, describes her discomfort at an informal teacher gathering:

I'm uncomfortable with these tests. It's not just biology teachers and the HSA. We know that we are being compared to other teachers at other schools. It's not fair. They don't take our students into consideration. I don't want to be compared to another teacher in another school.

It seems that in BCPS, accountability may be pervasive throughout the educational culture and spreading to subjects that are not even tested by the state.

In order to monitor teachers effectively using the proxy of student scores, teachers need to do a lot of grading. In the past, teachers could go over homework in class or have students do worksheets without collecting them. Currently, at HHS the three biology teachers must grade and record every assignment. The environmental science teachers also must record and report their students' progress to Ms. Victoria, but the process is less rigid and formal. During a conversation after an observation, Ms. Khana reports grading over five thousand papers per semester. She says:

I spend hours and hours grading. It feels like all I do is grade papers. I cannot believe how much time it takes. I am exhausted by it almost every night.

When I ask the other teachers how many papers they grade, Ms. Lydia and Dr. Stevens both tell me that they have tried to figure it out, but quit when they got to "the several thousands." All of the teachers complain about grading and report that an overemphasis on grading student assignments interferes with their preparation and teaching.

The Intensification of Teaching

Apple (1992) calls the incursion on teachers' energy and time "intensification." He argues that intensification leads teachers to cut corners, so only what is "essential to the task immediately at hand is accomplished" (p. 426). He says that it forces teachers to rely on experts and second-guess their own expertise. Intensification promotes isolation from colleagues and supervisors because teachers no longer have time to engage with one another. It damages teachers' pride in their work because they are implementing someone else's conception of what should be done. The biology team at HHS has experienced all of these effects, and in some cases, intensification may have even surpassed Apple's depiction.

Cutting Corners

In terms of cutting corners, the teachers at HHS were remarkably organized and thorough, considering the circumstances. The mere fact that I was welcomed into the department indicates the willingness of teachers to share their time and experiences. Was their participation an act of resistance to the accountability climate? Perhaps, participation in this study allows teachers to engage in what they perceive to be professional work while their status and power are diminished by the accountability system.

Although they are under a major strain to meet deadlines, grade papers, monitor students, offer a variety of interventions, and prepare for class, the six biology instructors at HHS do a remarkable job managing multiple competing directives on their time and energy. Nevertheless, Ms. Calypso, for example, frequently has to attend IEP meetings or other interventions with special education students. Often the meetings interfere with her

scheduled class time. As a result, Ms. Harris and Dr. Stevens are placed in a position where they have to teach the entire class by themselves, a task they usually cannot handle well. Without a second teacher, the classes are chaotic and largely unproductive. Similarly, Ms. Lydia would sometimes tell me that she was unprepared for class due to grading demands or parent phone calls. Ms. Victoria often complains about being overworked and staying up all night. In class she is rushed and always seems to be a step behind where she wants to be.

Essentially, although the teachers manage to juggle many responsibilities, they are forced to cut corners. Unfortunately, the corners usually come from the time teachers spend preparing for class. Thus, their pedagogy, the work that defines teaching, is likely to suffer. In effect, intensification begins to erode the center of a teacher's identity, which is formed through interactions and work with students, rather than tasks like data entry and test score analysis.

Hiding Behind the Structures of Accountability

Apple (1992) also reports that teachers are forced to rely on experts. In the case of HHS biology, the teachers are essentially left alone by the school administration. Ms. Victoria describes the teachers' relationship with the school administration:

I'm not sure...I guess that I just don't know if they are really aware. I don't know if they are aware. For them it's just about the HSA score. They have no idea what's involved to get these scores. And, it's frustrating.

In the case of relying on experts, as a department, the biology team has isolated itself from reliance on others, for now. It is impossible to predict what would happen if the environmental science intervention fails, and ESOL or SPED students' HSA scores do not improve in 2007-2008. For now, however, the administration has given the team

some space. Perhaps their reticence to be involved stems from a fear that the administration will be held accountable for the interventions. If Ms. Victoria makes the decisions, she can be blamed for their failure to produce results.

Although the biology team essentially may be left alone by the school's central administration, the structures of accountability are still present. They are the experts that teachers must turn to for guidance. The HSA becomes the curriculum guide, and BCPS accountability structures are the experts. Ms. Khana describes the County curriculum modules:

I like them, but they're not perfect, far from it. I would definitely use some of my own stuff. I fit it in here and there, but I am encouraged to use [the County curriculum]. I know if I use it, they cannot complain that I'm teaching the wrong stuff. I have to be careful, you know.

Here Ms. Khana demonstrates that her own expertise is diminished because she can only "fit it in here and there." Further, she understands the rules of accountability. Its structures serve to deskill her (Apple & Jungck, 1990). She likely knows that if she follows the BCPS rules, the county has a harder time making the case that she is responsible for her students' failure on the HSA. She cannot defend herself against the onslaught of accountability. However, she may be able to pass the responsibility on to the curriculum modules. If she follows directions, she can always blame student failure on the county's protocols for teaching biology. As a result of the high-stakes, teachers voluntarily may give up their expertise in order to protect themselves from the ramifications of their students' failure. They may not want to be experts because experts are held accountable for their students' performance. Perhaps they seek to find safety hiding themselves inside automatons who simply implement the county's accountability measures.

An Isolated Community

In the case of HHS, intensification has, on one hand promoted isolation, but on the other hand, it has increased cooperation based on accountability. As I mentioned before, the biology team is generally a tightly knit group of teachers who work together for the good of the students at HHS. Nevertheless, accountability does undermine departmental unity. I already have outlined a rift caused by replacing environmental science with pre-biology. Also, competition between teachers fosters isolation. Accountability, relegating teachers to a state akin to solitary confinement, treats each teacher as a single-unit out of context. By comparing teachers across schools, accountability ignores the teachers' backgrounds, experiences, school climate and colleagues. Ms. Khana describes feeling alone:

I'm alone out there. It's just me and my students. No one [at HHS] knows the students and what they need better than me. I know what I have to teach them and I teach it. Yet I still have to do everything to make sure that I am covered.

Ms. Khana recognizes that she is alone. She ultimately is responsible for her students' performance. It seems that she is resigned to her work as a biology teacher within the climate of accountability. Ms. Khana confessed to me on several occasions that it is hard to imagine what some of her most challenging students go through outside of school. She knows that some of her students will fail, and many will struggle to pass. According to *No Child Left Behind*, Ms. Khana is solely responsible for her students, and she may be scared, so she protects herself behind the rules of accountability. If she follows the rules, they may become responsible for her students' failure and success.

While, on the one hand, defining teachers as single, isolated units, accountability creates a sense of community along two lines: enforcement of the accountability

mandates and a teacher support group. In terms of enforcement, Ms. Victoria is at the helm of a constant dialogue between teachers. HHS has developed an electronic system of communication between her and all the teachers, who are required to submit electronically their students' attendance and grades weekly. Ms. Victoria regularly accesses the teachers' data to ensure that they are in compliance with the accountability rules and standardized curriculum. Ms. Victoria gets frustrated if teachers are late in submitting scores, if they omit assignments, or if they go off the predetermined curricular track. For instance, Ms. Victoria is very uncomfortable with Ms. Lydia moving the biotechnology lesson to follow genetics, rather than preceding it, and she frequently complains that she has too many meetings with teachers because they do not properly use the database.

On the other hand, the HHS biology and other science teachers have created a community to support one another. The clearest example deals with discipline. Dr. Stevens is the first to address this issue:

As a new teacher I am starting to learn. The others told me. We deal with discipline in-house. I get the feeling that the administration doesn't really like to deal with classroom problems. They just want the students to attend class. I don't want to go to my first evaluation and have them tell me that I can't manage the kids. So Ms. Calypso and I just try to deal with the discipline here.

Perhaps Dr. Stevens is right, and the administration is concerned with attendance over discipline because attendance is a factor in making AYP. It is possible that the administration tries to conceal problems by leaving discipline to teachers. Whatever the cause, it is obvious that biology teachers do not send students to the administration. Although I witnessed some disrespectful behavior, including cursing at a teacher and fighting, I never witnessed a teacher summon security or send a student to the office.

They always handle it with the help of their colleagues in the department. Another teacher may come in to take a walk with the student. The student may be sent to another teacher's classroom. In other words, the department creates a community of support for one another because the teachers, as a group, are isolated within the school.

Ms. Lydia clearly articulates how the department's isolation fosters an atmosphere of support between teachers:

There is more tension with the administration because they have to perform. They have to get the number. They don't treat teachers like professionals. I've seen it over and over again, especially with classroom management and new teachers. [New teachers] do not have the skills. They have to practice after coming in, but they are expected to have a veteran's approach to managing the classroom. Well, they do not have that. They are often driven out of school because they don't have it. If they enlist the help of administration, they are looked at as failures. Instead of administrative support, they get none, they are labeled, and then they are driven out. Veterans, we know better than that. Science has developed a plan because this is not the first time and the first group of administrators.

Ms. Lydia provides a vivid example of educational splintering along the lines of administration, new teachers, and veteran teachers. Although ultimately *No Child Left Behind* judges each person individually, individuals at HHS form coalitions based on their status. The veteran teachers work together to protect individual members from being labeled as a "failure," a word with many negative ramifications attached to it by the accountability movement. Thus, the label induces individual actors to protect themselves, but it also encourages groups of teachers to support one another.

The Psyche of Accountability

Entering the classroom determined to erase the body and give ourselves over more fully to the mind, we show by our beings how deeply we have accepted the assumption that passion has no place in the classroom. Repression and denial make it possible for us to forget and then desperately we seek to recover ourselves, our feelings, our passions in some private place—after class. (hooks, 1994, p.192)

According to Apple (1992) another effect of intensification is damage to teachers' pride in their own practice because they are implementing someone else's curriculum. The story at Halbert High School is more complicated than a simple loss of pride. Nieto (2006) finds that teachers who teach socially marginalized students most effectively, share a sense of mission, empathy for their students, the courage to challenge mainstream knowledge and conventional wisdom, improvisation, and a passion for social justice. In other words, they are both compassionate and passionate. On a daily basis, I see how some of these qualities are being slowly eroded by the mandates of accountability and how deeply accountability tentacles have penetrated into the psyche of biology teachers at HHS.

Frustration

First, it is important to note that with the exception of Ms. Lydia, no teachers admitted that *No Child Left Behind* and high-stakes testing have influenced their passion for teaching and for their students. Passion is integral to teachers' work. Teachers define themselves by their passion and desire to help students, and they use it to surpass the challenges that they face in their work (Fischman, DiBara, & Gardner, 2006). Consequently, it may be difficult for teachers to admit that their passion for teaching has been compromised, especially if they buy-into components of accountability. Thus, it is quite likely that Ms. Lydia admitted a diminished passion for teaching because she is whole heartedly critical of *No Child Left Behind*.

On the last day before her leave of absence for the remainder of the school year, Ms. Lydia describes the climate at HHS as frustrating:

It's a very frustrating time for teachers. I never thought that I would see it like this, but I did.

Ms. Lydia ardently denies any positive influence that *No Child Left Behind* may have on teaching, and then she goes on to describe, in greater detail, how the policy has hindered her interest in teaching:

No, *No Child Left Behind* has not eliminated my interest in teaching science. It has stymied my interest. It's put a lid on my creativity, I believe. I don't want to have anything to do with it. I don't like having to define my children and myself by numbers. We are more than just data. I want to look at my profession through a bigger lens. What is the climate? How do I address differences? How can I connect with students? How can I improve their confidence and self-esteem? That is what I want my job to be about.

The frustration oozes out of Ms. Lydia's words. She likely is alluding to the limited definition of teaching and learning that is embedded in the structures of accountability. As a veteran teacher, she may be frustrated by being judged based on numerical inputs into a database. She understands that a lived curriculum cannot be captured meaningfully in summation by attendance and assessment data points.

Interestingly, Ms. Lydia, like all six teachers, never expresses that she has had a problem with accountability in principle. The teachers' welcoming response to this research may be an indication, in and of itself, of a culture of transparency within the biology team at HHS. The teachers are more than willing to open their classrooms and teaching practice to parents, administrators, and colleagues. This year, Ms. Calypso mentored Irina, a special education paraprofessional student. Last year, Dr. Stevens did his student teaching with Ms. Lydia. Throughout my period of observations, each teacher whom I observed had at least one visitor in the classroom. The veteran teachers are all confident in their professionalism, and the new teachers acknowledge their inexperience

and welcome advice and guidance from their veteran peers. The teachers are open, even eager for meaningful reflection and growth, and they would welcome more involvement, support, and constructive criticism from the administration.

What frustrates the teachers about the current manifestation of accountability is that it de-professionalizes teachers and their work. The teachers are willing to be evaluated, but they want the evaluation to be meaningful and reflective of their practice. They want their work to be duly valued and respected. Ms. Lydia describes how she feels about the way that accountability views teachers:

[Accountability structures] don't trust anybody. They don't trust the educators. I mean we're like the slimy salesmen or something. We're not professionals, and if they don't view us that way, then who does? Where's the respect? Who respects the U.S. teacher? The rest of the world respects their teachers, but we don't. Therefore, the kids don't.

It is difficult to imagine job satisfaction under these conditions. Apple (1992) argues that accountability threatens teachers' satisfaction because they are disconnected from the curriculum. Although he may be correct, his rationale is likely to be secondary in its deleterious influence on teacher satisfaction. Pride is difficult to muster when teachers, at least structurally, are viewed as implementers, often poor implementers. A teacher who creates knowledge and understanding in the curriculum commands respect. A transmitter of knowledge who is measured by her ability to avoid interfering with the smooth transmission of standards to students is more difficult to respect. It seems impossible for teachers be passionate about their work without being respected for it.

Passion and Teaching

Not only do teachers not acknowledge that *No Child Left Behind* has influenced their passion for teaching, they actually respond definitively and succinctly to the

contrary. As soon as I ask whether *No Child Left Behind* has influenced her passion for teaching, Ms. Calypso responds, “No, not at all.” Ms. Harris, Dr. Stevens, and Ms. Khana echo Ms. Calypso verbatim. Ms. Victoria chimes in with, “No! I’m still very passionate.” Her tone berates me for even asking the question. At first, the teachers’ responses did not make sense to me. I asked this question of the teachers during our final interview. By then, we had already had one interview. I had observed them for at least fifteen classes. I had numerous informal conversations with all of them, and every teacher including Ms. Victoria, the staunchest supporter of accountability, had offered critical insights into how the policy’s implementation has made their job more prescriptive and mundane. Further, they had criticized many aspects of the policy and/or its implementation. Nevertheless, when I ask teachers whether *No Child Left Behind* has affected their passion, the type of open-ended question that normally elicits a reflective response, the teachers immediately shut down and refuse to consider such a possibility.

Maslow (1943) posits that people are motivated by a certain order of needs. Our needs at the bottom of the pyramid, such as for food and safety, must be satisfied before we can pursue the higher needs of love, self-esteem, and self-actualization. Both physical and emotional needs, however, are not linear. For example, although hunger may drive a person without food more than love, hunger does not make the need for love go away. In other words, we have many simultaneous needs that define who we are. While we may focus on one at any particular time, they all have to be nurtured or else we lose part of our selves.

Passion is part of teaching, and teaching embodies passion. Throughout their careers, teachers struggle to maintain their passion and enflame new passions. Passion

contributes to a successful and long teaching career (Intrator & Kunzam, 2006). Whether it is a passion for helping students, opening young minds, social justice, fairness, arts, sciences, or empowerment, all teachers must have passion. What happens when a teacher loses passion? The teacher loses energy, drive, focus, and most importantly, compassion. A teacher can be overworked or disrespected, but if a teacher loses her passion, she loses her purpose. Ms. Harris captures the importance of passion:

Of course, I still love it. It's still being with the kids. It's still teaching. I am making a difference in their lives. It's just hard.

Despite teaching a subject that she does not know as a substitute teacher, Ms. Harris captures the importance of her work. She reveals her love of the students, which was evident from my observations. She loves them for who they are, not who the accountability system says they are. Education policy makers and implementers often overlook the importance of teachers' passion and purpose in educational change (Hargreaves & Fullan, 1998). Like hunger, passion cannot be ignored, but accountability policies actually may go beyond simply ignoring teachers' passion. They often undermine it, and a system that erodes passion, rather than fostering it, fundamentally is flawed because it undermines the work of teachers.

Five out of six teachers did not seriously consider whether accountability affects their passion because the proposition of losing their passion understandably makes them uncomfortable. In the words of Ms. Lydia:

As a doctoral student, I have seen and I follow the policy, and what researchers are saying. And that gives me comfort. Without it, I wouldn't be sitting here now. I don't know how the other teachers do it. What gives them comfort?

Ms. Lydia sees how accountability undermines her work and being. She finds comfort in her knowledge of what she calls the “big picture,” but what about other teachers? How do they maintain their purpose? They have to protect their purpose and passion. They must hold on to them in order to have “comfort.”

That may be why the rest of the teachers could not admit that their purpose is being compromised by accountability policies. They need their passion to do their work. Perhaps nothing can take teachers’ passion away, but accountability structures attempt to undermine it. This realization makes Ms. Lydia’s candor tremendously powerful because she is willing to consider one of the darkest corners of accountability’s affect on teaching. *No Child Left Behind* inadvertently attempts to undermine teachers at their very core, separating them from their purpose. On the day before her leave of absence, Ms. Lydia exudes frustration. Part of her may have been beaten down by teaching biology at HHS. As the greatest critic of accountability on the biology team at HHS, her anger allows her to criticize *No Child Left Behind* at the deepest personal level, her purpose for teaching. On the other hand, her criticism may attenuate her ability to inoculate herself from *No Child Left Behind*’s deleterious effects on her psyche.

Dividing Teachers

Despite their denial, *No Child Left Behind* likely has influenced teachers’ passion for their profession, at least to some degree. The teachers’ silence creates an opportunity to examine the significance of their silence, but it also poses the challenge of having to read between the lines in order to gain an understanding of some of the ways in which accountability mandates and structures have influenced teachers’ passion. First, it is important to note that not one teacher considers the possibility that *No Child Left Behind*

has inflamed their passion for teaching. Even Ms. Victoria, an outspoken supporter of *No Child Left Behind*, quickly defends her passion, and then changes the subject. Why? If *No Child Left Behind* does enhance the quality of education, wouldn't teachers' interest and motivation likely be stimulated by it, at least to some degree? The teachers' silence on this matter reveals volumes about how they view accountability and the role it plays in forming their professional identity.

During my observations, I see extreme passion and compassion from all the teachers. I witness expressions of passion ranging from a gentle pat on the shoulder for moral support during a test, to an emotional conversation about a student's identity in school, to genuine empathy for student stress in the classroom. I also observe the "teacher-as-thing" outlining all the facts that students need to know for the next test, announcing the correct answers to an assignment that would have been much more meaningful to students if it had remained open ended, espousing the importance of proper form for answering the BCRs, and implementing other accountability mandates that leave no space for passion or creativity.

Perhaps teachers are able to manage living in this world of tensions, but at the very least, teachers likely have to struggle to protect themselves from the contradictions. During our conversations about accountability and *No Child Left Behind*, teachers were unwilling to take ownership of their own words. In other words, whenever they discuss the effects of accountability on the classroom, they always talk about another teacher or usually teachers, in general. For example, when I ask Ms. Harris whether *No Child Left Behind* has had any effect on her, she responds, "Personally? Me? No, no, I don't think so." As soon as she utters these words, she goes on to describe how last semester she

dealt with chronic absenteeism from students who had failed the HSA and needed to be retested. She describes the effect on her classroom:

They were constantly being pulled from my class, two at a time. It was non stop. Nearly every day, students were missing. They were some of my most struggling students. They needed to be there. I had to just keep covering the same material over and over so no one would fall behind...It was really boring.

The complexities that Ms. Harris negotiates are profound. In the same breath, she says that *No Child Left Behind* has not influenced her, and then she immediately describes how she has struggled with its ramifications in her classroom. The example she provides is a specific consequence of the strain that testing puts on the curriculum. Ms. Harris seems to know what it means to be a good teacher. She likely wants to protect her students from any negative ramifications of high-stakes testing, but she cannot. The HSA impedes upon her curricular intentions. Simply put, over-testing takes time away from teaching. Ms. Harris has to forgo her understanding of best practices in order to accommodate the mandates of accountability. In fact, because she was referring to her ESOL classes, the HSA that her students had failed was English, so Ms. Harris offers an example of how the English HSA affects the science curriculum.

The current manifestation of accountability appears hard pressed to get teachers to love their work, feel compassion for students, and be passionate about their purpose. *No Child Left Behind* is about oversight, management, testing, and data. These are not “real” components of teaching to teachers. Evaluating teachers on the basis of a singular summative measure lacks pedagogical sensitivity. Accountability does not have to be manifested this way. Accountability can be holistic, meaningful, and pedagogically sound, but current structures often ignore teachers’ individuality and humanity. Simpson

describes educators who are “alienated from educational systems that preferred technology to humanity” (Berman, Hultgren, Lee, Rivkin, & Roderick, 1991, p. vii). As accountability is currently constructed, teachers are being alienated from an educational system that prefers testing and data to a meaningful effort at improving teaching and learning.

Summary

The veteran biology teachers at HHS say that they have been feeling the pinch of accountability since the new curriculum and the HSA was introduced in 1999-2000. Currently, the grip of accountability on the biology curriculum creates enormous pressure on teachers. Its present implementation mistrusts the professional qualifications of teachers. It monitors them to ensure compliance, with a narrow and rigid understanding of teaching and learning. The structures put in place by accountability policies place tremendous pressure on teachers in and out of the classroom. Despite being adept at managing multiple competing demands on their time, as products of their environment, teachers internalize the pressures, which in turn, have a profound influence on the biology curriculum.

Further, as students, beginning with the class of 2008-2009, do not graduate because they failed the biology HSA, the pressures of accountability on teachers are likely to continue to mount. It is possible that students will receive a reprieve from the graduation requirement. MSDE recently began working on a “Bridge Plan” that involves locally administered projects developed for students who have not passed the HSAs by their senior year (MSDE, 2007a). Nevertheless, the accountability climate generally seems to be expanding. Ms. Lydia describes the ongoing encroachment of accountability:

All states have to have proficiency by, what is it, 2014. You can see the pressure running up to that. All schools want to do is pass the test. Actually, they really just want to be left alone. But they have to play. But the rules are impossible. Meanwhile, they have to realize that they cannot meet [100% proficiency]. It may take a while as the bureaucracy churns, but they will realize it...So, for now, I'm under the pressure of having to meet something that can never be met, as a teacher.

Ms. Lydia plainly explains that, in the near term, the pressure of accountability on teachers will only continue to mount. As the mandates of accountability move forward, they are bound to collapse under their own impossible proposition, that “no child is left behind” by 2014. In the meantime, however, it is important to understand how the pressure actually undermines the work of teachers, both “as-beings” and “as-things.” As accountability policies strive to regulate teachers, they do so by applying pressure, both intended and unintended, that manipulates the work of teachers. The pressure creates a place in-between where teachers would choose to be and where accountability wants them to be. In the final theme chapter, I turn to an examination of how teachers engage in the tensions of being-in-between competing paradigms of their work.

CHAPTER VII: TEACHER-IN-BETWEEN

...splitheads have glimpses or peeps into another side of themselves because they feel split. As such, they are dimly aware that they may be living according to others' expectations as opposed to living from their core. Social habits prevent psychic movement. These characters allow their social selves to take the place of their authentic selves, until the difference between the two is blurred. What is reality, what is illusion; what is the true self, what is the false? Such a person is playing a part in a script written by others. (Doll, 2000, pp.82-83)

In this chapter, I turn to a critical exploration of how teachers negotiate through the tensions that I unearth in the previous two chapters. Teachers in the current accountability system receive countless mixed messages that call them in multiple directions. Do teachers teach to standards or to the high-stakes test? Do teachers decide to vary their pedagogy at the possible expense of leaving their students less prepared for the HSA? Do teachers reconcile their own understanding of quality teaching with the accountability mandates, or do they protest in silence? The myriad demands on teachers can be frustrating and confusing. Ms. Harris describes her difficulty traversing through the different and competing signals:

Look, I'm a new teacher. I'm not even certified in science. I don't know what I'm doing, but I get told so many different things. We need to teach special ed kids differently, but we have to cover biology, so just look for worksheets. We don't want to bore kids who have already taken biology by teaching it twice, but teach biology because we need to prepare them for the HSA. I mean, it's crazy!

Ms. Harris brings up two powerful tensions that confuse her. She seems to be unable to reconcile the competing sets of signals that she receives logically, yet she comes to the classroom and teaches. Although she sees contradictions, she must get around them in order to practice her profession. Being able to live in these tensions has become part of Ms. Harris' job in the current educational climate.

Ms. Harris provides two examples that may only apply to her, but, through this research, I have discovered that every biology teacher at HHS sees serious tensions embedded in accountability. They receive and reconcile myriad messages in and out of the classroom. For example, each teacher in this study mentions having to compromise on the number and quality of labs, creativity in their teaching, and diverse pedagogy in order to cover all the material that students need for the HSA and the County Final Exam. How do teachers make such compromises? How can they forgo their own expertise and understanding of quality teaching? What are the underlying reasons and mechanisms behind teachers' negotiations through the accountability system?

Teachers must reconcile competing signals by developing curricular spaces in-between the tensions. Ms. Khana alludes to such a space when I ask her about tensions in the curriculum:

I like the county curriculum. It's not perfect but I like it. It covers the material, makes students think. It's got some faults too. I teach it 'cause I'm supposed to, but I like it too. Sometimes I add my own stuff to it, when I need to show the kids something.

Ms. Khana compromises between her own understandings of the curriculum and the county's mandates. She generally follows the curriculum quite literally, but occasionally she fills in what she considers to be a crack. For example, when she introduces the unit on evolution, Ms. Khana tells her students:

You have to know Charles Darwin. Charles Darwin isn't even on the test [the HSA], but as far as I am concerned, when you talk about evolution, you need to know Charles Darwin.

Although a minor example in terms of scientific significance, Ms. Khana's words clearly offer an example of what she considers to be non-negotiable. Every teacher likely has aspects of the curriculum that are not to be compromised. Otherwise, they are able to find

a space in-between their own understandings of teaching and the demands of high-stakes testing.

Pinar (2005), referencing Aoki, refers to this type of reconciliation as in-dwelling. He uses in-dwelling to explore how teachers negotiate between curriculum-as-plan and curriculum-as-lived experience. In this chapter, I examine in-dwelling between accountability signals and teachers' own constructions of good teaching. I explore how teachers create in-between curriculum places and how those places manifest themselves in the curriculum. Specifically, I uncover how teachers negotiate and understand the terrain of morality, expertise, and pedagogy.

In-Between Values

No Child Left Behind appears to regard education as a technical endeavor of transmitting predetermined knowledge and skills to students. For better or worse, teachers are the means through which the knowledge gets transferred. To the chagrin of accountability structures, classrooms contain more than knowledge. Teaching involves a multiplicity of meanings that teach students knowledge and skills, but also morals and values (Oser, 1992). Despite often being excluded from the curriculum conversation, this paradigm of teaching that allots equal weight to both the practical and the moral dimensions of the craft, opens the curriculum to numerous possibilities, as well as pitfalls. Teaching is a moral practice and teachers are moral actors (Pring, 2001). In the world of the classroom, innumerable moral interactions and exchanges occur between teachers and students. Some meanings are obvious, such as a teacher establishing particular codes of conduct. Others are more subtle, evidenced by tone of voice or a

particular curricular focus. Regardless of how messages are relayed, the moral standing and foundations of teachers are translated into classroom values.

While *No Child Left Behind* formally may ignore moral dimensions of the curriculum, its measures and structures send moral messages to and through teachers. Accountability influences teachers' behaviors and perspectives, which have ramifications in the moral arena of the curriculum. In other words, the decisions that teachers make have a moral dimension. For example, at Halbert High School, environmental science teachers accept teaching a course without a written curriculum. Biology teachers know that their students will be taking an exam at the end of the course that will determine whether they graduate from high-school, and they know that they will be compared with colleagues based solely on their students' test scores. These manifestations of accountability policies carry serious moral curricular outcomes. In this section, I examine how teachers traverse the moral terrain of accountability. Specifically, I focus on two ways that teachers deal with the moral implications of accountability: avoidance and submission.

Avoidance

Throughout my time at Halbert High School, I frequently witness teachers be compelled to avoid or deny the moral implications of *No Child Left Behind*. One of the clearest forms of denial comes from Dr. Stevens and Ms. Harris, who use their status as new teachers to avoid a critical examination of accountability and its effects on the curriculum. When I ask Dr. Stevens to discuss his view of accountability and *No Child Left Behind*, he responds:

I have nothing to say about it. It's how it was when I came in, and I don't know how it was before... What can I say? I don't really have an opinion about it. What can I compare it to?

In order to probe further to see if I can get Dr. Stevens to consider accountability from a different perspective, I ask him if it has had an influence on his teaching. He responds with, "Not really." Then, I ask if he is cognizant of it in his work, and he responds, "Not so much."

One explanation for Dr. Stevens' responses to my probes is that accountability structures have forced him into denial because the current manifestation of accountability has certainly influenced his work. It seems curious that he has "nothing to say" and doesn't "have an opinion" about the landmark educational policy of the current administration. He is teaching a class created to help struggling students pass the HSA. As a retired chiropractor, he is prohibited from teaching biology because he is not "highly-qualified" in the subject. His teaching is monitored to ensure appropriate alignment with the biology curriculum. Most importantly, accountability has challenged his very understanding of teaching. Dr. Stevens says:

Teaching, I always thought, more or less, you teach them how to think and how to figure things out, not necessarily to memorize facts. But, if they need to find what the fact is, they know where to go get it. They know how to figure it out. Whereas, with *No Child Left Behind*, I think they might be geared more towards just facts. If you know this, you answer the question this way, you are fine. You don't have to understand it, but if you answer it and get it right, you pass. Originally, coming in, I thought teaching would be a lot of teaching them how to think, and at times now you find out that it's teaching them facts, so they could know this because that's what they have to know.

Essentially, accountability has fundamentally transformed the focus of Dr. Steven's work. He is less concerned with teaching students how to think and more focused on transmitting predetermined facts to students. Nevertheless, Dr. Stevens claims to not have

an opinion of the legislation or accountability in general. What stands behind this denial? Why is Dr. Stevens reticent to critically examine *No Child Left Behind* and its accountability structures?

Ms. Harris also seems to avoid a critical examination of the implications of *No Child Left Behind*. When I ask her about possible tensions between accountability and her teaching, she responds:

So far, not really. Not that I'm aware of. I mean it's my first year teaching. I don't really know. Ms. Calypso begged me to teach.

Perhaps, she lacks some confidence in her own expertise, but Ms. Harris does have a strong passion for teaching. She has known that she would be a teacher since childhood. She has expressed anger about the HSA and the overt focus on biology in environmental science. Nevertheless, Ms. Harris fails to make an overt connection between accountability mandates and their manifestations in the curriculum. She claims to have no opinion of accountability policies, including the HSA. There seems to be a disconnect between the effects of accountability and how new teachers speak about the effects of accountability. In many cases, teachers seem to deny or avoid considering the influence of accountability policies on their own teaching. Why?

Why? How can Dr. Stevens and Ms. Harris possibly claim to have no opinion of something that has so profoundly affected their work? A variety of reasons may contribute to the teachers' reticence to share their opinions about *No Child Left Behind*. However, Dr. Stevens' words may provide insight into his thinking. He says about accountability, "It's how it was when I came in, and I don't know how it was before." These words are analogous to Billy Joel's song "We Didn't Start the Fire." In the chorus, Joel sings "We didn't start the fire. It was always burning, since the world's been turning.

No we didn't light it but we tried to fight it." Joel lists the twentieth century's momentous global events, suggesting that we can not be held responsible for the current status quo because the world has always been in a troubled state. The song's lyrics can be interpreted in two ways. Either an individual can avoid responsibility for the status quo by holding his predecessors accountable, or the status quo is normal because today's circumstances are the same as they always have been.

In the case of current accountability practices at HHS, both perspectives may be viable. On the one hand, teachers, especially new ones, can avoid responsibility for the degenerative outcomes of accountability. As novice teachers, they can claim ignorance or innocence. After all, they had nothing to do with the enactment of accountability policies. They were not even in the profession when *No Child Left Behind* was passed. According to this logic, they cannot be held morally responsible for the deleterious effects of accountability structures on teaching and learning. From a moral perspective, teachers can avoid feeling guilty for tricking students into taking biology two times in a row, attempting to jam discreet facts into students' minds, and abandoning many of their best practices. Perhaps, most importantly, they can avoid feeling responsible for their students' failure on the HSA. McNeil (2000) found that teachers within accountability systems that include a penalty for students, like failure to graduate, first and foremost want their students to graduate. For the most part, teachers are less concerned with being censured than they are about punitive outcomes for their students. Biology teachers at HHS are no different. All six of them are concerned about their students' graduation. Thus, while they do everything they can in the classroom to help them pass, they may be protecting themselves from guilt by rationalizing their own culpability in the process. As

Ms. Lydia points out, “No one seems to be resisting *NCLB*,” so denial may be another way of avoiding moral responsibility for some of its negative effects on students.

The other interpretation, that the status quo is normal because today’s educational climate is the same as it has always been, also may allow teachers to avoid the moral backlash of accountability by normalizing its presence in the educational system. In other words, if accountability is a normal, integral part of education, there is nothing a teacher can or should do to avoid it. It is a regular part of the curriculum to which teachers must ascribe because it is an integral part of their profession. Teachers particularly may be susceptible to this perspective because accountability appears to be omnipresent. BCPS and HHS seem to be saturated with accountability policies and mandates, and no single individual is responsible for originating accountability programs. Thus, if the teachers can accept that accountability structures are fundamental parts of the educational system, they can morally justify their actions and inactions because there is no other way to teach. In the words of Dr. Stevens, “That’s how it was when I started teaching. I don’t know anything else.”

Because they are new teachers, Ms. Harris and Dr. Stevens particularly may be susceptible to denying their responsibility for the outcomes of *No Child Left Behind*. Nevertheless, veteran teachers also exhibit signs of denial. At times, Ms. Khana denies the very premise of accountability, that teachers are responsible for their students’ performance. Ms. Khana explains one of the reasons that she transferred to HHS from another school in the district:

In my case, it really hasn’t affected my teaching because I’m pretty much going to do my best one way or another...Let me give you an example. I had such a high failure rate last year in my biology classes due to attendance. Rather than looking at the attendance aspect of it, the tables

were turned and they said, “What are you going to do about it?” I’m the one who created the problem, see. That’s when it could become a major problem.

Ms. Khana avoids responsibility for her students’ failure. She attributes their performance to a difficult home life, poor reading skills, jobs, familial responsibility, and lack of motivation. In this description and all of my conversations with her, Ms. Khana never considers her own culpability in her students’ performance. Her position can be summarized with her own words:

Even the slow readers can pass. I teach them everything they need to pass the HSA and the County Exam. It’s just a matter of whether they do the work.

Ms. Khana seems to suggest that the teachers and students both have a well defined role. Teachers provide knowledge and skills to students using a variety of methods, and they evaluate whether students have received and retained that knowledge. Students use a variety of techniques – like listening, studying, and reading – to acquire the information, and they develop the skills necessary to demonstrate their knowledge on instruments provided to them – like worksheets, homework, quizzes, and tests. Although these particular roles can and should be ascribed to teachers and students, they do not capture the entire act of teaching and learning. They reduce teachers and students to independent individual units with a limited relationship, defined boundaries, and little obligation to one another.

Ms. Khana’s defined roles for teachers and students ignore the complexity and humanity of the classroom. Teaching is boundless and interactive. Ellsworth (1997) calls teaching a transitional state in which “boundaries between choice and closure, responsibility and manipulation, self and other” are blurred (p. 167). Ironically, Ms.

Khana's proposed vision of teaching violates both *No Child Left Behind's* and Ellsworth's understandings of teaching. Ms. Khana's view seems to contradict *No Child Left Behind's* premise that teachers and schools are responsible entirely for their students' learning, ignoring students' individuality and life circumstances. On the other hand, her view, and the view of accountability policies, is significantly more limited than Ellsworth's description of teaching because it ignores the lived interactions that constantly occur in classrooms.

Thus, Ms. Khana avoids the premise of accountability, but her avoidance may have serious ramifications, moral and otherwise, because she limits her own understanding of her work and responsibilities. My observations of Ms. Khana's classes support this interpretation. On one hand, she preaches diligence and responsibility to students, and her pedagogy is teacher-centered, where knowledge travels in one direction from teacher to student. On the other hand, more than any of the other teachers, she alters her teaching to meet the demands of accountability. She uses the county curriculum, almost exclusively, despite only being mandated to administer pre-assessments and post-assessments for each chapter; she includes practice test questions and HSA preparation on a daily basis; and she constantly refers to biology content in terms of its likelihood of appearing on the HSA.

At the same time, Ms. Khana denies that accountability has affected her teaching. Ironically, her denial demonstrates just how deeply she has been affected by accountability. Apple (1992) suggests that the intensity and pervasiveness of accountability cause it to become normalized to teachers' work. The normalization of accountability may be an explanation for Ms. Khana's perspective. She tells me that

accountability “really hasn’t affected my teaching because I’m pretty much going to do my best one way or another.” She propagates the myth that accountability will motivate teachers to work harder. As I discussed in Chapter Six, passion is much more likely to drive teachers in their work than the accountability movement. Ms. Khana’s narrow definition of accountability allows her to avoid a critical reflection of its influence on her classroom. She tries just as hard as she would without accountability, but what is she trying hard to do? She whittles down biology knowledge and skills into simple and discreet concepts. She treats students as receptacles of test taking strategies. She limits her pedagogy and eliminates labs from her classes. These practices have become normal to her. They are her work. Ms. Khana may have found a curricular place where she can avoid a critical analysis of how *No Child Left Behind* has influenced her understanding of teaching and learning. In fact, the normalization of accountability, may even allow Ms. Khana to deny that accountability has influenced her teaching.

Submission

In addition to allowing teachers to avoid moral responsibility for their actions as implementers of accountability, the normalization of high-stakes testing may be indicative of teachers’ submission to the almighty status quo. In other words, if teachers submit to the ethos of accountability and the practices associated with it, their pedagogical actions can be judged by a whole new set of standards. Rather than defining their work through human interaction, creation of knowledge, and critical inquiry, teachers can view their practice through the lens of accountability, where teachers’ work is defined by statistics, like students’ test scores and attendance. In terms of responsibility to students, instead of being responsible for facilitating an awareness of self, critical

thinking skills, and an integrated understanding of biology, teachers are responsible for getting students to attend class and pass the biology HSA and county exam.

The differences between the teachers' personal sense of responsibility and responsibility within the context of accountability are profound. Although all six biology teachers, including Ms. Lydia, submit to accountability to some degree, Ms. Victoria offers the clearest and most informative example of submission. She says that her beliefs about teaching align with accountability:

I kind of like the *No Child Left Behind Act*. I like that teachers are accountable. I like that you have to raise your standards as a teacher because this test is going to reflect how well you are going to provide instruction. So you want to do a good job, and if kids don't get it, this is an incentive for the teacher too, to help them get it.

No Child Left Behind has captured Ms. Victoria's educational focus. With this statement, she seems to uncritically submit to accountability. Although she finds the policy to be motivational, a clear benefit to education, she does not consider that certain forms of motivation may be more beneficial to students because some goals may be more worthwhile. How does *No Child Left Behind* motivate teachers and to what ends?

Ms. Victoria expects that high-stakes testing will encourage teachers to raise their own "standards." Here, Ms. Victoria is ambiguous about which and what standards. The word "standard" has military etymological roots, originating from the Germanic *estandard*, a rallying place (Harper, 2001). Currently, one of its meanings, presumably the one referenced by Ms. Victoria and *No Child Left Behind*, is "an accepted measure of comparison for quantitative or qualitative measure" (Berube, 1995, p. 1075). Ironically, both of these definitions inform the structures of accountability and Ms. Victoria's submission to them.

In general, *No Child Left Behind* subscribes to the paradigm of quantitative measure, and teachers are, in fact, evaluated according to a particular standard, the percentage of the teacher's students who pass the HSA. Further, teachers are compared to one another on the basis of the established standard. Unlike some comparisons, *No Child Left Behind* does not explicitly place teachers into a zero-sum game that must have winners and losers. Theoretically, all teachers can have all of their students pass the HSA (as *No Child Left Behind* prescribes for the year 2014). Nevertheless, teachers recognize that, in the words of Ms. Lydia, "[They] are placed in an impossible predicament." All students cannot and will not pass the HSA. Therefore, the teachers are in a zero-sum situation, where ultimately, the best teachers, as defined by the accountability standard, will be those who are teaching the best students. Some teachers will inevitably be teaching failing students. In a bizarre twist, rather than quality teachers producing the best students, the best students produce quality teachers.

Ms. Victoria, however, does not talk about any standards; she uses the words "your standards," which seemingly suggests that teachers should raise their own standards. Unless a teacher has submitted to the principles of accountability completely, she will not define her work solely by students' HSA scores. Although all of the biology teachers at HHS do include, at least to some degree, preparing students for the HSA as part of their job, they all have standards beyond that for themselves and for their students. Further, the motivation for the standards is mainly internal, rather than external. For example, in terms of the HSA, all six teachers express deep concern that their students may fail to graduate if they fail the high-stakes test. The apprehension stems from their

own conscience and empathy for their students. Ms. Calypso's greatest concern about the HSA is

self-perception. Tests don't define who we are or what we do. Many bright students don't pass exams, but have skills that will help them be successful in many areas. But these four tests really dictate how they perceive themselves and their future.

Ms. Calypso, and the other teachers, also may be worried about how they look if students fail or how their students' failure may affect their professional career. However, in all of our conversations and my observations, their greatest concern is always the effect that the HSA may have on their students' lives. Ms. Lydia even expresses "surprise" at how much the prospect of student failure bothers her.

Using the term "standard," Ms. Victoria may be referring to *No Child Left Behind* as a rallying call for teachers. The problem with this way of thinking is that the teachers do not primarily define themselves as they are defined by accountability principles. They are motivated by a desire to teach and to improve the lives of their students. HSA passing rates do not reflect their personal standards for success. However, if teachers submit to the principles of accountability, like Ms. Victoria, the principles of *No Child Left Behind* can provoke them to work harder. In fact, teachers who buy-in to policies are more likely to be more motivated in their work (Elmore & Sykes, 1992). Buying-in is convenient, a path of least resistance, for teachers in a climate saturated with accountability and high-stake repercussions for noncompliance. For instance, it may be expedient to define success according to HSA scores if HSA scores monopolize the discourse around student achievement. Redefining success, however, translates into teaching practice, and teachers' work also becomes redefined. If teachers submit to accountability and are motivated by its principles, they are likely to pursue its outcomes. In other words,

accountability reinforces itself. As teachers submit to the policy because of its powerful position in and above the curriculum, they are likely to further its principles.

As an example, in her description Ms. Victoria adopts a fundamental, yet highly controversial, tenet of accountability, that the HSA, or another state's high-stakes summative assessment, reflects quality instruction. The tenet has two components that must be true for the entire proposition to be true. First, the HSA is an authentic and legitimate evaluation of students' work. Second, teacher quality can be evaluated legitimately according to students' performance on a test. In Maryland, AYP is measured by two factors, attendance and MSA or HSA test scores. Districts, schools, and teachers are evaluated by these two criteria. They are used to justify sweeping curricular changes, such as the restructuring of an entire course. A realization that the prevalence of accountability is promulgated by false pretences would undermine much of the current work of schools and teachers, which reflects current accountability mandates. In other words, a critical perspective regarding the value of the HSA would render much of HHS biology team's current efforts meaningless, and even possibly immoral. Instead, Ms. Victoria automatically propagates what is perhaps the most controversial assumption behind *No Child Left Behind*, that a summative standardized test adequately reflects quality teaching.

In her submission, Ms. Victoria even adopts the technical language of accountability, beyond simply discussing teacher standards. She says that the HSA measures how well teachers "provide instruction," as if instruction can be scooped up and handed over to students. Instruction is not matter, nor is it finite. It is a process, the exchange between teachers and students. The accountability movement, however,

considers knowledge to be a commodity. DeLossovoy and McLaren (2003) argue that “Trends in accountability reify the consciousness and creativity of students into simple scores and indices according to the logic of commodification” (p. 131). Like economists who categorize society, including people, according to monetary value, *No Child Left Behind* expects educators to characterize students and teachers according to scores and indices. Like money, the test scores become a universal equivalent that can be manipulated and compared across any and all environments.

As an administrator, high-stakes accountability compels Ms. Victoria to reduce students and teachers into units that can be readily manipulated. When I ask Ms. Victoria whether *No Child Left Behind* has influenced her work with students, she immediately responds according to the categories established by the legislation:

Oh yes! We do Specific Projected Intervention Monitoring, where each kid’s grades and objectives are monitored from the first day of school till the HSA. So if they don’t master an objective, we have to re-teach it. Intervention strategies range from before-school tutoring, after-school tutoring, lunch tutoring, peer work... Well, it turns out when you do the spreadsheet by race, gender, FARMs. You look at all that data and most of the kids that need a lot of interventions are Hispanic males, followed by African Americans.

In her work supervising teachers and students, Ms. Victoria commodifies her students and teachers’ work according to the language of *No Child Left Behind*, which needs to be deconstructed in order to understand its influence on the curriculum. Ellsworth (1997) suggests that the “work of deconstruction is to trouble every definition of teaching and studenting” (p. 140). What does the language of accountability say about the learning process? Who does it think that teachers and students are?

First and foremost, Ms. Victoria names her work, and it is not called teaching or even supervising. Ms. Victoria’s work is called “Specific Projected Intervention

Monitoring.” In order to monitor students and ultimately evaluate teachers, students’ experiences during the school year are summarized by grades and objectives. If students fail to meet their objectives, or if teachers fail to ensure that students meet their objectives, students receive “interventions,” not extra-help or more teaching.

Accountability even compels Ms. Victoria to organize her students according to numbers on a spreadsheet. The students are reduced to their race category, and the categories are compared by numerical data.

Further, Ms. Victoria uses the data she collects to commodify and admonish teachers:

I like that teachers are held accountable. We should either help them improve or help them leave. So we hold them accountable, but nothing happens to them. Oh look, you have a 40% passing rate on the HSA, and the most you can do is say, “Okay, you aren’t going to be teaching biology anymore.” But they’re going to go teach somebody else. You can’t get rid of them. So we are holding them accountable, but there are no consequences for them. “Ooh! I had a 40% passing rate. What are you going to do about it?” Nothing.

Ms. Victoria’s submission to accountability demonstrates how *No Child Left Behind* imposes itself on teachers. As an administrator and biology leader, Ms. Victoria holds institutional power over the teachers in the department. Evaluating science teachers is part of her job. Right or wrong, she uses her definition of quality teaching to describe and evaluate teachers, who may face repercussions based on her evaluations. Thus, they have to accept her criteria and be influenced by it, at least to some degree. As such, they, themselves, must submit partially to the principles of accountability in order to fit into her description of their job.

An examination of Ms. Victoria's submission to accountability and *No Child Left Behind* provides insight into the mechanisms of accountability. Ms. Victoria does not espouse her own view of teaching and learning. She is forced to act as a bullhorn of *No Child Left Behind*. As an enforcer of the accountability system, she must propagate accountability's views of the curriculum as her own. Based on many conversations with her, I know that Ms. Victoria does not believe that teachers should be evaluated and fired wholly according to their students' test scores. As a teacher, she knows that teaching and learning cannot be reduced and manipulated as a number. Watching Ms. Victoria be-with her students demonstrates her role as a teacher, rather than a "provider of instruction," yet, for some reason, she supports and promulgates the accountability system through the HHS Science Department.

In describing how she feels about accountability, Ms. Lydia ponders why Ms. Victoria supports *No Child Left Behind* and what her submission means for the Science Department:

I don't know how protected we are. Ms. Victoria pretty much ascribes to accountability. I'm not sure why. I think she cares about the kids and this is her way of showing that they are proficient. She doesn't want them not to graduate. I think that must be the thing. It's the only thing that I can see.

Ms. Lydia's words demonstrate the threatening power of accountability. Ms. Lydia is justified in feeling unprotected. Ms. Victoria tells me that she judges teachers according to the values established by the accountability system. In fact, she is skeptical of Ms. Lydia's test preparatory abilities. The blame for Ms. Lydia's discomfort, however, does not lie with Ms. Victoria. She does not have the agency to protect or threaten the teachers. She is not in control. Accountability is the agent of authority. Ms. Victoria's role is defined by accountability. Like an actress chosen to play a part, she is simply cast

in the role of teacher/administrator in the climate of accountability. Part of her job is to evaluate teachers and organize the curriculum according to the mandates of accountability. Anyone else would do the same thing or be replaced.

Although Ms. Lydia ascribes the threat to student graduation as the mechanism of control over teachers, the influence of accountability may be significantly greater than the stakes associated with it. Accountability imposes itself upon the actors in the system of education, both structurally and emotionally. Ultimately, teachers have little choice but to submit. I presume that structural submission precedes emotional submission.

Accountability creates a rigidly defined space for teachers to do their work. It is the space predominantly occupied by teacher-as-thing, an implementer of education. Teachers are pressed into a practice that may contradict the knowledge-producing curriculum paradigm. Thus, they are given an impossible task of submitting to the technical paradigm of accountability that violates the reality of the lived curriculum. All six biology teachers at HHS find themselves negotiating this hazardous terrain.

In-Between Skills

Apple (1992) and Apple and Jungck (1990) argue that the structures of accountability serve to de-skill the biology teachers. At HHS, accountability measures have placed a tremendous burden on teachers. Teachers must spend valuable hours grading, inputting data, analyzing data, and developing interventions. Accountability creates pressure to teach to a high-stakes test, compelling the teacher to act as a test preparatory coach. Knowledge is reduced into bits, so teachers must be experts in accountability biology rather than just biology. Consequently, teachers are de-skilled and re-skilled. Their teacher-as-being skills are underutilized, while teacher-as-thing skills are

developed. In other words, teachers gain expertise in test preparation, teacher-centered pedagogy, data analysis, etc. Are these skills useful? Perhaps they should be part of teachers' practice. However, in the current accountability climate, they seem to have gained supremacy at the expense of other teaching practices. Thus, the very nature of expertise comes into question and begins to be redefined. In effect, teachers must "in-dwell" between competing demands that call for different skill sets, which has powerful ramifications on teacher practice, both in and out of the classroom.

Competition

Competition is a process or variety of habitual behavior that grows out of a habit of mind. (Willard Beecher, cited in Genn, 2007, ¶ 5)

As teachers are compelled to engage with accountability, the engagement becomes a habit of mind. Whether teachers submit to the policy or deny its effect on their practice, accountability becomes part of the process of their work. High-stakes accountability reduces biology knowledge into discreet bits of information, and it assigns numerical data driven values to actors in the educational system – districts, schools, administrators, teachers, and students. Like variables in an equation, the data and bits are manipulated to fit into an accountability plan, which imposes a numerical label that can be used as a ranking of every teacher. Unsurprisingly, I found that biology teachers compete with one another for a higher ranking. However, what I found to be a major surprise is that Ms. Lydia, who disagrees most vehemently with *No Child Left Behind*, appears to be the most competitive teacher.

Throughout my period of observations and during our conversations, teachers often cited their professional accomplishments. It is possible that they wanted to impress an outsider who will tell their story, or they simply are proud of their impressive

professional and educational accomplishments. The teachers presented a variety of accomplishments; their understanding of their practice is holistic, and they take pride in their work. Both Ms. Calypso and Ms. Victoria are seeking National Board Certification in biology. Ms. Calypso is very proud of taking forty-four environmental science students, many of whom are students with special needs, to the Annual BCPS High School Student Inquiry Conference. Ms. Victoria often cites her students' dedication and giving character. Her students participate in community service, tutor struggling peers, and raise money for school events and charities. She equates their success with her own. Ms. Lydia regularly refers to her academic and professional achievements and her knowledge of education. Ms. Khana frequently highlights her twenty years of teaching experience. These examples tell a story of who the teachers are and what they value outside the restrictions of accountability.

Many of the accomplishments that teachers cite, however, do pertain to professional success as defined by accountability. Ms. Khana, who, according to accountability standards, has a much longer road to success than her colleagues because she teaches the two general biology classes, highlights her teaching rather than her students' test scores:

None of my students should fail by rights. I have taught them everything they need to know. I know the test and I have taught them what they need to get ready.

On the other hand, Ms. Victoria, who teaches the top students at HHS, describes work in terms of students' performance on the HSA:

Oh, my kids? They should know every single question that is on this test. None of them should fail the HSA. They will all pass the test. It is easy for them. They are ready.

Even environmental science teachers, whose students do not take the HSA at the end of the course, subscribe to the hierarchy of accountability. Ms. Calypso cites her efforts at preparing students for the HSA:

The test requires reading, so I work on their literacy. The test requires graphs and charts, so I include graphs and charts. We give them BCRs to practice. They should be ready. I am worried that they still won't be able to sit still and pay attention for three hours, but they should be ready.

Whether teachers teach top students or students with special needs, and whether they teach biology or environmental science, teachers often highlight their strengths in accountability terms. Every teacher, except Ms. Harris, claims to be ready for the inclusion of biology on the HSA, and all of them claim to know how best to prepare their students for the high-stakes test.

As new teachers, Ms. Harris and Dr. Stevens are the least competitive, especially in terms of accountability. Both teachers take responsibility for some of their students' creative projects and other classroom-based accomplishments. Neither of them, however, makes any claims about students' performance on the HSA. Ms. Harris, for obvious reasons, claims no expertise in biology or the HSA. Although Dr. Stevens may possess the requisite content knowledge, he prefers to defer to his more experienced colleagues. As such, Ms. Calypso essentially supervises their practice. Perhaps Ms. Harris and Dr. Stevens feel like they cannot compete with their colleagues. Regardless of their rationale, both first year teachers prefer to deny or avoid *No Child Left Behind*, rather than compete according to its standards. Their approach may be the safer one because if they do not compete, they cannot be disappointed.

All of the above examples are instances of positive competition where teachers claim excellence without undermining the work of their colleagues. Teachers are willing

to share trade secrets. They collaborate on developing the HSA Review for Biology. They meet to discuss best practices and effective interventions for students. When problems or conflicts arise, the teachers support one another. Consequently, during my second to last observation of her class, I was quite surprised to discover that Ms. Lydia tries to give herself a competitive edge over her colleagues. She inadvertently revealed her “secret weapon against the HSA” during a casual conversation.

Ms. Lydia’s replacement for her leave comes to observe the class, and the three of us stay after to discuss the biology textbooks. As expected, all the biology classes at HHS use the same textbook, which Ms. Lydia calls the Whale Book because it has a picture of a whale on it. Ms. Lydia, however, uses two additional textbooks as supplements to the regular curriculum, the Elephant Book and a Prentice Hall Book. She frequently uses both the Whale Book and Elephant Book during class, and she has enough copies of the books for all her students. She only has one copy of the Prentice Hall book, so she makes photocopies of certain sections for her students. Until this conversation, I did not know why she used all three books. Addressing her replacement and me, Ms. Lydia explains:

The Whale Book is too superficial for the HSA and County Exam. I use the Elephant Book because it’s deeper and more thorough, but it’s too advanced for many of the students. They don’t get it unless I explain it to them. But this other book, it’s my secret weapon. Two years ago my apprentice discovered that Prentice Hall published a textbook that was being used by Maryland to write the HSA.

When I ask Ms. Lydia how her apprentice found out, she says that her apprentice heard rumors that Maryland uses the textbook. Ms. Lydia and her apprentice reviewed the book and recognized questions in the textbook that she had seen on the HSA. Now, Ms. Lydia is convinced that the state uses the book to write the biology HSA.

In light of such a profound discovery, I express surprise to Ms. Lydia that I had not previously heard of her treasure. She responds:

I try to keep it quiet. I don't share the book with the other teachers. Every little bit counts. I can use the edge. And Halbert High School doesn't have the resources to buy the book for all the classes anyway. We just bought the Whale Book. I don't want to make the other teachers feel bad.

As I listen to this explanation, I cannot help but make the comparison to professional sports. In professional sports, athletes are sometimes induced to take performance enhancement drugs in order to beat their competition. In her description, Ms. Lydia relies on a textbook that gives her an "edge" over other teachers. Ms. Lydia's retention of the Prentice Hall book for herself, offers a vivid example of how accountability promotes a ruthless winner-take-all type of competition.

The reduction of teachers and students into data mirrors the reduction of athletes into statistics in professional sports. Young (1986) argues that sports statistics serve to "colonize the beauty, elegance, joy, and despair of physical performance" (Young, 1986, p. 6). Athletes are reduced to batting averages, a free throw percentage, or 100 meter speed. Such statistics serve to individualize athletes, promote the individual over the team, and accentuate differences by making distinctions between similar performances. For example, in terms of performance, there is a negligible difference between a 9.86 second and a 9.85 second 100 meter dash, but one sprinter is a "winner" and the other is a "loser." Young uses a structural examination of sports to argue that a desire to market sports to spectators drives the need to commodify athletes for consumption.

The current manifestation of accountability seems to have a similar effect on teachers. It colonizes the beauty, elegance, joy, and despair of pedagogical performance. Teachers are evaluated according to the numbers they input into the central database, and,

even more importantly, by HSA scores and County Exam scores. Like the work of athletes, the work of teachers is reduced to securing higher scores and statistics. Because the process of teaching and learning becomes narrowly defined and reduced to a set of scale scores, teaching is localized to an individual teacher, and the differences between teachers are quantifiable. This premise contradicts the real teamwork and multiplicity of interactions involved in the teaching and learning process. Learning is not isolated or compartmentalized. In the words of Ms. Calypso:

We can accommodate difference but we cannot eliminate it. Behavior changes gradually and slowly. *NCLB* holds teachers accountable and if students fail, teachers are told that they are crummy. But growth and progress is cumulative, not instantaneous.

As with sports that are marketed to consumers, the premise behind accountability is to give outsiders the ability to compare schools. Like sports fans, the outsiders – community, politicians, media, or business – vary in their familiarity and involvement with education. After all, teachers are the experts in education like athletes are the experts at their sport. Therefore, in order to make education more accessible, it is “simplified” for its consumers with little consideration for the effects on the curriculum.

As a result, education is driven by a simplified set of measurable characteristics – HSA scores, County Exam scores, attendance, “highly-qualified teachers,” AYP, etc. Therefore, it may not be surprising that teachers compete like athletes by the standards of their spectators. However, with this rationale, I would expect that a teacher who has submitted to the policy would be the most competitive. Instead, Ms. Lydia, the greatest critic of *No Child Left Behind*, is highly competitive and withholds what she considers an invaluable resource from her colleagues. In fact, she appears to be more competitive than any other teacher on the biology team. On the surface, Ms. Lydia’s actions may be

counterintuitive, but a deeper analysis of the relationship between teachers on the biology team provides insight into Ms. Lydia's actions.

The degree to which teachers submit to accountability policies influences the social dynamics of the biology team. In a study of Kentucky's high-stakes language arts portfolio assessment, Scott (2005) found that accountability coalitions form along two poles, those teachers who buy-in and those who do not. At HHS, accountability policies definitely create coalitions, and they form along poles, but the division is not solely influenced by buy-in to the policy. The story is more complicated. Ms. Lydia certainly finds herself to be isolated from the other teachers, but her isolation may be the result of multiple outcomes of accountability.

One of the main reasons for her isolation may be the division that occurred at the beginning of the year about environmental science being remade into pre-biology. At that time, she isolated herself from the other environmental science teachers who work together under Ms. Calypso. Their partnership serves to foster a bond between Ms. Harris, Ms. Calypso, and Dr. Stevens. Ms. Lydia could have been a part of the partnership if she were willing to submit to the demands of accountability. On the other hand, other factors likely influenced her decision or, at least, facilitated the division. Ms. Calypso, as a special education teacher, co-teaches with both Dr. Stevens and Ms. Harris, which makes it easy for the three of them to plan and interact. Further, as new teachers, both Dr. Stevens and Ms. Harris are willing to follow Ms. Calypso's lead. Ms. Lydia, who is older and equally experienced, may be less likely to follow Ms. Calypso. Also, if Ms. Lydia were included, the four teachers would have to schedule specific meetings in order to engage with one another, which would create a further burden on them.

As is, the demands of teaching a pre-biology course without a curriculum plan create fissures between teachers. Ms. Calypso refused to co-teach with Ms. Harris for the last two weeks of school because she believes that Ms. Harris did not fulfill her responsibilities throughout the year. She releases some of her resentment toward Ms. Harris:

Ms. Harris does not live up to her end of the bargain. I am just frustrated. The students are deprived. The teacher doesn't contribute, and I build resentment.

Ms. Calypso's frustration may be warranted, since Ms. Harris is new, a substitute teacher, and does not know science. It would be practically impossible for her to "live up to her end of the bargain." Thus, Ms. Calypso, Ms. Harris, and Dr. Stevens are brought together by accountability measures, but at the same time, the measures create unreasonable pressures on the team, which result in frustration and division. In other words, accountability structures can foster both unity and division among teachers.

In addition to creating a rift between her and the other environmental science teachers, Ms. Lydia's revolt at the beginning of the year frustrated Ms. Victoria. Although Ms. Lydia ultimately pursued a biology centered curriculum in environmental science, she still is perceived to be teaching too much environmental science in environmental science. Further, it is likely that Ms. Lydia's outward noncompliance has had negative ramifications for how her biology courses are perceived. In general, Ms. Victoria sees Ms. Lydia as somewhat of a pariah and fears that her students will be less prepared for the HSA and County Exam. Although Ms. Lydia's view of accountability colors Ms. Victoria's perception of her work, Ms. Victoria also is likely influenced by her friendship with Ms. Calypso. After the environmental science rift, it was likely more

natural for Ms. Victoria to be partial to her friend Ms. Calypso and, as a result, Ms. Harris and Dr. Stevens.

Ms. Khana seems to be uninvolved in coalitions. For the most part, she keeps to herself and does what she is told. As the newest “highly-qualified” biology teacher at HHS, she teaches the two general biology courses. Ms. Khana outwardly subscribes to accountability, and she is under little scrutiny from Ms. Victoria. It seems that Ms. Khana consciously works to remain unnoticed in the department, and she generally is successful in her pursuit. Thus, although she is not a member of any inter-team coalitions, she remains out of the fray by conspicuously following the rules and keeping to herself. Ms. Calypso, who also is able to avoid attention, is highly critical of *No Child Left Behind*, but her criticism is more subtle. Because she does not outwardly challenge Ms. Victoria, she retains her trust. Thus, she is given more freedom to adjust her curriculum to follow her own view of teaching.

Ms. Lydia, on the other hand, seems to do the opposite of Ms. Khana and Ms. Calypso. She outwardly criticizes accountability policies. As a result, she draws attention to herself. Although she essentially follows the rules, her indiscretions, like switching the curricular order of biotechnology, are visible to her colleagues and Ms. Victoria. Therefore, it is not simply her lack of submission to accountability that pits her against her colleagues. How Ms. Lydia criticizes accountability may be what isolates her. She is an outspoken critic of *No Child Left Behind*, and she is unwilling to defer to Ms. Victoria, who outwardly supports the policy. Consequently, her minor deviations from the mandates of accountability get attention, and she becomes someone who Ms. Victoria considers to not be “a team player.” Ironically, in terms of her practice, Ms. Lydia may be

as complicit to the demands of accountability as the other teachers, who appear to be more submissive to the policies.

Ms. Lydia's outsider status is likely what drives her to seek an "edge" over her colleagues. She is outnumbered five to one in the department, so she needs something extra to compete with the other biology teachers. Regardless of her personal views, success as defined by accountability measures is currently the most prominent to those around her. Ms. Lydia's students' test scores are guaranteed to be noticed. If her statistics are as good as those of her colleagues, she, at least partly, is vindicated for her position on accountability. Thus, Ms. Lydia has a tremendous incentive to horde her "secret weapon." Ironically, Ms. Lydia, an outspoken critic, provides an ultimate example of submission to the structures of accountability. She submits herself to something she scorns.

Ms. Lydia's guilt about her decision is obvious. One of the justifications for her "secret weapon" is that she does not want to make the other teachers "feel bad" because HHS cannot afford new books. BCPS and HHS allocate many resources toward implementing accountability measures. If Ms. Lydia is correct and Maryland uses the book for writing the HSA, HHS may be likely to buy at least one copy of the text for each teacher. Besides, Ms. Lydia bought her own copy; yet, she does not offer the school or the teachers an opportunity to buy the book. Thus, the justification is likely an excuse for her conscience, nothing more.

Further evidence of Ms. Lydia's guilt arises two days later during our second interview, when I ask her how *No Child Left Behind* has influenced her relationship with colleagues. Ms. Lydia underscores the friendly competition between the biology teachers:

It is not a competitive group. I shouldn't say that. It's a competitive group, but it's a friendly competition. We share, and we hope everybody is successful. We are collegial. There is nothing that we hoard or keep from one another to look better. We share everything and encourage everybody to be the best they can.

Of course, Ms. Lydia does “hoard” a textbook “to look better.” Her concealment demonstrates the degree of her guilt by revealing that Ms. Lydia considers her actions to be wrong. She commends her department for not exhibiting such behaviors. Thus, Ms. Lydia knows that what she does is wrong, but she does it anyway. The power of accountability is so great that it compels her to violate her own morals and beliefs about what is right and wrong, creating what is likely to be quite an uncomfortable in-dwelling for Ms. Lydia.

The Eight Hundred Pound Guerilla in the Classroom

An educational film's address to the student...invites her not only into the activity of knowledge construction, but into the construction of knowledge from a particular social and political point of view. This makes the “viewing experience” and the sense that we make of films not simply voluntary and idiosyncratic, but relational – a projection of particular kinds of relations of self to self, and between self, others, knowledge, and power. (Ellsworth, 1997, p. 25)

Dewey (1916) argues that education is a social experience. The classroom sets a context for students in which particular lessons are embedded. Thus, the medium of education gradually develops “a system of behavior, a certain disposition of action” (p. 10). In other words, participants in the classroom do not learn simply the curriculum. They make meaning from the curriculum, which includes the role of teachers and students. These meanings generally fall under the realm of the hidden curriculum, rather than the overt curriculum (Pinar, Reynolds, Slattery, & Taubman, 2000). The overt curriculum includes the objectives, indicators, test taking strategies and anything else that

is embedded in state science standards. The hidden curriculum falls outside the scope of the course materials and scheduled lesson plans (McLaren, 1989). It deals “with the tacit ways in which knowledge and behavior get constructed” in the classroom (p. 183).

Accountability often frames the role of the teacher as “neutralizing, eliminating, or distracting students from the differences between what the curriculum ‘says’ and what a student gets” (Ellsworth, 1997, p. 41). The classroom is uncontrollable. The curriculum is informed by broader social structures, participant relations, and interactions in order to produce knowledge (Pinar, 2002). Fundamentally, the curriculum cannot be the passive vehicle of policy aims. Rather, the curriculum must create meaning. Thus, in terms of the overt curriculum, accountability policies attempt to narrow and control the role of teachers by standardizing knowledge and assessing it through a high-stakes state-wide standardized test. Within the climate of accountability, the overt curriculum casts the teacher as a transmitter of standards. These changes in the overt curriculum have a powerful influence on the meanings made in the hidden curriculum. Accountability practices do not seek to reconstruct social stratification or challenge the status quo. On the contrary, research suggests that accountability policies reinforce racial and class hierarchies (Darling-Hammond, 2004; Irons & Harris, 2007).

Accountability policies, however, have significantly influenced the relationship between participants in the biology classroom at HHS. They have created a new agent in the classroom, the high-stakes test, namely the HSA and County Exam. The high-stakes test has ascended to hold great power in the curriculum. In the words of Ms. Lydia, “BCPS has been attempting to align itself with Maryland curricular and teacher standards since the HSA was introduced.” The district has developed a curriculum that, according

to Ms. Victoria, “is perfectly aligned with the HSA.” It includes all the necessary topics and indicators, and it trains students in multiple choice questions and BCRs. “HSA,” “BCR,” and “County Exam” are accepted components of the classroom lexicon. Teachers regularly include test preparatory drills and worksheets into their lesson plans, and they often bring attention to what is or is not covered on the exam.

While addressing her students during class, Ms. Khana succinctly describes the role of the HSA in the curriculum:

I’m going to tell you exactly what you need to know for the County Exam. I hate that it is all about the test, but we are in a test driven society. So deal with it. I hate that it’s all about tests.

Ms. Khana’s statement highlights two ways in which accountability has influenced her classroom. First, her teaching has become “all about the test.” More than any other teacher, Ms. Khana focuses her assignments on the HSA. She closely follows the county curricular modules, and highlights parts of the curriculum that are most likely to appear on the HSA. She includes test preparation in her lesson plans on a daily basis, and she often tangentially alludes to the test during class.

Ms. Khana is not alone. The other teachers also regularly include the HSA and County Exam in their lessons. Ms. Victoria explains how significantly the HSA influences her lesson plans:

My least favorite aspect of teaching science is teaching to the test...I make the syllabus over the summer. I sit down with my time line. I get the assessment date in June or July. July 1st is when I think they come out. I back map from the test date...There are all these components, so by the time you finish, you know that you have this many days for this unit and so many days for this unit. Then, I’ll make my exams. We basically have a test every fifteen days in this class. So I can’t go over that because that messes up the whole schedule.

The HSA is a focus of Ms. Victoria's curriculum. She organizes her entire course – including labs, tests, and time spent per topic – around the HSA. Essentially, the HSA governs the scope, sequence, and duration of the course.

Both Ms. Khana and Ms. Victoria underscore the frustration that they feel about the role of the HSA. Ms. Khana “hates” that it is all about the test, and “teaching to the test” is Ms. Victoria's least favorite aspect about teaching science. On the surface, it may be tempting to treat the HSA as a minor curricular nuisance or annoyance that teachers have to accommodate, like taking attendance at the beginning of class. However, Ms. Victoria's description demonstrates how truly prominent the HSA is in the curriculum. It is not something minor that teachers must heed. On the contrary, it is a powerful force that governs how teachers structure and teach their courses.

In addition to redefining teachers' practice, creating more work, and increasing stress on teachers, accountability has redefined the role of the teacher in the classroom, likely contributing to teachers' frustration. When teachers organize and design courses, they translate district and school policies, their own understanding of quality teaching, the context, the needs of their students, and available curricular materials into a course. These components of the course are embedded in the planned-curriculum and manifested in the lived-curriculum, and they each play a role influencing the classroom experience. The prominence of high-stakes accountability has amplified the influence of policies on the curriculum, particularly those pertaining to preparing students for the HSA. Of course, the other factors continue to inform the curriculum, but preparation for the high-stakes test now serves as its foundation.

As a result, the role that teachers have in influencing the curriculum is diminished significantly. In other words, the HSA has gained agency, and teachers have lost it. Ms.

Lydia describes the relative role of teachers and the HSA in the classroom:

The HSA is everything. Teachers are nothing. Nobody listens to us. Everyone listens to the stinkin' test. What we want or what we think is best doesn't matter. It doesn't matter that we know what's best for the students or how to help them learn better. We have to have them ready for the test. We have to stick to the program that is given to us.

Ultimately, teachers are power-less against the HSA. In the end, the HSA is the arbiter of what students must learn. Even in environmental science, the test looms over teachers and students and enforces compliance with its curricular specifications. Parents, too, may serve to empower the HSA. Parents are likely to demand that teachers prepare their children for the test because they need to pass it in order to graduate. In essence, the HSA is the expert on the curriculum because it wields the most power. Students, teachers, administrators, schools, and districts suffer when students fail the test.

Further, the HSA has shifted the relationship between teachers and students. Students are aware of the presence of a new expert in the classroom. In the hidden curriculum, studying for the teacher's exam sends students the message that the teacher is the biology expert who knows what is important and how to evaluate students' understanding of the subject matter. Now, with the presence of high-stakes testing, students and teachers work together to prepare for the HSA. On the one hand, the new relationship undermines teachers' expertise. They no longer decide what to include in the curriculum, how to structure the content, and how to evaluate students. The HSA makes many of these decisions now. On the other hand, teachers have gained a new expertise. They are now experts on the HSA and preparing students for a high-stakes test. Rather

than turning to teachers for their knowledge of biology, I witnessed many students turn to teachers for their expertise in the biology HSA. While teachers' expertise in teaching has been reduced, it has been replaced by an expertise in test preparation. In other words, under the umbrella of accountability, students' respect for teachers partially has been supplanted by a respect for a high-stakes test and test preparation. Like an eight hundred pound guerilla, the HSA has bullied its way toward supremacy in the biology classroom. Teachers are left to redefine their expertise in-between two identities, teacher and test-prep coach.

In-Between Pedagogy

According to *Webster's New College Dictionary*, pedagogy is the "art or profession of teaching" (Berube, 1995, p. 809). Although quite limited, this definition highlights an interesting, yet false, dichotomy between the "art" and "profession" of teaching. *Webster's* definition seems to imply that teaching can be either an art *or* a profession, but in actuality, teaching is both. On one hand, teachers are professionals, like doctors and engineers. Teachers are experts in their subject area, and they know multiple methods for teaching diverse groups of students. On the other hand, teachers are artists, like writers and musicians. Much of teaching is intuitive. It forms within teachers and the classroom. Being with students in the classroom comes from innate understandings, as well as professional preparation. Quality teachers are both strong professionals and creative artists whose talents and training work together in the classroom.

Not Artistic Professionals

Interestingly, the accountability system seems to diminish both the art of teaching and the profession of teaching. In terms of art, the standards movement does

acknowledge the importance of “creativity” in teaching. The *National Science Education Standards* call for the expression of creativity in the science classroom:

Science is a discipline in which creative and sometimes risky thought is important. New ideas and theories are often the result of creative leaps. For students to understand this aspect of science and be willing to express creative ideas, [teachers] must support...a diversity of experience, ideas, thought, and expression. (NRC, 1996, p. 46)

In addition, the *National Science Education Standards*, the *Benchmarks for Science*

Literacy, Maryland State Science Standards, and BCPS curricular documents

acknowledge the importance of creative pedagogical approaches, especially for diverse groups of students (AAAS, 1993; BCPS, 2003; BCPS, 2004c; MSDE, 2002b; MSDE, 2000b). Ms. Calypso highlights the importance and nature of creativity in her science classroom:

As a special ed teacher, I have to be creative in the moment. My kids can't pay attention for more than five minutes. I need to keep them interested and engaged. That can't be planned. I can't tell you how often I wing it. I can just see how what I have planned isn't working, so I need to feel where the class is and go with it.

Mandating creativity, however, may be a paradigmatic paradox. Creativity is art, originating from inside the teacher-creator in response to external stimuli. As Ms. Calypso points out, pedagogy cannot always be planned. Sometimes it emerges from the “feel” and character of the class on a particular day. Artistic teaching is impossible to standardize and evaluate with a high-stakes summative assessment, composed solely of multiple choice questions and BCRs. Therefore, although the accountability movement may acknowledge the “art of teaching,” the structures that have arisen to implement and enforce accountability generally exclude the artistic side of pedagogy.

On the other hand, while accountability structures de-emphasize the art of teaching, they also de-professionalize teachers. By treating teaching as something that can be mandated, governed, and promoted, *No Child Left Behind* reduces teaching to a set of curricular inputs. In other words, pedagogy is reduced to a set of prescriptions that teachers must follow. In the words of Ms. Lydia:

[Accountability structures] try to control the teacher. They make everything prescriptive. The teacher just reads the curriculum and repeats it like a parakeet. Every teacher says the same thing. Thank goodness we don't have that at Halbert yet. I wonder how long we will be protected.

The BCPS biology curriculum modules are not prescriptive like some curricular materials that require teachers literally to read the entire lesson from a book, and current school policies grant the biology team a degree of autonomy so long as the teachers work to accomplish the goals of *No Child Left Behind*. Nevertheless, Ms. Lydia notes a trend toward prescription. Although the teachers at HHS are not yet required to serve as microphones for printed materials that align with the philosophies of accountability, they are told how to write exams, what kinds of questions to use for evaluating students, exactly which topics to teach, and how to define success.

These efforts to control teachers underscore the de-professionalization of teaching. Professionals are trusted as experts. They understand their craft better than anyone else. Therefore, others turn to them for their knowledge and expertise. The current manifestation of accountability, by its very nature, mistrusts teachers' expertise. The premise behind accountability is that teachers need to be monitored in order to ensure that they are doing their job well. Accountability policies are meant to serve in lieu of a supervisor or a time clock at a factory job. In some ways, accountability treats teachers as unskilled or narrowly-skilled laborers, who are implementing procedures on

an educational assembly line. Thus, according to the accountability system, teachers are not professionals, artistic or otherwise. They are implementers who follow the rigid guidelines established by outsiders to define their work.

Understanding Pedagogy Within Accountability

What, then, is pedagogy under the climate of accountability? Accountability seems to hold pedagogy between art and standardized technique, or between profession and unskilled implementation. Its location in-between these understandings can be confusing, like standing on a ceiling or falling up. When teachers become immersed in the bubble of accountability, some of their own basic tenets of instruction and understandings of pedagogical processes are challenged, like an intuitive understanding of mechanics is challenged in zero-gravity space. As a result, teachers may lose perspective about what they consider to be good teaching, like astronauts in space who may become comfortable floating.

Dr. Stevens describes teachers' complete immersion within the structures and processes of accountability:

I don't have a different way of looking at [teaching]. You go into it, and this is the way it always was. It's kind of like this is the way it will always be, so I don't know. I don't know how else to look at it.

Dr. Stevens underscores that he is a new teacher who came to the profession after the passage of *No Child Left Behind*. He suggests that as a new teacher, he may not have the larger perspective that veteran teachers possess. However, other teachers who have experience teaching prior to *No Child Left Behind* find themselves equally enveloped in accountability mandates, despite their previous experience in teaching. In the words of Ms. Victoria:

There's just so much pressure. It's everywhere. I mean it's still fun. It's still biology. I'm still with the kids, but it's not how it used to be. I am always doing something [to comply with the mandates of accountability]. I always feel the test. It's always there.

These words come from a supporter of *No Child Left Behind*. Ms. Victoria and Dr. Stevens demonstrate the constant presence of accountability. Although teachers are motivated and guided by a multitude of factors, accountability informs, to some degree, most of their curricular decisions.

Thus, teachers are challenged to understand their own practice. They are torn between their own understandings of pedagogy and accountability's understandings of pedagogy. A particularly clear example of such pedagogical in-dwelling deals with a standardized curriculum. When I ask the teachers about possible positive effects of accountability on biology teaching, four teachers suggest that it offers all students biology through a standard curriculum. In other words, they claim that the standardization that has resulted from *No Child Left Behind* includes all students in the biology curriculum, regardless of their background. On the other hand, when I ask teachers about some of the negative effects of accountability, three teachers suggest that a standard curriculum reduces creativity, prohibits teachers from covering topics in-depth, and induces teaching to the test.

Interestingly, one teacher, Dr. Stevens, espouses both views. He captures the tension between defining pedagogy according to his own understandings of quality teaching and best practices and subscribing to the definition held by *No Child Left Behind*. When I ask Dr. Stevens if accountability has positively influenced science teaching, he responds:

I think [accountability] has made [the curriculum] more standardized. Everybody teaches the curriculum this particular way. Like if you're going to get into thermodynamics or something like that or modern physics. You're pretty much all going to teach the same thing in modern physics. You're not going to bounce out to something odd in relativity or string theory or something like that. You're not going to give them something like that because they're going to need the basic stuff for modern physics. It kind of standardizes [the curriculum]. I don't know if it's good or bad.

Immediately, after his response, I ask Dr. Stevens if accountability has negatively influenced science teaching. His answer reveals his space of in-dwelling:

That would probably be the same answer. You can't go off on a tangent. Well, you could. I wouldn't say that it's wasting time, but if you wanted to go discuss string theory or something like that in physics, you could, but you would be taking time away from teaching stuff that would help [students] with the standardized test. Even though I wouldn't call that a waste of time, it would be time spent other than on the test.

As a new teacher, Dr. Stevens is likely still formulating his teacher identity. He is searching for who he is in the curriculum. He understands the powerful influence of accountability structures. He feels the pressure to see the curriculum as a finite, definable entity that can be standardized and controlled. On the other hand, he recognizes that the curriculum cannot be standardized because it is informed by its participants, teachers and students. Dr. Stevens realizes that some of the best teaching and learning emerges unplanned from the curriculum. A standardized curriculum ignores the infinite pedagogical possibilities that regularly arise in the classroom. In fact, standardization and prescription largely reduce teaching and learning by stunting organic curricular growth and diversity.

Dr. Stevens' opposing views of the standard curriculum reflect two different pedagogical paradigms. The positive view of the standard curriculum evinces the equality discourse of *No Child Left Behind*, which holds that all students should get quality

teaching, regardless of race, class, and location. The challenge to this discourse comes from the reality of teaching, which is captured by the second view. Rather than offering quality instruction to all students, a standardized curriculum ends up reducing pedagogy to its most common denominator by offering all students instruction that has been torn apart by accountability mandates. Even the *Benchmarks for Science Literacy*, which was developed as part of the standards movement, argues that common goals do not require a common curriculum (NRC, 1996). A common curriculum implies that all students are the same. As Ms. Calypso points out, “A core curriculum is important, but it needs to be more of a guideline that allows teachers to deviate from it and adapt it to their own particular classrooms.” A standard curriculum cannot meaningfully address the individual needs of students. Instead, with a standardized curriculum under *No Child Left Behind*, all students receive a superficial survey biology course with a focus on test preparation techniques.

Dr. Stevens’ response to my question about the negative effects of accountability highlights another pedagogical place where teachers in-dwell. Teachers in the current accountability system must navigate between their work as teachers and their work as test preparation coaches. Akin to their negotiation between a standard and open curriculum, teachers understand the infinite opportunities that come from teaching, as well as the severe limitations of test preparation pedagogy. Nevertheless, the powerful stakes of accountability induce teachers to teach to the HSA because, to them, the benefits out way the costs.

All six teachers are extremely concerned with the HSA graduation requirement. Ms. Khana speaks for all the teachers:

Graduation, that's the big concern. If students don't pass, they don't graduate. That's a problem. They should all be able to pass [the HSA], but some of them won't and then they have to go through life without a high school diploma.

The teachers all realize the tremendous stakes associated with the HSA. The test can change students' lives in positive as well as negative ways, a proposition that places unreasonable pressure on both teachers and students. Dr. Stevens names why it is inappropriate and unjust to attach such severe stakes to a single test:

What happens if they have a bad day? What happens if they have the stomach flu for three days before they have the HSA's. They go in. They take the HSA. They blow it. They still can take it again, but, you know, it's not like they can take it really quickly. They have to wait until it comes up again. It looks bad. Like I said, a kid has one bad day and blows the HSA. Now that's high-stakes.

The HHS biology team is well aware of the high-stakes associated with the HSAs. Although some of them may consider the test to be bankrupt and unjust academically, they must defer to it because of its powerful implications.

On the other hand, teachers do recognize that they are not solely responsible for students' success and failure. They understand their limited ability to get all students to pass the test. Teachers understand the decisive role the HSA will play in students' lives, but not all students grasp the test's significance. Therefore, not all students are willing to spend their time being prepared for a summative exam. Ms. Victoria comes to students who are at-risk of failing the HSA at the end of school and, in her own words, "drags them to the after-school HSA prep course." Ms. Khana often complains that students do not come to her lunchtime HSA review sessions, even when she makes it mandatory. Ms. Harris describes her frustration with students in her class who previously failed the English HSA and, now, they have to retake it in order to graduate:

They don't realize how important it is. Some of [the students], you would notice that they wouldn't even go to their tests. An aid would send them a pass or send me a pass [to take the English HSA], and I found out that a couple of them wouldn't even go. Or they didn't even want to leave. I'd say, "Here, you have to go at 10:15. Please, leave and take your test." They'd say, "Well, I don't want to go." I'd tell 'em, "You have to. Don't you want to pass? You have to do this to graduate. But they don't want to, which, to me, I'm baffled by it. I mean, why don't you want to graduate? Why don't you want to succeed?"

Are teachers truly powerless to get some of their students to understand the significance of the HSA, or does high-stakes accountability necessarily exclude certain students from the curriculum?

Teachers use every tool in their accountability set to induce students to prepare for the test. Unfortunately, the tools they need are missing from the toolbox. They need creativity, imagination, and freedom to reach students who have difficulty learning. In order to help students make meaning of learning and connect with their own education, teachers need to express themselves "in a number of different languages" (Greene, 1995, p. 57). These different languages are underutilized by the monolithic structures of accountability which prescribe one-size-fits-all solution to academic underachievement. Thus, teachers may be left powerless to reach some of their students. Although they may be frustrated with students' unwillingness to prepare or even take the HSA, teachers ultimately accept that they cannot ensure that every student passes. In other words, to some degree, they may let themselves off the hook, which likely makes their in-dwelling between test preparation and quality teaching more bearable.

Nevertheless, teachers do see graduation as a foundation on which all else is built in life. They see themselves, at least, partially responsible for their students' success. They know how difficult it can be to go through life without a high-school diploma.

Therefore, teachers are willing to sacrifice some of the benefits of quality teaching in order to prepare their students for a test they must pass for graduation. Part of teachers' responsibility under the system of accountability is to negotiate between their role preparing students for a test that opens a gate to the next stage of life and teaching them the knowledge, skills, and values they will need for life.

Summary

Within the system of accountability, teachers stand on a shaky foundation. Teachers must act in accordance with their own curricular understandings, while attending to the mandates of accountability. Often, these demands call teachers in different directions. As a result, their footing is unsteady. They are compelled to work amidst myriad unresolved contradictions. Their role in the curriculum becomes a balancing act between what they believe to be quality teaching practices and the mandates of accountability. Teachers must compromise some of who they are by dwelling in-between two beings, their teacher-self and their accountability-self. They concede some of their moral obligations to students, their own professionalism and skills, and their understandings of quality pedagogy.

Using the words of Doll (2000), in the climate of accountability teachers are splitheads who are, at least partially, "playing a part in a script written by others" (p. 83). They have to exchange their true self for an accountability self. Being true to their teacher-selves and their students can be supplanted by implementing standards and assessments onto their students, which sometimes contradicts their pedagogical beliefs. As teacher-implementers, teachers run through someone else's curriculum. They do not make their own course or their own meanings. Due to the dynamic nature of students and

the learning context, a curriculum is ever changing and continuously elicits new meanings from students and teachers (Applebee, 1996). The standardized course laid out by accountability, however, is static. It ignores the participants and strips them of their identity and individuality. Thus, teachers must negotiate their selves through a terrain informed by a policy that ignores their expertise and individuality.

In the next and last chapter, I turn to the implications of teachers' negotiations. I examine how the three themes that I have developed in the last three chapters inform the research questions in this study. I also explore some of the paradoxes that are a consequence of the current manifestation of accountability. Then, I turn the implications of this study for science teachers, the curriculum, and the accountability system, paying particular attention to the impending scheduled amplification of the science accountability structures. Finally, I offer some critical propositions to address accountability structures based upon what I have learned through this research.

CHAPTER VIII: RECONCILING COMPETING HORIZONS

Re-Viewing Horizons

Over the last three chapters, I have outlined three themes that emerged from my study of six biology teachers at Halbert High School: “re-defining science teaching,” “the pressure mounts,” and “teaching in-between.” These themes capture how significantly high-stakes science accountability as mandated by *No Child Left Behind* has influenced the biology curriculum, specifically, and the science curriculum as a whole. The current manifestation of educational accountability has had a profound effect on the practice of teachers and their own perceptions of their work. As a result, the six teachers who participate in this study find themselves negotiating through numerous competing horizons that reflect the current educational climate under the mandates of accountability as dictated by *No Child Left Behind*.

In Chapter V, I uncover how biology, as the only science tested on the High School Assessment, has swept through the science curriculum at Halbert High School. It has usurped environmental science and even seems to be infiltrating the physical sciences by way of matter and energy. I outline how the HSA supplants the science standards as the driving force behind the science curriculum. It limits pedagogy and dramatically alters the content of the biology classroom. The emphasis on scientific inquiry found in state and national standards conspicuously is absent from the summative assessment, and, as a result, from the biology curriculum. Further, the HSA has altered the content of biology itself and reduced it to a set of standardized discreet knowledge bits that can be tested through superficial multiple choice and Brief Constructed Response (BCR) questions.

In Chapter VI, I explore how accountability has increased the amount of stress on teachers at Halbert High School and prescriptively limited and controlled their work. By its very nature, the current manifestation of accountability appears to mistrust educators. Its ethos holds that teachers must be monitored by outsiders to ensure that they are, in fact, doing their work properly. In reality, it ignores their productive-ness and limits teachers to implementing someone else's curriculum and knowledge. Accountability places teachers into a tightening vice that strictly defines their identity (as "highly-qualified" or not "highly-qualified") and places strict restrictions on their practice and the very meaning of their work. Time that teachers could spend producing curricular knowledge is now monopolized by grading for the purpose of reporting, producing and analyzing quantitative data with limited meaning, as well as implementing remediation and intervention programs for students at risk of failing the HSA. Accountability attempts to commodify teachers into universal discreet units that can be compared with one another according to students' test scores. As such, accountability has increased teacher frustration and likely diminished their passion for their profession.

In Chapter VII, I examine how teachers at Halbert High School find themselves in-between the mandates of accountability and their own understandings of quality teaching. I explore how teachers have compromised their own values. In order to justify their own culpability in the new curriculum, teachers either deny the influence of accountability on their own practice or submit to the prescriptions of *No Child Left Behind*. As a consequence of high-stakes accountability, teachers have been, in many ways, de-professionalized. The HSA has become the curricular expert in the classroom, derogating teachers to being test preparatory experts and fomenting competition along the

lines narrowly defined by accountability structures. As such, teachers find themselves in-dwelling between competing understandings of their pedagogy, curriculum, and profession.

In this chapter, I turn to the “lessons learned” from this case study for and about teachers who may find themselves in similar teaching circumstances (Stake, 1995). I uncover how the three themes developed in the previous chapters elucidate the research questions that guide this study. Specifically, I examine all three levels that my questions explore: the tensions embedded in the policy, how teachers perceive and understand the tensions, and the influence they have on teachers’ practice and the biology curriculum. Then, I explore the implications of this study for science teachers and the accountability system. Finally, I offer some suggestions for the future of accountability, both for science and the general curriculum, based upon what I have learned through this study.

A Return to the Research Questions

The overarching research question that is the focus of this study asks: How do high school biology teachers negotiate the explicit and implicit messages regarding high school biology accountability policies governing their work? This question is examined through three auxiliary questions and sub-questions. In this section, I turn to an examination of these auxiliary questions. Specifically, I examine how this study and the three themes developed in the previous three chapters illuminate and inform the research questions.

Tensions Embedded in Accountability Policies

The first auxiliary question asks, what, if any, tensions are embedded within and across the national science standards, Maryland State science standards and the high-

stakes assessment, and the district's biology curriculum? This question includes three layers of policy messages – the national, state, and district levels that are designed to influence science teaching. In chronological terms, the first documents to be written were the national science standards, first the *National Science Education Standards* in 1993 and then the *Benchmarks for Science Literacy* in 1996. Maryland subsequently developed science standards and core learning goals and the High School Assessment. Finally, Buckley County has developed science blueprints, frameworks, and curriculum modules, which it continues to revise.

Both district and state documents purport alignment with national science standards (MSDE, 2000b; MSDE, 2002b; Townsend et al., 2003). However, policy documents often make stated claims that turn out to be unfounded (Elmore & Sykes, 1992). Both sets of national science standards were written by two separate and autonomous organizations. Although Leonard and Penick (2005) report that the congruency of the two documents is “remarkable,” I expected that the documents would deviate from one another with enough significance to result in mixed signals from the national level. The mixed signals would then be incorporated into state and district level documents, which would result in tensions at all three levels. For the purposes of this study, however, I found the messages from all the standards documents that I reviewed to be remarkably unified, especially in terms of their emphasis on scientific inquiry-based instruction, scientific processes and skills, and diverse pedagogical approaches that meet the individual needs and learning styles of students. Contrary to my expectations, I found essentially no pedagogical mixed messages between the national, state, and district policy layers.

However, I did find mixed, and opposing, messages between the standards documents and the high-stakes test. While the standards call for diverse, process-oriented, and student-centered pedagogy in the science curriculum, the HSA promotes standardization, teacher-centered pedagogy, and an emphasis on biology content over processes and skills. Despite the proposition that the HSA assesses Maryland's science standards in their entirety, I find that the HSA only assesses a small fraction of the standards. Besides the obvious exclusion of all sciences except biology, the HSA is aligned strictly to specific indicators derived from the standards. Further, it evaluates each indicator discreetly such that questions align to a single narrow topic (MSDE, 2002d; MSDE, 2003b; MSDE, 2004d; MSDE, 2005c; MSDE, 2006c). Thus, while the standards call for an integrated approach to science that includes other disciplines like mathematics and technology, the HSA promotes an understanding of biology as a discipline that can be subdivided into discreet pieces of knowledge. In terms of depth, the structure and format of the HSA mainly calls students to regurgitate superficial knowledge. Overall, the HSA and the science standards promulgate very different conceptions of science and, as a result, science teaching practices.

Teachers' Perceptions of Policy Messages

Based upon the interviews and observations, I uncovered numerous tensions that teachers at HHS experience, but the sources of the tensions are not the mixed messages between the HSA and the science standards. When I formulated the research questions, I realized that I could not simply assume that teachers would perceive possible mixed messages embedded in policy documents because they may be unfamiliar with all or some of the documents that I examine for this study. All six teachers, including the

environmental science teachers, are reasonably familiar with the HSA. Ms. Calypso, Ms. Victoria, and Ms. Lydia are all quite versed in the standards documents. They have been exposed to the standards through education and professional development in BCPS. Further, Ms. Calypso and Ms. Victoria are seeking national certification in biology, which requires an intimate knowledge of both the *National Science Standards* and the *Benchmarks for Science Literacy*. Because both Maryland and BCPS are aligned with the national standards, Ms. Calypso and Ms. Victoria are familiar with all three policy layers. Ms. Lydia has studied these documents as part of her graduate studies. Further, all three teachers have many years of experience teaching at BCPS, which includes and builds upon the standards in professional development workshops and teacher handbooks. The other three teachers do not appear to have any content knowledge of the standards documents. Dr. Stevens and Ms. Khana have heard of the *National Science Standards* and the *Benchmarks for Science Literacy* but claim to be unfamiliar with their content. Ms. Harris, who is certified in physical education, not biology, has not heard of the science standards and does not know the meaning of scientific literacy, a central tenet of the national standards.

Ironically, I did not find that familiarity with the documents is related to the tensions that teachers experience. In other words, all six teachers perceive competing tensions, but the tensions that they experience do not stem from the extent of teachers' knowledge of the standards' content. Even Ms. Harris, who has almost no familiarity with what are considered to be standards-recommended practices in science teaching, experiences some of the same tensions felt by the veteran teachers, Ms. Victoria, Ms. Lydia, and Ms. Calypso.

If the dissonance between standards and the HSA is not a source of tension, even for teachers who know the content of the standards, what is? In formulating this auxiliary question, I also expected that teachers would experience other tensions based upon different pedagogical beliefs and practices, as well as the school's social and professional environment. Throughout the duration of this study, I found major contradictions between accountability structures and teachers' pedagogical knowledge, content knowledge, beliefs, and experiences. These contradictions, not those associated with the standards, were the major sources of tensions for all six biology teacher participants.

Addressing her students during a rare few minutes of downtime at the end of a class, Ms. Victoria provides a vivid example of the contradiction between accountability and her pedagogical understandings:

I'm going to be evaluated either next period or sixth period. [The evaluator] is just going to see me lecture. That's terrible....But I need to lecture. I'm sorry, but I can't waste time with fun activities. I need to get you guys ready for the exam. You guys need to be ready. The only way you'll know everything is if I tell it to you.

After class, when I talk about the evaluation with Ms. Victoria, she tells me that "good teaching engages students and ensures that they get what the teacher tells them." She describes lecture as an "opportunity for a teacher to tell students what they need to know in a short amount of time." According to Ms. Victoria, lecture almost certainly ensures that some students "won't get it," but she has so much material to cover that she cannot address the individual needs of each of her students. In other words, Ms. Victoria seems to have a solid understanding of what she considers to be good pedagogy, but accountability mandates contradict her understandings and pull her in another direction.

In addition to contradictions between pedagogical understandings and accountability pressures, I also find tensions between teachers' understandings of biology content and the mandates of accountability. Ironically, the most obvious example is provided by Ms. Khana who teaches more to the test than the other teachers. When I ask her how *No Child Left Behind* views science, she responds:

I don't teach science. I teach facts. Simple facts for that matter. These students don't know biology when they get out of my class. I try. I try to tell them to see how important science is. How much of a role it plays in their lives. I try to get them to watch the discovery channel and read outside [of class]. But I just don't have time in here. I have to get them ready. I have to teach them what they need to know for the test. It's really a shame.

Like the example with Ms. Victoria, Ms. Khana knows biology. She has a Master's Degree in biology. She has twenty years of teaching experience and considers herself a scientist. The accountability mandates, however, call her to teach disconnected facts that are aligned with indicators. She is aware that, as a discipline, biology is a process toward integrated understanding not a compilation of indicators, yet in the classroom she lives the contradiction between teaching her understanding of biology and teaching HSA biology.

Although the two examples that I provide come from veteran teachers, Ms. Harris and Dr. Stevens experience the same tensions with pedagogy, and Dr. Stevens perceives them with the content. They both are aware that teachers should try to help students learn, think critically, and understand material. They both express frustration and, as new teachers, surprise with the overemphasis on memorization. Dr. Stevens says, "When I started teaching, I thought that I would be primarily teaching students how to learn, not how to memorize facts." Despite their expectations and pedagogical understandings, the

teachers find themselves in an educational environment that promotes memorization and superficial regurgitation.

Interestingly, teachers' notions of quality biology teaching are mostly congruent with the content of the standards documents. However, even the three teachers who know the content of the standards only report contradictions in terms of their own understanding of quality teaching. No one mentions any contradictions between the HSA and science standards. The teachers essentially seem to disregard the science standards, even if they agree with them. In other words, although many of the principles embedded in the science standards comport with teachers' understanding of quality science pedagogy, unlike the HSA, the standards do not explicitly influence pedagogy, even when teachers are familiar with them.

The science standards appear not to have penetrated the curriculum at HHS for two reasons. First, the standards do not have a mechanism to induce changes in teachers' behavior. Teachers may exhibit some of standards-recommended practices in their work, but they do not credit the standards for their practice. In other words, all five teachers who were prepared to teach science have learned much of the content espoused by the standards in science pedagogy courses and professional development activities. However, they do not attribute their knowledge to the standards documents. In effect, the teachers see the documents as a reflection of quality teaching, rather than as something that induces quality teaching.

Secondly, the standards are overshadowed by the HSA. At the study's onset, I anticipated that one policy layer may be more important to teachers' work than another. In terms of influence on teachers and the biology curriculum, the HSA trumps all of the

other documents that I reviewed. The HSA influences professional development activities, monopolizes teachers' meetings and inter-departmental faculty collaboration, influences the curricular choices that teachers make in the classroom, and, in many ways, redefines the role of the biology teacher. Unlike the science standards, the HSA has the "stakes" attached to it that induce teachers to change their practice. If the teachers were ever to consider whether to teach to the standards or teach to the HSA, they likely would be induced to teach to the HSA even if they found the standards to be more aligned with their understandings of quality science instruction.

Ms. Victoria's dilemma about being observed demonstrates the distinction between the influence of the standards and the HSA. In the middle of April, essentially one month prior to the administration of the HSA, Ms. Victoria is slated to be observed as part of an evaluation process that teachers undergo every five years. Ms. Victoria knows that the observation will occur on a particular day, but she does not know exactly when during the day it will occur. During a break she tells me:

I am concerned that [the evaluator] will just see me lecture. I know that I should be engaging students in activities, asking them questions, making sure that they get it. But I don't have time for that. I can't do it right now. We have so much to cover [to get the students ready for the HSA]. I have to lecture.

In this example, Ms. Victoria is torn between what the science standards call her to do in the classroom and what the HSA calls her to do. Her understanding of quality pedagogy, which coincides with the science standards, and the HSA pull Ms. Victoria in opposite directions. She does not appear to consider disregarding the HSA. She knows that lecture may be a less effective strategy for her students' overall understanding of biology, and she is aware that she may get a poor rating by the evaluator, but she still chooses to teach

to the HSA. In a high-stakes accountability system, the HSA defines success. Ms. Victoria is likely to know that, for her overall evaluation, a higher rate of passing on the HSA is more important than an observation. The underlying message is that the definition of quality teaching, according to high-stakes accountability, is more important than standards-recommended practices and teachers' own understandings of best science teaching practices.

All six teachers perceive major tensions between their own constructions of good teaching and the HSA. In many ways, the tensions arise from the HSA's narrow definition of biology, and science in general. The teachers hold a broad view of the science curriculum based mainly on their educational background and professional experiences. The HSA, on the other hand, narrowly defines biology as a set of discreet knowledge bits, and it subscribes to the technical paradigm of teaching and learning where teachers transmit predetermined knowledge to students. Ms. Lydia describes the disconnect between teachers' understandings of their profession and the pedagogical demands of the HSA:

The HSA is not about teaching. It has nothing to do with teaching. It doesn't consider teachers or teaching. It's about facts. We are just supposed to tell students the facts. Labs, activities, inquiry – that's all extra. It has nothing to do with the HSA.

In her disappointment with the high-stakes test, Ms. Lydia highlights an important point about the HSA and biology pedagogy. The aspects of teaching that Ms. Lydia deems best practices – “labs,” “activities,” and “inquiry” – are all considered “extra” because of the HSA. In other words, the HSA compels teachers to exclude what they would consider quality pedagogy from the curriculum because it is not tested on the summative examination. The HSA does not prohibit quality teaching, but it reduces it. Teachers feel

like they cannot both adequately prepare their students for the HSA and implement their own understandings of quality teaching. In other words, fear that their students will not pass the HSA may drive teachers to abandon some of their own best teaching practices.

Teachers' Interpretations Shape Their Teaching Practice

The teachers who participated in this study do not appear to be influenced significantly by the content of national or state standards documents. As mentioned above, they, however, do respond quite vigorously to the messages they perceive from the high-stakes test. Based on the literature, I expected that teachers would respond to curricular policy messages differently, both in magnitude and in kind. In other words, some teachers might react more strongly than others, and their reactions might be quite different. Interestingly, I found a great deal of variation in terms of how teachers perceive high-stakes accountability messages, but in terms of how the messages influence practice, I found less variation than expected.

The six biology teachers at HHS greatly differ in their opinion of *No Child Left Behind* and high-stakes accountability. Ms. Lydia has the most disdain for the policy and its prescriptions, while Ms. Victoria sees the most potential in the policy's intentions. All six teachers are drawn to the possibilities of no child actually being left behind. Ms. Calypso, for example, highlights the attention that *No Child Left Behind* has given to special education students:

The premise of *NCLB* has made teachers aware, especially in science, which can be an uncomfortable environment for minorities who view science as too hard. Teachers must now pay attention to minorities. They have to try to help you learn regardless of background. Just the phrase helps teachers be cognizant of all students.

Teachers are drawn to the inclusive message, but they often are put off by how the policy's intentions are manifested in the school and the curriculum. This disjunction between intentions and practical implementation is a major source of tension for teachers. The policy's stated aims call teachers to comply with the mandates of high-stakes accountability, while the deleterious effects of the policy's implementation cause teachers to be critical of accountability.

Teaching adjustments. Regardless of how teachers feel about the policy, they all adjust their practice in response to accountability messages. Further, the degree to which teachers align themselves with the policy and its stated goals does not seem to affect significantly the degree to which the policy influences teachers' work. In other words, although she expresses the most disdain for the policy, Ms. Lydia is as influenced by the policy messages as Ms. Victoria who supports *No Child Left Behind*. Ms. Khana provides an interesting example because her understanding of quality teaching most closely aligns with the technical curricular perspective of high-stakes accountability, yet she does not agree with many of the policy's prescriptions and demands on teachers. Ironically, Ms. Khana readily implements many of the demands of accountability without buying-in to the policy. Consequently, she perhaps is the least influenced by high-stakes accountability mandates, but it is difficult to tell because she subscribes to the policy's curricular premise that teachers are transmitters of knowledge determined outside the classroom. Thus, in many ways her pedagogy would likely remain unchanged if accountability mandates were lifted.

In the classroom, high-stakes testing has caused all six biology teachers to adjust their pedagogy in response to policy structures. The language of accountability – “HSA,”

“County Exam,” and “BCR” – is incorporated into the biology classroom at HHS. Teachers often allude directly to the content of the HSA as part of a planned lesson or during impromptu tangential references to the summative exam. Teachers include “warm-up” activities to prepare students for the test, and they often highlight and repeat commonly tested topics. Because the HSA tests biology content in discreet bits, teachers emphasize single topics and concepts without integrating them into the body of biology knowledge. At times, teachers even emphasize single words and phrases that appear on the HSA. Further, because the HSA does not pay attention to scientific processes in a meaningful way, teachers can pay short shrift to analysis, labs, and critical reflection. In other words, teachers are more likely to emphasize test preparation and fact memorization than scientific inquiry-based instruction.

High-stakes accountability also monopolizes teachers’ work outside the classroom. Teachers spend many hours grading students’ work in order to report multiple data points for each student that are used to chart their progress toward passing the HSA and County Exam. Professional development centers on accountability related activities. During the summer, teachers are assigned to a weeklong review of the County Exam. Biology teacher meetings involve analysis of student data and discussions about curricular adjustments to address the mandates of accountability. In essence, high-stakes accountability has infiltrated the culture of the school, shifted the biology curriculum toward summative test preparation, and influenced teachers’ work inside and outside the classroom.

Understanding of science. In terms of teachers’ understanding of the discipline of science, I did not uncover any significant changes. Teachers certainly teach a biology

course that is reduced to a set of discreet knowledge bits, and they focus much more on knowledge than scientific process skills. In effect, teachers teach HSA biology rather than regular biology. It seems, however, that the teachers consciously are aware of what they are doing. They seem to realize that what they teach is different from the discipline that they learned and/or practiced. Thus, they are acting. They teach something different from what they know. Although they likely perceive a powerful tension between their understanding of biology and the mandates of the HSA, they are compelled to reconcile or live with the tension and teach HSA biology in order to prepare their students for the high-stakes exam.

Thus far, the process of teaching HSA biology does not appear to have altered teachers' understanding of the discipline. Although the accountability system challenges some of the core principles of science, like inquiry and inductive reasoning, teachers' understandings of the discipline are difficult to undermine. All of the biology teachers, except Ms. Harris, are experts in science. All of them studied science in college. Dr. Stevens and Ms. Calypso worked as practitioners. Ms. Calypso, Ms. Victoria, Ms. Khana, and Ms. Lydia have been teaching science for many years. The accountability system may not have the power to influence teachers' knowledge of the discipline that they teach. On the other hand, it is possible, that over time teachers may begin to absorb some of the lessons of their own teaching. Like actors who take on some of the mannerisms of their character, teachers of HSA biology may begin to lose some of the connections between biology topics, or some of their understandings of the scientific inquiry process may begin to atrophy.

As a non-expert, Ms. Harris may provide an informative example. As a teacher with certification in physical education, she came to teaching environmental science with almost no background in science. By her own admission, she learns the subject as she teaches it to her students:

I learn the material right along with them. I just stay one chapter ahead. They learn biomes. I learn biomes. They learn the nitrogen and carbon cycles. I learn the nitrogen and carbon cycles. It's hard. Sometimes I don't know what I am talking about, but I do my best.

Like her students, Ms. Harris is learning accountability biology in environmental science. How does that inform her understanding of the discipline? Does she see biology as a disconnected set of facts? Does she gain any appreciation for the scientific process? Although I am not in a position to answer these questions definitively, my observations of her teaching do not suggest that Ms. Harris is developing scientific inquiry skills or an integrated understanding of biology. If biology teachers do not possess a strong background in science, which may become increasingly more common as the demand for biology teachers increases in response to accountability mandates, the messages that they receive from the accountability system may substantially inform their own understanding of the discipline of science.

Teachers' role in the classroom. High-stakes accountability in Maryland has had a significant influence on the teachers' role in the classroom. In the case of all six teachers who participated in this study, the HSA and County Exam hold a great deal of authority in the classroom. The high-stakes test has ascended to a position of power, and it competes with teachers for agency in the classroom. The test manipulates the role of the teachers both at the level of the overt curriculum and at the level of the hidden

curriculum. In general, the high-stakes accountability system influences both pedagogy and the social climate and culture of the science classroom.

In terms of the overt curriculum, BCPS has developed curriculum modules and a County Exam that teachers consider to be well aligned with the HSA. In the words of Ms. Khana, “If students just know what’s in the curriculum [modules], they will get a 100[%] on the HSA.” The modules cover all of the tested topics and indicators, and they train students in the format, style, and lexicon of the HSA. On unit exams, students are evaluated only with multiple choice questions and BCRs. Further, teachers do their part to infuse the HSA into the curriculum. Teachers regularly include test preparation drills and worksheets into their lesson plans, and they highlight frequently tested topics. They habitually allude to the HSA during their lessons and frequently remind students how to write BCRs correctly. “HSA,” “BCR,” and “County Exam” are normal components of the classroom lexicon. The terms even may be used to define colloquial terms like “essay.”

The HSA controls lesson development. Technically teachers are free to develop their own lessons. In fact, the only parts of the curriculum modules that BCPS requires teachers to employ are the pre- and post-tests for each unit. However, teachers are bound by the demands of accountability. They are driven to teach to the test for two reasons. First, they feel compelled to help their students succeed on the HSA. They understand how important passing is for their students’ future, and they choose to teach HSA biology rather than regular biology because it has more immediate consequence for their students to graduate from high school than to have a strong, integrated understanding of biology. Further, fear of failure compels teachers to give up their own agency to accountability. If

they let the accountability system make curricular decisions and their students still fail, they can blame the county curriculum or school policies. On the other hand, if teachers exert their own agency by creating lesson plans that deviate from the rules of the accountability system, they can be held responsible for student failure. In other words, by giving up agency to the accountability system, teachers may be able to inoculate themselves from some of the “stakes” associated with student failure.

In terms of the hidden curriculum, which deals with tacit ways that knowledge and classroom dynamics get constructed (McLaren, 1989), the HSA has shifted the relationship between teachers and students. Students are aware of the presence of a new expert in the classroom. When teachers write their own lesson plans and exams, students subliminally receive the message that the teacher is the biology expert who knows what is important for students to learn and how to evaluate students’ understanding of the subject matter. Now, because of the influence of high-stakes accountability, students and teachers work together to prepare for the HSA, the new curricular expert. The new classroom dynamics undermine teachers’ pedagogical and disciplinary expertise because accountability pressures transfer many curricular decisions to the exam. Teachers no longer decide what to include in the curriculum, how to structure the content, and how to evaluate students. Many of these decisions are now made for the teachers by the accountability system.

While relinquishing some pedagogical expertise, teachers have gained a new type of expertise. They are now experts on the HSA and preparing students for a high-stakes test. Rather than turning to teachers for their knowledge of biology, students turn to teachers for their expertise in the biology HSA. The graduation requirement has made

passing the HSA has become a top priority for students. Therefore, they seek support and instruction that helps them pass the exam and ensure that they graduate. As such, teachers' traditional expertise in teaching has been reduced, but it has been replaced by a new expertise in test preparation. In other words, under the umbrella of accountability, respect for teachers partially has been supplanted by a respect for the test and preparation for it, and teachers have adapted to their new environment. In order to be considered successful, teachers have become experts in accountability, the HSA, and test preparatory pedagogy.

The passion for teaching. Apple (1992) argues that accountability leaves teachers less satisfied with their job and less passionate about their work. The only teacher to admit that high-stakes accountability has influenced her passion for teaching is Ms. Lydia. In fact, the remaining five teachers adamantly denied that their passion for teaching has been eroded by accountability. Teachers must have passion in order to teach. Passion is part of teaching, and teaching embodies passion. Throughout their careers, teachers struggle to maintain their passion and enflame new passions (Intrator & Kunzam, 2006). The need to retain their professional purpose may be why the rest of the teachers could not admit that their passion is being compromised by accountability policies. The teachers need their passion to do their work. Although they may not extinguish teachers' passion, accountability structures seem to undermine it. Perhaps it is too early for teachers to see changes in their passion. After years of working in a high-stakes accountability climate, teachers may become more aware the policy's influence on their passion for their work.

High-stakes accountability inadvertently wears on teachers' passion. The current manifestation of accountability is unlikely to induce teachers to love their work, feel compassion for students, and be passionate about their purpose because *No Child Left Behind* is about oversight, management, testing, and data. These current components of the profession are unlikely to stoke teachers' passion for their jobs. Teachers are unlikely to derive purpose and meaning from preparing students for a summative high-stakes exam. Further, evaluating teachers on the basis of a singular summative measure lacks pedagogical sense. Evaluation does not necessarily have to take the form of a single high-stakes test. In fact, the *National Science Education Standards* suggest that "Assessments can be done in many different ways. Besides conventional paper and pencil tests, assessments might include performances, portfolios, interviews, investigative reports, or written essays" (NRC, 1996, p. 6). Accountability can be holistic, meaningful, and pedagogically sound, but current structures often ignore teachers' individuality and human sensitivity. As accountability currently is constructed, teachers are being alienated from an educational system that prefers testing and data to a meaningful effort at improving teaching and learning.

The most compelling evidence for the deleterious effects of accountability on teachers' passion for their work occurred at the end of the school year. Of the six teachers who participated in this study, only three are returning to HHS for the 2007-2008 school year. Ms. Harris, who is a permanent substitute and may be expected to leave, reports that she "will never teach science again." Her proclamation probably results from her unfamiliarity with the subject matter and inexperience, but it is likely that the accountability climate and pressure associated with it only encouraged her decision. Ms.

Khana, who has been teaching for nineteen years, is leaving the profession. When I ask her why, she responds with reflection, “I’m tired. It’s time for something new.” Ms. Victoria, the strongest supporter of *No Child Left Behind* in the HHS biology department, left HHS for a school in another county in Maryland where she will have “less pressure and stress.” Ms. Lydia took an early leave in May of 2007 and applied for a job with the College Board over the summer. She only decided to return to HHS for “one more year” in August after she was not offered the position. Ms. Lydia vows to put in for academic leave in spring of 2008. Because Ms. Lydia decided to return so late, she is not teaching biology, a subject she has taught each of the eleven years that she has been at HHS. Interestingly, she is far from disappointed. She says, “It’s great! There’s no pressure!” As a result of teacher attrition, for the 2007-2008 school year, HHS has four new biology teachers, some of whom are not certified.

Clearly, high-stakes accountability seems to influence teacher passion and job satisfaction. Accountability undermines teachers’ interest in their work. The tension between accountability and passion may divide a teacher into two identities, along similar lines as Aoki’s (2005) teacher-as-being and teacher-as-thing. The teacher-as-being with passion, excitement, and compassion has little space in the accountability system. Instead, the current manifestation of accountability promotes an emotionless teacher-as-thing who is simply a transmitter of biology information. Thus, the HSA creates an inverse relationship between teacher-as-being and teacher-as-thing. While the current structures of accountability expand, the space for teacher-as-being decreases.

The Paradoxes of Accountability

I turn, now, to an examination of some of the paradoxes presented by the high-stakes accountability policy. I do not present the paradoxes in order to solve them, for paradoxes, by their nature, do not have simple solutions. In fact, they may not have any solutions at all. As such, teachers live and work among myriad paradoxes. The pedagogical relation between teacher and student may be inherently paradoxical (Ellsworth, 1997). Ellsworth writes, “Pedagogy poses problems and dilemmas that can never be settled or resolved once and for all” (p. 8). Thus, the purpose of presenting the paradoxes is to engage with the tensions that teachers’ experiences have uncovered about accountability policies and practices. The tensions uncovered in this study are consistent with the many paradoxes highlighted by the literature on *No Child Left Behind*.

A thoughtful engagement with curricular and pedagogical paradoxes often is lacking in an educational system that attempts to “fix” problems quickly with specific interventions. Policymakers often place bandages on leaks without a meaningful consideration of how their proposed solution will reverberate throughout the complex, multi-layered educational system and how it will interact with policies already in place. Contradictions and tensions are inevitable with such a piecemeal approach to education policy development. For example, the manifestation of *No Child Left Behind* contradicts many previous initiatives, such as Physics First in BCPS, Maryland’s MSPAP, and Connecticut’s student evaluation system (Sternberg, 2006). In fact, Connecticut sued the U.S. Department of Education because the testing mandates of *No Child Left Behind* contradict the state’s assessment program, which has been recognized nationally for closing the achievement gap and increasing accountability (Blumenthal, 2006).

Ironically, a national policy whose intention is to increase accountability contradicts a state's policy heralded for increasing accountability.

One reality that this study has demonstrated is that the current manifestation of accountability has penetrated the science curriculum deeply and influenced teachers' work significantly. Within the high-stakes accountability system, teachers find themselves negotiating between their own constructions of the curriculum and pedagogy and the demands of the accountability system. Teachers' complex negotiations, in some cases, may cause teachers to act in ways that contradict their own understandings and best practices. Teachers' responses to accountability pressures even may cause them to undermine their colleagues' efforts and students' learning. At the very least, the high-stakes system undermines teachers' expertise and places tremendous pressure and stress on teachers to comply with accountability policy mandates. In other words, teachers are oppressed by the paradoxes embedded in the educational system. They are pulled in multiple directions, and their pedagogy is a manifestation of their lived reality within the current educational system.

Must Accountability Fail?

From an economic perspective, the entire reaction to accountability may be a paradox. Of the \$5,395,836,962 of Maryland State education funds that went to local school districts in the fiscal year 2007, only \$849,425,012, or approximately 15.7%, came from federal funds (MSDE, 2007b). In BCPS, state aid accounts for 19.1% in the fiscal year 2007.¹ Therefore, the federal contribution to the district budget through the state is about 3.0%. Additionally, direct federal government grants account for 3.5% of the district's 2007 budget, so the total federal contribution to Buckley County Public Schools

¹ Taken from the BCPS website.

is about 6.5%. Does 15.7% of the state's budget and 6.5% of the district's budget warrant a massive effort to comply with the mandates of *No Child Left Behind*? What percentage of the state and district budget goes to the implementation of high-stakes accountability mandates?

Another important contradiction is one that I have raised previously, but is worth revisiting in light of this examination. The very structure of the current manifestation of accountability may be paradoxical. The system relies on the alignment of standards and assessments, both of which are developed outside of schools and the classroom. In science, the standards call for scientific inquiry-based instruction, which subscribes to a constructivist paradigm of learning (AAAS, 1993; NRC, 1996). The very nature of science content standards may contradict their own pedagogical prescriptions. By definition, standards are objective in that they exist outside of the classroom experience. They have truth and meaning that must be discovered in the classroom (Crotty, 1998). In other words, educational standards align with a positivist epistemology. By its own admission, however, BCPS wants to align its science curriculum with the constructivist learning paradigm, where students would create truth and meaning through engagement in the biology classroom. Therefore, science standards may not be an appropriate method for promoting the goal of scientific inquiry-based instruction in the classroom. Further, a single summative exam is an inappropriate assessment tool for inquiry-based instruction because student evaluation should match pedagogy. Perhaps, this epistemological contradiction partly explains why science standards do not have a serious influence over the biology curriculum and inquiry-based instruction is largely absent from the classroom.

A further source of contradictions within the current system stems from the misalignment of standards and assessments. The practices of curriculum, pedagogy, learning, and assessments are connected and interdependent (Carr et al., 2000). In other words, quality assessments should reflect curricular goals and intentions coherently (Treagust, Jacobowitz, Gallagher, & Parker, 2001; Wilson & Berenthal, 2005). If assessment results are to be used to provide feedback for teachers in order for them to make curricular adjustments that will help students attain understandings according to Maryland state standards, then it is essential that the standards and assessments are aligned. The notion of alignment, in and of itself, presents a paradox. The concept of curricular alignment is embedded in the technical curricular paradigm (Pinar, Reynolds, Slattery, & Taubman, 2002). The technical perspective of curricular alignment implies that teaching and learning are consistent and homogeneous across teachers, schools, and districts (De Lissovoy & McLaren, 2003). Despite the linear, technical intentions of accountability mandates, alignment is impossible in the lived reality of schools and classrooms.

The logic of the high-stakes accountability system breaks down because the assessments do not align with the standards. Although each question on the HSA is connected with an individual indicator from the state standards (MSDE, 2002d; MSDE, 2003b; MSDE, 2004d; MSDE, 2005c; MSDE, 2006c), my review of the HSA demonstrates that the connections between questions and indicators are superficial at best. The NRC (1996) argues that assessments should not check simply whether students “memorized certain items of information. [They should] probe for students’ understanding, reasoning, and use of that knowledge – the skills that are developed

through inquiry” (p. 6). The biology HSA neither aligns with the standards, nor does it heed their calls for quality assessment.

Because the HSA fails to align with the standards, the accountability system lacks coherence. Although the intention of accountability may be for students to learn the standards, which national organizations and states have taken great pains to develop and promulgate throughout the educational system, the outcome of incoherence is that standards and assessments remain independent entities in the curriculum. Because the assessment carries the high-stakes, it influences teachers and the curriculum to a much greater extent than the standards. Thus, rather than having a meaningful interplay between standards, pedagogy, curriculum, and assessments, the current accountability system promotes a disjointed curriculum and pedagogy driven by the high-stakes assessment, not the standards or teachers’ expertise.

Even if standards and assessments were aligned, the curricular intentions of the standards may prove very difficult, if not impossible, to fulfill. Case studies of teachers attempting to invoke more inquiry-based science instruction as propagated by national standards (Anderson, 1995, 1996) demonstrate many of the dilemmas science teachers face. When teachers invoke new inquiry techniques, they experience a tension between their pedagogical options, as well as tensions between their options and other demands on their work. Although some of the teachers’ dilemmas may be more perceived than real, they still influence teachers’ work.

Anderson outlines five tensions for teachers attempting to switch to inquiry-based instruction. First, teachers think that there is never enough time to do everything that needs to be done. The curricular changes promoted in the *National Science Education*

Standards and the Benchmarks for Science Literacy only place greater time demands on teachers. Second, science teachers generally think that the national science standards portray an unrealistic ideal that conflicts with the realities of the classroom. Thus, teachers may consider the changes invoked in the standards to be inappropriate for their classrooms. Third, expectations for both teachers and students are deeply ingrained within a school's culture. A change in teachers' roles probably is essential if the mandates of the standards are to be achieved in the classroom. However, it is difficult for teachers to change their roles without a change in the school's culture. Fourth, the culture of schools, especially among science teachers, is infused with the proposition that preparation for the next level of schooling is the purpose for their work. Although research suggests that their fears are groundless, science teachers often think that scientific inquiry-based instruction does not prepare students for the rigors of the next level of schooling. Finally, teachers often perceive a contradiction in the provision of "science for all" students. Many teachers see a tension between providing quality instruction for the able and willing students, while simultaneously providing for the less able or uninterested students. My observations of teachers and conversations with them support many of Anderson's claims. Thus, if contradictions between the standards and the HSA were overcome, districts, schools, and teachers likely would face serious challenges still to proper implementation of national and state standards. Thus, they still may be unable to comply adequately with the current demands of the accountability system.

Another possible paradox in the current accountability system deals with what accountability excludes from the educational process. A fundamental principle of *No*

Child Left Behind is that all students can be taught to reach academic “proficiency” (U.S. Congress, 2002b). The legislation provides that districts, schools, and teachers must be held accountable if students fail to reach the goals established for them. If these entities fail to induce student success, the legislation prescribes certain interventions like school choice, supplemental educational services, school restructuring, and school reconstitution. In its consideration of educational inputs and outputs, the current manifestation of accountability ignores everything that occurs in students’ lives outside of schools.

In the dentist analogy that I received from Ms. Calypso, the dentist brings up this dilemma to the patient:

Don't you see that dentists don't all work with the same clientele, and that much depends on things we can't control? For example, I work in a rural area with a high percentage of patients from deprived homes, while some of my colleagues work in upper middle-class neighborhoods. Many of the parents I work with don't bring their children to see me until there is some kind of problem, and I don't get to do much preventive work. Also, many of the parents I serve let their kids eat way too much candy from an early age, unlike more educated parents who understand the relationship between sugar and decay. To top it all off, so many of my clients have well water which is untreated and has no fluoride in it. Do you have any idea how much difference early use of fluoride can make?

Research demonstrates the myriad disadvantages lower class and minority students face outside of school, leaving them less prepared for success in school than their middle-class, white peers (Collins & Yeskel, 2000; Hart & Risley, 1995; Kozol, 1992; Lee & Berkham, 2002; Rothstein, 2004). Even before children enter kindergarten, the achievement gap between poor and middle-class students is tremendous (Hart & Risely, 1995; Lee & Berkham, 2002).

All of these studies suggest that children's out of school experiences, based on social and economic differences, have a significant influence on their academic achievement. In fact, after outlining the many disparities between students from different socioeconomic levels, ranging from how parents communicate with children to disparate rates of asthma, Rothstein (2004) concludes that an average annual increase on social spending of \$156 billion would be required to close the achievement gap. Rothstein argues that social factors outside of school play a significantly more important role in the achievement gap than what occurs during school. Whether Rothstein is correct or not, children's experiences and environment outside of schools play an important role in student success.

The current manifestation of accountability ignores these experiences. Kantor and Lowe (2006) argue that while *No Child Left Behind* intensifies the importance placed on education at the federal level, it simultaneously contributes to the diminution of political support for a more expansive view of public social provision. In other words, it allows the federal government to ignore social factors outside of school that influence education. In the words of Ms. Lydia:

Teachers and schools are an easy target. It is easier to put everything on the doorstep of teachers or blame teachers, rather than to take a critical look at the problems and inequality in society. They can't fix the school system if they can't fix society.

By ignoring socioeconomic inequality, *No Child Left Behind* may place teachers in an impossible position and, thus, be ineffective at achieving its goal of leaving no or fewer children behind because it ignores the influence of students' experiences outside of school.

The Paradoxical Response

Perhaps the greatest paradox of the current high-stakes accountability system deals with the depth and breadth of the response to its mandates. Many top-down educational policies have had only a limited influence on the curriculum and classroom practice, but accountability policies are influential (Malen & Muncey, 2000). The impact of accountability was obvious to me at the onset of the study. Almost immediately, I began to wonder how accountability so deeply has penetrated the curriculum. Are all of the efforts to prepare students for the high-stakes test at the district, school, and classroom levels a reasonable response to the high-stakes HSA? Is it reasonable for teachers to abandon many of their best practices in response to accountability mandates?

On the surface, the clear answer to these questions is that “high-stakes” associated with student evaluation are the explanation for such an extensive response to the accountability system. The message of the current accountability system is clear. If enough students fail the test, the district, school, administration, and teachers are labeled “failing.” Severe penalties result from such a label. Thus, fear is a powerful motivator in the proliferation of responses to accountability messages. This study’s findings contribute support to previous studies that claim that high-stakes accountability influences teachers’ actions through fear (Craig, 2004; McNeil, 2000; Pringle & Carrier-Martin, 2005). I found that biology teachers at HHS primarily fear the ramifications of students’ failure. They understand the consequences for students who will not receive a high-school diploma. Further, biology teachers at HHS are aware that they are being compared to their colleagues on the basis of their students’ test scores. Even a teacher like Ms. Lydia, an experienced teacher who wholeheartedly disagrees with *No Child Left*

Behind, fears that she will not compare well to her colleagues at HHS and across the school district. As a result, she makes a concerted effort to prepare her students for the HSA, even at the expense of quality pedagogy and the rest of the biology curriculum. Ironically, Ms. Lydia facilitates the implementation of the very policy she disdains. Thus, teachers may be complicit in fomenting the very contradictions that undermine their work.

Fear can cause an irrational response, but so can being forced to reconcile irreconcilable paradoxes. A statement that Ms. Victoria made during one of my first observations of her class underscored a critical paradox of accountability. After a class where she spent at least twenty minutes explaining how to write a good BCR, Ms. Victoria tells me:

The HSA isn't even hard. It tests like fourth or fifth grade life science. I don't know why they make the test so easy.

If the test is so easy, why is Ms. Victoria, who teaches pre-IB Biology to the school's highest achieving students, worried about the HSA? Why does she shape her curriculum to prepare her ninth grade students for a test they could have taken years ago?

Both Ms. Lydia and Ms. Khana, the other two teachers who teach the "official" biology course, also consider the test to be "easy." Both teachers do highlight that the questions are sometimes wordy and may be confusing to students, an observation that I verified through my review of the published HSAs. Ms. Khana thinks that the HSA tests students at the middle school level (a claim that my review would refute). All three teachers who teach the official biology course adamantly agree that all of their students should be able to pass the test because it is easy. In the words of Ms. Khana, "There is absolutely no reason why any of my students should fail the HSA." Why then did almost

thirty-nine percent of students at HHS fail the biology section of the HSA in the 2005-2006 school year?¹

The incongruencies between teachers' perceptions of the test, the test itself, and student scores on the test indicate the depth of the contradictions within the accountability system. The "high-stakes" and fear of consequences have resulted in a large scale response to accountability mandates. However, the responses may not achieve the aims of the policy. They often undermine teachers' best practices; they may not improve student test scores; and they often are disproportionate to the calls of accountability. Even Diane Ravitch (1995), an ardent supporter of the standards and accountability movement through the 1990s, writes that the current manifestation of accountability "has unleashed an unhealthy obsession with standardized testing that has reduced the time available for teaching other important subjects" (Ravitch 2007, p. A25). This "obsession" may be a result of living inside the tensions of accountability and attempting to resolve irreconcilable policy messages.

The Ramifications of Paradox

Throughout this study, I uncover teachers' experiences with inconsistencies and tensions that arise from the accountability system. This study underscores how teachers meander through a series of policy contradictions. Because of the difficulty of living within the paradoxes of accountability, teachers' responses to messages they receive are often, themselves, contradictory. On the surface, the focus on the HSA is a reasonable response to a nearly thirty-nine percent failure rate on the biology section of the HSA in the 2005-2006 school year. Thirty-nine percent of the high school's students would

¹ Taken from the Halbert High School Maryland High School Assessment Biology Improvement/Intervention Plan.

certainly be an unacceptable number to not graduate. Teachers' fears of student failure are real and valid. The accountability system has created a real crisis at HHS, which may call for drastic measures.

Before examining the response to the high-stakes test, it is important to underscore that the HSA is, itself, a result of a paradox embedded in the current accountability system. The superficial level at which the HSA tests biology may be another unintended ramification of the current accountability system. *No Child Left Behind* may induce states to develop easy assessments unwittingly. The legislation lays out penalties for schools and districts that are based on student failure, but it leaves the development and administration of assessments up to the states (U.S. Congress, 2002a). Therefore, states may develop easier tests so more students can pass, which would help schools and districts avoid corrective action based on assessment results. In fact, students who make significant gains on the state developed tests in reading and math do not score higher on the NAEP (Balfanz, Legters, West, & Weber, 2007). Thus, the HSA is likely to be a source of contradictions because it is a response to the paradoxical prescriptions of *No Child Left Behind*, not a rationally designed evaluation strategy for high school biology courses.

Perhaps, the HSA's paradoxical origin partly accounts for the potential crisis caused by a thirty-nine percent passing rate on the biology HSA. However, are the measures that HHS is taking an appropriate response to the crisis, or do they contribute to the policy contradictions? Should teachers abandon their own best practices in an effort to improve student test scores? One of the more powerful, and perhaps the most damaging, consequence of the current manifestation of accountability is that the system

promotes a focus on testing over learning. This consequence, of course, is not part of the design of *No Child Left Behind*, but it is a result of a series of contradictions and misassumptions. The logic of accountability is linear. It follows Tyler's (1949) technical curricular model. Teachers input information into students who then output the information onto an exam. The current system is predicated on the assumption that a single summative exam can measure students' learning. Further, accountability presumes that if students learn they will perform well on a single, narrowly constructed test. In reality, often both of these are false assumptions. Learning is chaotic, abstract, and certainly non-linear. Evaluation of students (and teachers) must reflect the teaching and learning process. Therefore, evaluation needs to be multi-dimensional, ongoing, and based on best practices.

Because teachers are well aware (perhaps instinctively) of these misassumptions, they bypass the false linear process on which accountability is predicated. Rather than focusing on student learning, they focus on the test. Teachers understand that teaching and learning result from complicated connections and interactions between teachers and students (Ellsworth, 1997). Because accountability largely ignores these complexities, teachers attempt to bypass them and focus directly on the test. In other words, they focus on the testing process, rather than the learning process. This may be why they abandon their best practices. They are good for learning, not test preparation. In the long run a focus on learning likely would raise test scores. The current system, however, measures success according to *annual* progress. This unreasonable mandate of *No Child Left Behind* induces schools and teachers to cut corners and focus on test preparation. In the

meantime, by abandoning best practices and focusing their pedagogy on test preparation, teachers may be doing a tremendous disservice to their students.

In order to help the thirty-nine percent of students who failed the biology HSA, teachers need to expand what Ellsworth (1997) calls their pedagogical toolbox. Submission to the HSA removes tools from a teacher's toolbox. In other words, the HSA seems to limit teachers' pedagogical options. As a result, rather than including more students in the learning process, teaching to the HSA excludes students who need teachers to incorporate diverse pedagogical techniques into the curriculum in order to succeed in school. As Ms. Calypso points out, such limited pedagogy particularly may be detrimental to students with special needs. At a time when they should be using all their best practices to reach more students, and even expanding their pedagogical repertoire, teachers actually are retreating into a narrow technical paradigm of teaching and learning.

Rather than addressing the real crisis of students falling behind, teachers may be perpetuating the inequality of the status quo. In the words of Ms. Calypso, many of the students who fail the biology HSA

need inquiry. They need diversification. They cannot just have more of the same. They cannot just sit there and listen to lecture for an hour. It won't work. They will be even more turned off from school than they already are. Especially in science. They need teachers who recognize their learning needs. Teachers who can address their issues. That's what *No Child Left Behind* should do for students.

Ms. Calypso highlights the irony of teachers' responses to accountability pressures.

Although their intentions are good in the face of a real crisis of student failure that needs to be addressed, the actual manifestations of teachers' intentions may be inappropriate for improving student learning, particularly in the case of the most struggling students.

Further, this reaction actually may increase the amount of stress and pressure that

teachers experience because they behave contrary to their best instincts. In other words, their best intentions cause teachers to forgo their own understandings of quality teaching, an action likely to be quite stressful.

Implications of Teachers' Negotiations Through Accountability

Now that I have uncovered some of the paradoxes and contradictions of high-stakes accountability, I turn to an exploration of the implications of this study for science teachers, schools, and the accountability system. Now is a difficult time for teachers. Teachers find themselves negotiating through a system that is “frighteningly anti-intellectual and antidemocratic” (Noddings, 2007, p. vii). Ingersoll (2003) reports that while the demand for teachers in all subject areas is rising, the annual turnover of teachers is around fifteen percent. The turnover is greatest in math and science. In terms of new teachers, Ingersoll finds that between forty and fifty percent leave the profession within five years. Interestingly, the data show that staffing actions like retirement, terminations, and school closings only account for about twenty percent of departures. Job dissatisfaction and desire to pursue a better career account for nearly fifty percent of departures. Thus, organizational conditions in schools are by far the greatest reason for teacher turnover. The three teachers who left the biology department at HHS were all dissatisfied with their work conditions. One more is looking forward to leaving at the end of the 2007-2008 school year.

Thus, schools are like revolving doors. Teacher recruitment alone cannot meet the staffing needs of schools because so many teachers leave dissatisfied with their career. Interestingly, the data that Ingersoll uses predate *No Child Left Behind*. The legislations' rigid definition of “highly-qualified” teachers make it more difficult to recruit teachers,

and the paradoxes and contradictions that make up the reality of working in the education system only likely will increase teacher attrition due to discontent. This study's findings certainly support such a claim. In order to improve satisfaction, teachers must be better prepared to work in the current system. In this section, I address some ways to improve the reality for teachers working within the paradoxical dimensions of the current high-stakes accountability system.

Reframing the Accountability System

High-stakes accountability cannot be ignored. It is pervasive throughout the system of education. Everyone involved feels it and reacts to it. Teachers are likely not the only educational actors to be living inside the tensions, but the current system seems to promote isolation. Each actor feels alone navigating his or her way through a barrage of paradoxes. Teachers are commodified into individuals who are measured and compared to one another according to numerical test data (DeLossovoy & McLaren, 2003). In order to improve the work of teachers, an obvious prerequisite for improving education, teachers and administrators need to be brought out of isolation. They must be provided supports that will help them respond to the unique challenges posed by the current system. Professional development is one way to support teachers within the current system.

Thus far, professional development for teachers and administrators likely has promulgated contradictions, rather than helping teachers navigate through them. Professional development activities often focus on test preparatory pedagogy, test data analysis, and intervention strategies for improving test scores. This type of professional development prepares teachers to respond uncritically to the accountability system. Thus,

it propagates the system's contradictions, but it does not help teachers cope with them. Professional development that ignores accountability is also inappropriate because it ignores a major reality of working in the system. If accountability is not addressed, teachers are likely to ignore what they learn because it does not attend to their reality. Thus, teacher preparation that may otherwise be useful would simply be a waste of time.

In order to prepare teachers to work with accountability, rather than be strangled by it, professional development activities should foster a critical engagement with the accountability system. Teachers and administrators should be encouraged to use accountability and the high-stakes testing as a starting point for improving education. In this study, I found the ethos of *No Child Left Behind*, that no child be left behind, an inspiration to all six teachers, yet the realities of the policy depressed their work. This paradox is unacceptable to the stated aims of the policy. Teachers must not fear the policy. On the contrary, if the policy's aims are to be achieved, accountability should be an energy source for education. Thus, professional development should encourage teachers and administrators to look beyond the test results. They must be convinced that they are held accountable for student learning, not only students' test scores. If accountability is reframed in this way, teachers will be encouraged to improve their practice and pedagogy, rather than being induced to abandon their best practices for a focus on test preparation.

Another way to bring teachers out of isolation is to engage parents and other community members, which is another unrealized intention of *No Child Left Behind*. Parent involvement in education is a central tenet of the legislation (U.S. Congress, 2002b). One important indicator in which the policy has failed to engage parents is that

only about one percent of eligible students transfer from “failing” schools, and fewer than twenty percent of eligible students receive Supplemental Educational Services (Ravitch, 2007). Both of these options require parents to make educational choices for their children. Further, Henderson and Berla (1996) argue that the language of accountability places a wedge between families and schools because outsiders are unfamiliar with the new lexicon. Ironically, while *No Child Left Behind* may mandate the publication and dissemination of an unprecedented amount of data on schools, the community at large may have difficulty understanding its meaning. If parents are not educated about their choices and meaningfully engaged in the educational process, many will remain at the sidelines.

Misunderstandings serve to magnify the detrimental effects of accountability. The larger community, including parents, does not understand the complexities of the current accountability system that are outlined in a one thousand page document. Non-educators mainly understand a few basic concepts that are actually misconceptions. A high-stakes test measures learning. If students fail the test, they are not learning. If enough students do not learn, the school is “failing.” Further, “failure” may even have economic ramifications. Poor test results may negatively influence property values inducing members of the community to intervene to demand higher test scores, but not necessarily improved student learning.

Such narrow understandings of accountability are only likely to promote fear and the obsessive focus on test scores associated with current manifestation of the system. If districts, schools, administrators, and teachers think that they are evaluated solely according to student test scores, they are likely to focus pedagogy on improving scores,

which is exactly what is currently happening. Meaningful parent and community engagement can alleviate some of the pressures that stem from ignorance of the policy. If non-educators are helped to realize that success cannot be measured by a single summative assessment, they are likely to change their perspective of teachers and schools. Consequently, educators would feel less subservient to the high-stakes test. They may still consider it an important part of students' evaluation, but they would feel supported to teach beyond the test, rather than to it.

Noddings (2007) argues that teaching responsibility “is the fundamental concept on which any reasonable concept of accountability must be built” (p. 38). A teacher's responsibility is to develop appropriate content and standards for each student or group of students. Teachers may fail to live up to their responsibilities, but accountability should abolish this failure, not teachers' responsibilities. In other words, rather than de-skilling teachers, accountability should encourage teachers to be experts. Such a reframing of the purpose of accountability would change the political energy around the policy. Teachers would be encouraged to see themselves as experts in teaching, not test preparation. They would be motivated to improve their practice. Rather than being a repressive educational albatross that attempts to control educational actors, it may help motivate and encourage educators to work toward the ethos of *No Child Left Behind*.

Researching the Accountability System

High-stakes accountability also has penetrated the research community. Researchers have responded to the prevalence of accountability mandates. Although few studies have focused on the effects of the current version of accountability on science education, over the last few years researchers have studied accountability extensively in

math and language arts education. This study's most significant contribution to the current literature is its focus on science teaching. Because *No Child Left Behind* is mandating science testing for the first time for the 2007-2008 school year, the proliferation of science accountability systems has lagged behind math and language arts. As such, research on accountability in science education has been limited. In this sense, Yin (1981) would describe this study as exploratory. In other words, at the onset of this research, I had some basic assumptions about the affect of accountability on science education based on research in math and language arts, but I did not know exactly how my assumptions would apply in science. As such, some of this study's findings that are germane to science education may serve as a basis for future research on accountability in science education.

Specifically, my findings that high-stakes testing disaggregates science content into superficial and isolated knowledge bits needs to be examined further. Do all accountability systems test isolated knowledge? How can large scale accountability systems be induced to incorporate evaluation methods, such as concept mapping and students' observation, that promote integration and depth? Another important consideration for future research is the hierarchy of knowledge created by the presence of a high-stakes test. Do all states focus on the same biology content as the biology HSA? Why do test writers focus on certain topics? How can the tests be more reflective of what scientists and science educators consider to be important biology knowledge?

Maryland's exclusive focus on biology limited this study to biology. Because each discipline in science carries its own set of assumptions, similar research should be done with physical science teachers in states where physical science is included on the

test. A comparative study between physical science teaching and biology teaching in Maryland may uncover additional influences of high-stakes accountability that were missed. It would be interesting to explore how a science that is not on the HSA is affected by accountability. Future research should also examine the cohort of students who went through elementary and middle school with almost no exposure to science. Because *No Child Left Behind* did not mandate science testing until the 2006-2007 school year, elementary schools only devoted an average of six percent of classroom time to science (NICHHD, 2005). What will happen with these students as they go through middle school and high school science?

Further, although much attention has been allotted to developing large-scale evaluation methods for scientific inquiry-based instruction (Doran, Boorman, Chan, & Hejaily, 1993; Hickey, DeCuir, Hand, Kyser, Laprocina, & Mordica, 2002; Zachos, Hick, Doane, & Sargent, 2000), this study's findings suggest that researchers may need to take a step back and focus on reconciling accountability with inquiry. In other words, the current manifestation of accountability may be incompatible with inquiry-based instruction because one subscribes to positivism while the other subscribes to constructivism. If scientific inquiry is to fit into an accountability system, research needs to focus on reconstructing accountability to comport with the constructivist paradigm.

Similarly, further research must be done on how to link standards, pedagogy, and assessments better. This study found that science standards may be well aligned at the federal, state, and district levels, but they have little influence on the classroom. On the other hand, pedagogy is driven by a high-stakes test that poorly reflects the standards and the district's curricular intentions. Therefore, research on alignment of assessments and

standards is insufficient because alignment exists outside the classroom. In many ways, it is pedagogically meaningless. Instead, research should focus on how standards, pedagogy, and evaluation can become more integrated in the classroom.

In addition to providing insights specific to science education, this study's focus on teachers presents insights into how educational accountability influences the classroom. Since teachers are the policy's ultimate implementers, their practice is integral to the success or failure of *No Child Left Behind*. Valli and Buese (2007) argue that "If policy expectations for teacher role change had benefited students, one could argue that the toll on teachers, although unfortunate, was for the greater good of students. But that did not seem to be the case" (p. 520). This study contributes to the growing body of research that suggests that high-stakes accountability has strained educators with minimal benefit, if not significant detriment, to students. Further research should be conducted on how to reframe accountability to encourage teachers to look beyond single, narrow measures of success and to engage in quality pedagogy. Accountability should not control educators through fear. This study demonstrates that the current accountability system may cause teachers to retreat from pedagogical innovation into test preparatory pedagogy. On the contrary, the mechanisms of accountability should promote reform where it is needed. When accountability promotes optimism, responsibility, job satisfaction, avenues for developing pedagogical expertise, and collaboration between teachers and administrators, high-stakes accountability is more likely to improve education for all students.

The Future of Accountability

In the near term, high-stakes accountability is likely to expand, especially if *No Child Left Behind* is reauthorized as planned in 2007, or debate on it ends in a stalemate until the next administration. This school year (2007-2008) is the first year where the law orders that science be tested at least once in grade spans 3-5, 6-9, and 10-12. Maryland, along with twenty-three other states, will soon mandate passing an exit exam in order to graduate from high school (CEP, 2005). Further, as each year draws closer to 2014 where every student in the United States is supposed to be “proficient” in reading and math, it becomes more difficult for schools to make AYP because an increasing number of students must pass the high-stakes exam. Therefore, districts and schools are likely to increase monitoring and pressure in an attempt to stay above the rising AYP tide. A piece of evidence for this increase in accountability at Halbert High School is the difference between the 2005-2006 and the 2006-2007 biology HSA intervention plan. In the 2005-2006 school year, Ms. Victoria developed what she calls a “general strategy,” but for the 2006-2007 school year she wrote a “much more involved” five page intervention plan with five intervention programs and forty-two additional strategies HHS can implement to raise HSA scores. Because Ms. Victoria left HHS, I do not have access to the intervention plan for 2007-2008, but it is likely to be very involved since all of the ESOL and special education students will now be taking biology rather than environmental science. Thus, some of the school’s most needy students need to be targeted for “intervention.”

In the slightly longer-term, some changes are likely to be made to the current accountability system because as people are becoming increasingly aware of some of the

tensions and contradictions embedded in the policy, resistance to *No Child Left Behind* continues to mount. For example, the National Education Association (NEA) has filed a lawsuit against the U.S. Department of Education (Chanin, Collins, & Pollard, 2005). The NEA and school districts in Vermont, Michigan, and Texas claim that the U.S. Department of Education has withheld funds illegally from school districts because it is unwilling to spend its own money on expenses resulting from the implementation of accountability measures. Although the case was dismissed in 2005, the plaintiffs have filed an appeal for which six additional states and the District of Columbia have filed amicus briefs supporting the NEA's lawsuit (NEA, 2007a).

In terms of the reauthorization, the NEA (2007b) would like the U.S. Congress to make three fundamental changes to the legislation. They want states to use more than test scores to measure student learning and school performance, reduce class size, and increase the number of "highly-qualified" teachers in schools. This study certainly demonstrates the need for the NEA's first recommended change. A single measure of student learning results in a narrow and fragmented curriculum that is focused on the test, and it creates a tremendous amount of pressure on teachers to prepare their students for a high-stakes test and for students to perform well on a single summative measure. The NEA's second and third recommendations may be worthwhile, but they are difficult to implement. Reducing class size and increasing the number of "highly-qualified" teachers places a financial burden on districts that are already unable to finance accountability. Further, under the current legislation, schools already have difficulty recruiting "highly-qualified" teachers. In the 2007-2008 school year, HHS has four new biology teachers, at least one of whom is not certified.

In addition to adopting the NEA's recommendations, Congress should repeal the universal proficiency requirement. It is both unreasonable and immoral to impose an impossible mandate on the education system. As the 2014 deadline draws nearer, an increasing number of schools will fail to make AYP. Accountability should measure progress toward improvement, not impose artificial mandates that sound good politically, but are impossible to achieve in practice. Interestingly, AYP may not even measure student performance. Whether schools achieve AYP actually may correlate more with varying implementation of *No Child Left Behind* across states than student achievement (Balfanz, Legters, West, & Weber, 2007). As such, AYP should be replaced with a system that requires schools to show improvement that is appropriate for their particular context. Further, rather than the federal government imposing threats and penalties on schools, states and districts should decide what to do with underperforming schools. They should help schools develop customized plans that lead to improvement in a reasonable amount of time. Equality is not sameness. The same prescriptions can not help different students in different contexts.

Although the NEA is one of the more high profile opponents to the current version of high-stakes accountability, many others are joining the opposition. According to the Communities for Quality Education, an advocacy group that tracks state actions on *No Child Left Behind*, all fifty states have taken at least some action against the shortcomings of the legislation (CQE, 2007). Perhaps, some of these states have realized that the drastic measures called for by the legislation are not worth six percent of a district's education budget, especially considering that over the last five years *No Child*

Left Behind has cost \$40 billion more than the federal government has allotted for its implementation (Vu, 2007).

In an effort to promote the reauthorization of the legislation, the U.S. Department of Education (2007) claims that *No Child Left Behind* has drawn attention to the achievement gap between white students and their minority peers, a claim that this research validates. Further, supporters of high-stakes accountability highlight that fourth graders performed better on the National Assessment of Educational Progress (NAEP) exam in 2005 than they have in the past thirty years, and the achievement gap in reading between White and African American students narrowed (NAEP, 2005). The results for eighth grade and math, however, showed no significant improvement.

Interestingly, both the proponents of reauthorization and many of its opponents lack a critical perspective of accountability. Proponents who cite improvements in test scores assume that tests are a solid measure of students' learning and achievement. Further, although more attention may be given to minority students, is the attention necessarily constructive and beneficial to their academic success? Opponents like the NEA who challenge *No Child Left Behind* on the basis of funding or teacher shortages may not address some of the underlying issues that this study has uncovered. By focusing on particular details, the current discourse about accountability may miss the larger curricular and pedagogical issues unearthed by this research.

On September 5, 2007, in an effort to promote the reauthorization of *No Child Left Behind*, Secretary Spellings said, "The beauty of *NCLB* is that it provides straightforward, unvarnished information on how students are doing" (USDE, 2007, ¶ 24). Perhaps the trouble with this pronouncement is the absence of critical implications.

The legislation has produced volumes of data and information on students, teachers, schools, districts, and states. That information is being used by advocates and opponents of accountability alike. But is the information meaningful? Does it truly capture teaching and learning? Does it meaningfully evaluate students, teachers, and schools? The findings of this study suggest that data collected by accountability may be of little value because of how significantly high-stakes testing has penetrated and transformed the curriculum. If accountability has influenced the curriculum so significantly, do the data report on teaching and learning or teaching and learning in accountability? Do they reflect students' knowledge and understandings or their ability to perform well on an assessment? If the latter is true, the data may need to be reconsidered in light of what is considered to be quality teaching and learning.

As a society, do we want to teach our students biology or HSA biology? Do we want students to learn the meaning of "BCR" before "essay?" Do we want them to hear about high-stakes exams in nearly every class period? If the discourse about accountability remains uncritical and does not address these underlying questions and the other themes uncovered in this research, the next reauthorization of the Elementary and Secondary Education Act (whether it be under the Bush Administration or its successor) likely will lack substantive change that really can improve teaching and learning. Without a critical examination of high-stakes accountability and its influence on education, teachers are likely to continue to traverse the tensions and contradictions between accountability policies and the lived curriculum without meaningful pedagogical change.

Summary

The three themes developed in this study uncover some of the complexities and contradictions that biology teachers experience in the current educational climate of high-stakes accountability. This study uncovered tensions at three levels. First, I explored the disagreement between the high-stakes test and the national and state science standards. Then, I examined the tensions between accountability mandates and teachers' understandings of quality instruction. Finally, I explored some of the ways in which high-stakes accountability has influenced biology teachers' practice. Specifically, I focused on pedagogy, teachers' roles in the classroom, their understanding of science, and teachers' passion for their profession.

Future research should focus on how this study's findings may differ in other accountability systems, especially those that include the physical sciences in their high-stakes assessment. Research also should focus on the how four years of excluding science from the curriculum has influenced the curriculum now that science testing is mandated by *No Child Left Behind*. In general, research should explore ways to reframe accountability from a system that controls teachers through pressure and fear to one that inspires teachers to improve their practice.

In light of this study's findings, the current manifestation of high-stakes accountability as legislated by *No Child Left Behind* is riddled with paradoxes and inconsistencies at all levels of the educational system. These paradoxes have a significant influence over the curriculum and teachers' work. As Congress and advocacy groups battle over the reauthorization of the legislation, they may not be adequately attentive to all the inconsistencies. Ironically, they are focused on narrow indicators like a lack of

finances, teacher shortages, inappropriate assessment strategies, overrepresentation of minority and special education students in developing AYP criteria, and inadequate resources for schools. Although each of these issues is significant and important, without a comprehensive discourse about the meaning and aims of accountability and quality teaching and learning, the many tensions and contradictions currently embedded in the system are likely to persist.

Appendix A: Letter of Invitation

Isaak Aronson
2700 Connecticut Avenue # 701
Washington, DC 20008
(202) 248-8944
iaronson@umd.edu

November, 2006

Dear Teacher:

I am writing to invite you to participate in a study that examines science educational accountability as mandated by the *No Child Left Behind Act*. I am conducting this study as a doctoral student in the Department of Education Policy and Leadership at the University of Maryland, College Park under the direction of Dr. Francine Hultgren.

The purpose of this case study is to explore how biology teachers understand and construct their practice in a high-stakes accountability. As I seek to understand science teaching under the umbrella of high-stakes testing, I will conduct approximately three weeks of classroom observations and two interviews that I will tape-record and transcribe. All interviews and observations will be reported anonymously. You will not be identified by name in the published findings or in oral presentations, unless you choose to have your name revealed. You will, however, be invited to adopt an alias for the purposes of my writing. After the research is complete, you are welcome to review the results.

This study will make an important contribution to understanding science educational accountability as mandated by *No Child Left Behind* and the policy's effects on the science curriculum. This study will be considered successful if it offers policy makers and implementers insights into how science educational accountability influences classroom practice. The pedagogical insights derived from this study should help enhance science teaching and learning.

I am interested in setting up initial interviews for late January or early February 2007. If you have any questions and/or would like to be one of the teacher-participants in this study, please contact me. Thank you, in advance, for your consideration.

Sincerely,

Isaak Aronson
Doctoral Candidate
Education Policy and Leadership
University of Maryland, College Park

Appendix B: Data Collection Timeline

January 2007	Document Review <ul style="list-style-type: none">– national and state standards documents– published biology HSAs– district curricular frameworks and blueprints
February 2007	Document Review <ul style="list-style-type: none">– national and state standards documents– published biology HSAs– district curricular frameworks and blueprints Pilot Interview (February 23)
March 2007	1 st Interviews (March 9 –March 16) Classroom Observations (begin March 16) Document Review <ul style="list-style-type: none">– school based policy documents– school based curriculum documents– county curriculum modules
April 2007	Classroom Observations Document Review <ul style="list-style-type: none">– school based policy documents– school based curriculum documents– county curriculum modules
May 2007	Classroom Observations Document Review <ul style="list-style-type: none">– school based policy documents– school based curriculum documents– county curriculum modules 2 nd Interviews (Ms. Lydia, Ms. Harris, Dr. Stevens)
June 2007	Classroom Observations (end June 4) 2 nd Interviews (Ms. Victoria, Ms. Khana, Ms. Calypso) Document Review <ul style="list-style-type: none">– national and state standards documents– published biology HSAs

Appendix C: Document Review Guide¹

The following list of questions and sub-questions will guide my document review:

- I. What situation resulted in the development of the document?
 - a. Who was involved in the development of the documents? What were their names, with what institution were they affiliated, and what were their respective roles in the project?
 - b. To what social, political, economic, or educational problem was the document attempting to respond?
 - c. What planning elements dominated the development process?
- II. What are the purposes of the document?
 - a. Who is the intended audience of the document?
 - b. At what level(s), if at all, does the document express its purpose?
 - c. What educational goals and aims are emphasized and what are their relative priorities?
 - d. What types of learning objectives are included and emphasized in the document?
 - e. Does the document hold a critical or reconstructionist perspective?
- III. What assumptions underlie the document's organization and approach to purpose or content?
 - a. What conceptions of learning, objectives, curriculum, and teaching underlie the document?
 - b. What aspects of a lived curriculum are likely to accompany the conceptions and perspectives espoused by the document?
 - c. To what extent is the document likely to play a hegemonic role in its purposes or content?
 - d. What epistemological assumptions, if any, are propagated by the document?
- IV. What are the ramifications of the document's use?
 - a. To what extent will the document be consistent with and appropriate for teachers' attitudes, beliefs, and competencies?
 - b. To what extent does the document take into account different needs of students?
 - c. What are the document's strengths?
 - d. What are the document's weaknesses?
 - e. To what extent is the document aligned with the other documents being analyzed?

¹ Adapted from Table 1.5 Curriculum Analysis Questions (Posner, 2004, pp. 20-22)

Appendix D: Informed Consent Form

Page 1 of 2
Initials _____ Date _____

Project Title	NEGOTIATING THE TERRAIN OF ACCOUNTABILITY IN SCIENCE TEACHING
Why is this research being done?	This is a research project being conducted by Isaak Aronson at the University of Maryland, College Park. We are inviting you to participate in this research project because you are a high school biology teacher in Montgomery County, Maryland. The purpose of this study is to explore how biology teachers understand and construct their practice in a high-stakes accountability environment.
What will I be asked to do?	<p>The procedures involve interviews and observations, which will provide text for analysis. The topics of focus for the interviews and observations will include: how science accountability effects teaching, teachers' understanding of the discipline of science, teachers' understanding of their role in the classroom, and teachers' passion for their profession.</p> <ul style="list-style-type: none"> • I understand that I will engage in approximately two tape-recorded and transcribed interviews, approximately three months apart, each approximately forty-five minutes in length; • I understand that my class sessions will be observed for approximately three weeks.
What about confidentiality?	<p>We will do our best to keep your personal information confidential. To help protect your confidentiality, notes, transcripts, and cassette tapes will be accessible only to the researcher and kept in a locked cabinet in his residence. At the completion of this study, the tapes will be erased, and written records will be shredded.</p> <p>You will only be referred to in the final document by first name or an alternative name, should you prefer. Your identity will be protected to the maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.</p>
What are the risks of this research?	There are no known risks associated with participating in this research project.
What are the benefits of this research?	The results of this research may help enhance your science teaching and students' learning. Additionally, it may offer policy makers and implementers insights into how science educational accountability influences classroom practice.

Project Title	NEGOTIATING THE TERRAIN OF ACCOUNTABILITY IN SCIENCE TEACHING	
Do I have to be in this study? May I stop participating at any time?	Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify.	
What if I have questions?	<p>If you have any questions about the research study itself, please contact:</p> <p style="padding-left: 40px;">Isaak Aronson 2700 Connecticut Avenue #701 Washington, DC 20008 iaronson@umd.edu 202-248-8944</p> <p>This research is being conducted by:</p> <p style="padding-left: 40px;">Dr. Francine Hultgren Department of Education Policy and Leadership University of Maryland College Park, MD 20742 fh@umd.edu 301-405-4562</p> <p>If you have any questions about your rights as a research subject or wish to report a research-related injury, please contact: Institutional Review Board Office, University of Maryland, College park, Maryland, 20742; (e-mail) irb@deans.umd.edu; (telephone) 301-405-0678</p> <p>This research has been reviewed according to the University of Maryland, College Park IRB procedures for researching human subjects.</p>	
Statement of age of participant and consent.	<i>Your signature indicates that: you are at least 18 years of age; the research has been explained to you; your questions have been fully answered; you freely and voluntarily choose to participate in this research project.</i>	
Signature and Date	NAME OF PARTICIPANT	
	SIGNATURE OF PARTICIPANT	
	DATE	

Appendix E: Interview Guide

Introduction of the Study to the Teachers

This study is about teachers' experiences with science accountability. I have already conducted a document review of relevant science education policy documents at the national, state, and district level. However, as *No Child Left Behind* seeks to mandate particular educational outcomes by influencing teaching and the curriculum, it is critical to understand how teachers are experiencing the implementation of the legislation. In other words, it is not only important to know what the policy says, but also what the teachers hear and how they respond.

The interview process is broken up into two parts. First, I will interview teachers prior to conducting any classroom observations. Then, after the observations, I will conduct the second interview to follow up on my observations, as well as to dig deeper into their experiences with teaching science within the current accountability climate.

I will touch on four topics:

- Background information in order to become acquainted with the teachers and their approach to teaching.
- Teachers' interest and passion for teaching science.
- The meaning and importance of science.
- Teaching practice and how it is influenced by *No Child Left Behind*.

The purpose of the first interview is to get to know the teacher and to get an overview of his or her experiences with science accountability. In the first interview, I will begin by getting to know why and how teachers came to teach science. Then, I will probe into why teachers consider science to be important for both students and society as a whole. At the end of the first interview, I will seek to gain a general understanding of how teachers view *No Child Left Behind*. For the second interview, I will seek to probe further into how teachers experience science accountability and the tensions that they perceive in the policy's mandates. I may also use the second interview as an opportunity to follow-up about aspects of the curriculum that I observed in the classroom.

The tone of the interviews is conversational. I would like teachers to elaborate as much as possible on the questions or concepts that I raise. There are no right or wrong answers. I simply seek to learn as much as possible about teachers' experiences with teaching science in the current educational accountability climate. I will start with some orienting questions in order to get to know you as a teacher.

First Interview

Introduction:

Background information

1) Please describe your professional background and experiences.

Potential probes:

- a. How long have you been teaching?
- b. How long have you been teaching biology?
- c. Have you taught other subjects? If so, how did you find yourself teaching biology?
- d. What is your educational background?
- e. Are you certified in biology?
- f. Do you work with students in any other capacity?
- g. Do you belong to any professional organizations? What is your level of involvement?

2) Please explain why you chose teaching as a career.

Potential probes:

- a. How did you come in to teaching?
- b. How would you define teaching?
- c. Do you hold a particular philosophy about teaching?

(Bridge): Now that we have discussed your professional experiences and perspectives, I would like to focus on science teaching in particular.

Passion for teaching science

1) Please explain why you teach science.

Potential probes:

- a. Is there something in particular that you enjoy the most about teaching science?
- b. How, if at all, has your interest in teaching science changed over the years?
- c. What is your least favorite aspect of teaching science?
- d. Would changing any school or district policies make you enjoy teaching science more?

(Bridge): Now that we have discussed teaching science, I would like to explore the meaning of science and its role in our society.

Meaning of science

1) Both national and state science standards highlight the importance of science in our society. In your view, what is the role of science in our society?

Potential probes:

- a. How is it integrated into our lives?

- b. Why is it important for non-scientists to understand science?
- 2) What are some of the most common misconceptions about science? How do you dispel them in your class?
 - 3) In the Benchmarks for Science Literacy it says, “In a culture increasingly pervaded by science, mathematics, and technology, scientific literacy requires understandings and habits of mind that enable citizens to grasp what those enterprises are.” What does scientific literacy mean to you?
 - 4) The National Science Education Standards call for a creative approach to teaching science. What does creativity in science mean? What examples might you give of creative approaches in your teaching?

(Bridge): After getting an overview of your perspectives on teaching and the meaning of science, I would like to get some greater detail on your particular teaching style and practice.

Teaching Practice

- 1) How would you define your teaching style?
- 2) Have you noticed differences in how students learn science? How do you address those differences?
- 3) The National Science Education Standards and Maryland State standards call for a greater focus on scientific inquiry based instruction. What does scientific inquiry based instruction mean to you?

Potential probes:

- a. Is it effective at helping most students learn science?
- b. Do you practice it? Please share some examples of what you do.
- c. In your view, what is the purpose of laboratory work in the classroom?
- d. What is the role of the teacher during laboratory work?

(Bridge): Now, I would like to explore how accountability has influenced your teaching.

Tensions between accountability and personal constructions of teaching

- 1) What is your view of *No Child Left Behind*?

Potential probes:

- a. What are some of its positive aspects?
- b. What are some of the legislation’s challenges?
- c. Has it changed your thinking in the classroom?
- d. Does *No Child Left Behind* affect you?

- e. Does *No Child Left Behind* affect your students?
- 2) Is there anything else that you would like to add about your experiences with teaching science and/or the mandates of *No Child Left Behind*?

Second Interview

Introduction: Now that I have been able to observe your work, I would like to follow up with some more poignant questions about your teaching within the current climate of educational accountability.

Tensions between accountability and personal constructions of teaching

(Bridge): I will start by exploring potential factors that may influence your teaching practice.

- 1) Please describe how your beliefs, background, and previous experiences influence your teaching.

Potential probes:

- a. How does your educational background inform your practice or beliefs about teaching?
- b. Have previous professional development activities informed your practice or beliefs about teaching?
- c. How does your previous classroom experience inform your practice or beliefs about teaching?
- d. How, if at all, does your relationship with colleagues inform your practice or beliefs about teaching?

- 2) Has the No Child Left Behind Policy affected what goes on in your classroom? If so, please describe how it has influenced your teaching.

Potential probes:

- a. Has the HSA had an effect on the way that you teach?
- b. Have you been influenced by national or state standards?
- c. What are some of your experiences with high-stakes testing? Provide some examples.
- d. Have your efforts at scientific inquiry-based instruction been influenced by accountability policies? If so, how?
- e. Is scientific inquiry-based instruction compatible with the mandates of *No Child Left Behind*?
- f. Has *No Child Left Behind* influenced your views and constructions of the science curriculum? If so, how?

- g. Has *No Child Left Behind* influenced how you view or work with your students? Any particular group of students? If so, how?

(Bridge): Now I would like to explore whether *No Child Left Behind* influences how you feel about teaching.

- 3) Do you experience any tensions between the mandates of educational accountability and your beliefs about teaching? If so, might you describe them?

Potential probes

- a. Does your current teaching style follow your beliefs about teaching?
- b. Do you sometimes accommodate your teaching to the mandates of the High School Assessment? If so, how?
- c. How would you compare an exam that you would write to the county exam?
- d. Several teachers mentioned that they have been conducting fewer labs. What is your experience with the amount of labs that you feel are required in your teaching?

(Bridge): Now, I would like to turn to a discussion about your views of *No Child Left Behind* and its influence on science teaching.

Accountability and science teaching

- 1) The Benchmarks for Science literacy argue that common goals do not require a common curriculum. Has *No Child Left Behind* influenced this relationship? If so, how?
- 2) How, if at all, has *No Child Left Behind* influenced the school's climate?

Potential probes:

- a. How, if at all, has it influenced your relationship with your fellow teachers?
 - b. How, if at all, has it influenced your relationship with administrators?
 - c. Do you feel included in curriculum or policy changes in your school?
 - d. Will some of your students fail the HSA? On what basis do you make that judgment?
- 3) Do you feel like you were prepared for the inclusion of biology on the High School Assessment? Why or why not?

Potential probe:

- a. When was the first time you heard the term "BCR"? HSA?
 - b. Why has the HSA gained such supremacy in the lexicon of the classroom?
 - c. In your opinion, how difficult is it for students to pass the HSA?
- 4) Has *No Child Left Behind* positively influenced science teaching? If so, how?

5) Has *No Child Left Behind* negatively influenced science teaching? If so, how?
Potential probe:

- a. Has it influenced your passion/interest in teaching science? If so, how?
- b. What is your greatest concern about the HSA?
- c. What is your greatest concern about student failure on the HSA?

6) Has *No Child Left Behind* influenced your understanding of the discipline of science? If so, how and why?

Potential probe:

- a. How does *No Child Left Behind* view science?

7) Is there anything that you would like to add about educational accountability and teaching?

Appendix F: No Dentist Left Behind

My dentist is great! He sends me reminders so I don't forget checkups. He uses the latest techniques based on research. He never hurts me, and I've got all my teeth. When I ran into him the other day, I was eager to see if he'd heard about the new state program. I knew he'd think it was great.

“Did you hear about the new state program to measure effectiveness of dentists with their young patients?” I said.

“No,” he said. He didn't seem too thrilled. “How will they do that?”

“It's quite simple,” I said. “They will just count the number of cavities each patient has at age 10, 14, and 18 and average that to determine a dentist's rating. Dentists will be rated as excellent, good, average, below average, and unsatisfactory. That way parents will know which are the best dentists. The plan will also encourage the less effective dentists to get better,” I said. “Poor dentists who don't improve could lose their licenses to practice.”

“That's terrible,” he said.

“What? That's not a good attitude,” I said. “Don't you think we should try to improve children's dental health in this state?”

“Sure I do,” he said, “but that's not a fair way to determine who is practicing good dentistry.”

“Why not?” I said. “It makes perfect sense to me.”

“Well, it's so obvious,” he said. “Don't you see that dentists don't all work with the same clientele, and that much depends on things we can't control? For example, I work in a rural area with a high percentage of patients from deprived homes, while some of my colleagues work in upper middle-class neighborhoods. Many of the parents I work with don't bring their children to see me until there is some kind of problem, and I don't get to do much preventive work. Also, many of the parents I serve let their kids eat way too much candy from an early age, unlike more educated parents who understand the relationship between sugar and decay. To top it all off, so many of my clients have well water which is untreated and has no fluoride in it. Do you have any idea how much difference early use of fluoride can make?”

“It sounds like you're making excuses,” I said. “I can't believe that you, my dentist, would be so defensive. After all, you do a great job, and you needn't fear a little accountability.”

“I am not being defensive!” he said. “My best patients are as good as anyone's,

my work is as good as anyone's, but my average cavity count is going to be higher than a lot of other dentists because I chose to work where I am needed most."

"Don't get touchy," I said.

"Touchy?" he said. His face had turned red, and from the way he was clenching and unclenching his jaws, I was afraid he was going to damage his teeth. "Try furious! In a system like this, I will end up being rated average, below average, or worse. The few educated patients I have who see these ratings may believe this so-called rating is an actual measure of my ability and proficiency as a dentist. They may leave me, and I'll be left with only the most needy patients. And my cavity average score will get even worse.

On top of that, how will I attract good dental hygienists and other excellent dentists to my practice if it is labeled below average?"

"I think you are overreacting," I said.

"Complaining, excuse-making and stonewalling won't improve dental health... I am quoting from a leading member of the DOC," I noted.

"What's the DOC?" he asked.

"It's the Dental Oversight Committee," I said, "a group made up of mostly lay persons to make sure dentistry in this state gets improved."

"Spare me," he said, "I can't believe this. Reasonable people won't buy it," he said hopefully.

The program sounded reasonable to me, so I asked, "How else would you measure good dentistry?"

"Come watch me work," he said. "Observe my processes."

"That's too complicated, expensive and time-consuming," I said. "Cavities are the bottom line, and you can't argue with the bottom line. It's an absolute measure."

"That's what I'm afraid my parents and prospective patients will think. This can't be happening," he said despairingly.

"Now, now," I said, "don't despair. The state will help you some."

"How?" he asked.

"If you receive a poor rating, they'll send a dentist who is rated excellent to help

straighten you out,” I said brightly.

“You mean,” he said, “they’ll send a dentist with a wealthy clientele to show me how to work on severe juvenile dental problems with which I have probably had much more experience? BIG HELP!”

“There you go again,” I said. “You aren’t acting professionally at all.”

“You don’t get it,” he said. “Doing this would be like grading schools and teachers on an average score made on a test of children's progress with no regard to influences outside the school, the home, the community served and stuff like that. Why would they do something so unfair to dentists? No one would ever think of doing that to schools.”

I just shook my head sadly, but he had brightened. “I’m going to write my representatives and senators,” he said. “I’ll use the school analogy. Surely they will see the point.”

References

- Adler, P.A., & Adler, P. (1994). Observational techniques. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 377-392). Thousand Oaks, CA: SAGE.
- Abelman, C., Elmore, R., Even, J., Kenyon, S., & Marshall, J. (1999). *When accountability, will anyone answer?* Philadelphia: Consortium for Policy Research in Education.
- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.
- Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). Retrieved on November 5, 2006 from <http://epaa.asu.edu/epaa/v10n18/>.
- Anderson, R. D. (1995). Curriculum reform: Dilemmas and promise. *Phi Delta Kappan*, 77(1), 33-36.
- Anderson, R. D. (1996). *Study of curriculum reform*. Washington, DC: U.S. Government Printing Office.
- Anderson, R. D., & Helms, J. V. (2001). The ideal of standards and the reality of schools: Needed research. *Journal of Research in Science Teaching*, 38(1), 3-16.
- Aoki, T. (1990). Beyond the half-life of curriculum and pedagogy. *One World*, 27(2), 3-10.
- Aoki, T. (2005). *Curriculum in a new key: Collected works of Ted T. Aoki*. Mahwah, NJ: Lawrence Erlbaum.
- Apple, M. W. (1992). Do the standards go far enough? Power, policy, and practice in mathematics education. *Journal for Research in Mathematics Education*, 23(5), 412-431.
- Apple, M. W. (1996). *Cultural politics and education*. New York: Teachers College Press.
- Apple, M. W. (2001). Markets, standards, teaching, and teacher education. *Journal of Teacher Education*, 52(3), 182-196.
- Apple, M. W. (2006). What can we learn from Texas about *No Child Left Behind?* *Education Policy*, 20 (3), 551-560.

- Apple, M. W., & Jungck, S. (1990). "You don't have to be a teacher to teach this unit." Teaching technology and gender in the classroom. *American Educational Research Journal*, 27(2), 227-251.
- Applebee, A. N. (1996). *Curriculum as conversation*. Chicago: The University of Chicago Press.
- Atkin, J. M., & House, E. R. (1981). The Federal Role in Curriculum Development, 1950-80. *Educational Evaluation and Policy Analysis*, 3(5), 5-36.
- Atkins, M. (1993). *Assessment issues in higher education*. London: Department of Employment.
- Arons, A.B. (1983). Achieving wider scientific literacy. *Daedalus*, 112(2), 91-122.
- Baxter, G., & Glaser, R. (1997). An approach to analyzing the cognitive complexity of science performance assessments. *CSE Report 452*. Los Angeles: CREST/UCLA.
- Becker, S. (1990). Generalizing from case studies. In E. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: A continuing debate* (pp. 232-242). New York: Teachers College Press.
- Bell, R. L., Blair, L. M., Crawford, B. A., & Lederman, N.G. (2003). Just do it? Impact of a science apprenticeship program on high school students' understandings of the nature of science and Scientific Inquiry. *Journal of Research in Science*, 40(5), 487-509.
- Berlak, H., Newmann, E., Adams, E., Archbald, D., Burgess, T., Raven, J., et al. (Eds.). (1992). *Toward a new science of educational testing and assessment*. Albany, NY: State University of New York Press.
- Berman, L. M., Hultgren, F. H., Lee, D., Rivkin, M. S., & Roderick, J. A. (1991). *Towards curriculum for being: Voices of educators*. Albany: State University of New York Press.
- Berryman, L., & Russell, D. (2001). Portfolios across the curriculum: Whole school assessment in Kentucky, *English Journal*, 90(6), 76-83.
- Berube, M. S. (Ed.). (1995). *Webster's II: New college dictionary*. New York: Houghton Mifflin.
- Balfanz, R., Legters, N., West, T. C., & Weber, L. M. (2007). Are NCLB's measures, incentives, and improvement strategies the right ones for the nation's low-

- performing high schools?. *American Educational Research Journal*, 44(3), 559-593.
- Black, P. (1994). Performance assessment and accountability: The experience in England and Wales. *Educational Evaluation and Assessment Policy Analysis*, 16(2), 191-203.
- Black, P., & Dylan, W. (2005). Lessons from around the world: How policies, politics and cultures constrain and afford assessment practices. *The Curriculum Journal*, 16(2), 249-261.
- Bidwell, C. E., & Yasumoto, J. Y. (1999). The collegial focus: Teaching fields, collegial relationships, and instructional practice in American high schools. *Sociology of Education*, 72(4), 234-256.
- Blumenthal, R. (2006). Why Connecticut sued the federal government over No Child Left Behind. *Harvard Educational Review*, 76(4), 564-569.
- Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal*, 42(2), 231-268.
- Boudett, K. P., Murnane, R. J., City, E., & Moody, L. (2005). Teaching educators: How to use student assessment data to improve instruction. *Phi Delta Kappan*, 86(9), 700-706.
- Bourdieu, P. (1977). *Outline of a theory of practice*. New York: Cambridge University Press.
- Bowe, R., Ball, S. J., & Gold, A. (1992). *Reforming education and changing schools: Case studies in policy sociology*. London: Routledge.
- Branford, J. D., & Donovan, S. M. (2005). Scientific inquiry and how students learn. In NRC, *How people learn science in the classroom* (pp. 397-419). Washington, DC: National Academies Press.
- Brees, K. (2003). Across the great divide: Bridging the communication chasm. *English Journal*, 92(6), 15-16.
- Browder, D. M., Karvonen, M., Davis, S., Fallin, K., & Courtade-Little, G. (2005). The impact of teacher training on state alternate assessment scores. *Exceptional Children*, 71(3), 267.
- Bruner, J. (1960). *The process of education*. Cambridge: Harvard University Press.
- Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation.

- Bybee, R. W. (1997). *Achieving scientific literacy: From purposes to practices*. Portsmouth, NH: Heinemann.
- Bybee, R. W., & Ferrini-Mundy, J. (1997). Editorial. *School Science and Mathematics*, 97(6), 281-282.
- Carr, M., McGee, C., Jones, A., McKinley, E., Bell, B., Barr, H., et al. (2000). *Strategic research: Initiative literature review: The effects of curricula and assessment on pedagogical approaches and on educational outcomes*. Wellington, New Zealand: Ministry of Education.
- Carraway, L. N. (2006). Improve scientific writing and avoid perishing. *The American Midland Naturalist*, 155(2), 383-394.
- Cawelti, G. (2006). The side effects of NCLB. *Educational Leadership*, 64(3), 64-68.
- Center for Quality Education. (2007). *State legislative summaries*. Retrieved on September 12, 2007 from <http://www.qualityednow.org/summaries/states/index.php>.
- Center on Education Policy. (2005). *States try harder, but gaps persist: High school exit exams 2005*. Washington, DC: Center on Education Policy Publications.
- Center on Education Policy. (2006). *From the capital to the classroom: Year four of the No Child Left Behind Act*. Washington, DC: Center on Education Policy Publications.
- Centolanza, L. R. (2004). New Jersey teachers believe testing compromises sound practices. *ERS Spectrum*, 18(4), 32-41.
- Cerrillo, T. L., Hansen, M. A., Parke, C. S., Lane, S., & Scott, K. (2000). *The relationship between MSPAP and science classroom instruction and assessment materials*. Washington, DC: Department of Education.
- Chanin, R. H., Collins, J. A., & Pollard, D. (2005). *Pontiac v. Spellings*. Washington, DC: United States Court of Appeals for the Sixth District.
- Clandinin, D. J. (1986). *Classroom practice: Teacher images in action*. Philadelphia: Farmer Press.
- Cohen, D. K., & Ball, D. L. (1990). Relations between policy and practice: A commentary. *Educational Evaluation and Policy Analysis*, 12(3), 249-256.
- Collins, A. (1993). Performance-based assessment of biology teachers: Promise and pitfalls. *Journal of Research in Science Teaching*, 30(9), 1103-1120.

- Collins, A. (1995). National science education standards in the United States: A process and a product. *Studies in Science Education*, 26(1), 7-37.
- Collins, C., & Yeskel, F., (2000). *Economic apartheid in America*. New York: The New Press.
- Comstock, D. E. (1982). A method for critical research. In E. Bredo & W. Feinberg (Eds.), *Knowledge and values in social & educational research* (pp. 370-389). Philadelphia: Temple University Press.
- Craig, C. J. (2004). The dragon in school backyards: The influence of mandated testing on school contexts and educators' narrative knowing. *Teachers College Record*, 106(6), 1229-1257.
- Creswell, J.W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, CA: SAGE.
- Cross, J. R., & Cross, T. L. (2005). Social dominance, moral politics, and gifted education. *Roeper Review*, 28(1), 21-29.
- Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. Thousand Oaks, CA: SAGE.
- Darling-Hammond, L. (2004). From "separate but equal" to *No Child Left Behind*: The collision of new standards and old inequalities. In D. Meier & G. Wood (Eds.), *Many children left behind: How the No Child Left Behind Act is damaging our children and our schools* (pp. 3-32). Boston: Beacon Press.
- DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, 37(6), 582-601.
- De Lissovoy, N., & McLaren, P. (2003). Educational 'accountability' and the violence of capital: A Marxian reading. *Journal of Education Policy*, 18(2), 131-143.
- Dewey, J. (1916). *Democracy and education*. Oxford: McMillan.
- Dillon, S. (2005). Teachers' union and districts sue over Bush law. *New York Times*, Retrieved November 12, 2006 from <http://select.nytimes.com/gst/abstract.html?res=F20B17FD3C550C728EDDA00894DD404482>.
- Doll, M. A. (2000). *Like letters in running water: A mythopoetics of curriculum*. Mahwah, NJ: Lawrence Erlbaum.

- Donmeyer, R. (1990). Generalizability and the general case study. In E. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: A continuing debate* (pp. 175-200). New York: Teachers College Press.
- Doran, R. L., Boorman, J., Chan, F., & Hejaily, N. (1993). Alternative assessment of high school laboratory skills. *Journal of Research in Science Teaching*, 30(9), 1121-1132.
- Downcounty Consortium. (2007). *High school program overview and pathways*. Silver Spring, MD: Division of Consortia Choice and Application Program Services.
- Dufour, R., & Eaker, R. (1992). *Creating the new American school*. Bloomington, IN: National Education Service.
- Duit, R., & Treagust, D. F. (1998). Learning in science – from behaviourism towards social constructivism and beyond. In B. Frasier & K. Tobin (Eds.), *International handbook of science education: Part 1* (pp. 3-25). Dordrecht, Holland: Kluwer.
- Eisner, E. W. (1995). Standards for American schools: Help or hindrance? *Phi Delta Kappan*, 76(10), 758-760, 762-764.
- Ellsworth, E. (1997). *Teaching positions*. New York: Teachers College Press.
- Elmore, R. C., & Sykes, G. (1992). Curriculum policy. In P. Jackson (Ed.), *Handbook of research on curriculum* (pp.185-215). New York: MacMillan.
- Erikson, E. (1986). Qualitative methods on research in teaching. In M. Wittrock (Ed.), *Handbook of research on teaching*, 3rd ed. (pp. 119-161). New York: MacMillan.
- Fensham, P. J. (1992). Science and technology. In P. W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 789-829). New York: Macmillan.
- Ferneding, K. (2004). The discourse of inevitability and the forging of an emergent social vision: Technological diffusion and the dialectic of educational reform discourse. In W. M. Reynolds & J. A. Webber (Eds.), *Expanding curriculum theory: Dis/positions and lines of flight* (pp. 1-18). Mahwah, NJ: Lawrence Erlbaum.
- Finley, S. J. (2000). *Instructional coherence: The changing role of the teacher*. Washington, DC: Office of Educational Research and Improvement.
- Finn, C. (1993). *We must take charge: Our schools, our future*. New York: Free Press.

- Firestone, W., Mayrowetz, D., & Fairman, J. (1998). Performance based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(1), 95-113.
- Firestone, W., Monfils, L., Camilli, G., Schorr, R., Hicks, J., & Mayrowetz, D. (2002). The ambiguity of test preparation: A multi-method analysis in one state. *Teacher College Record*, 104(7), 1485-1523.
- Fischman, W., DiBara, J. A., & Gardner, H. (2006). Creating good education against the odds. *Cambridge Journal of Education*, 36(3), 383-398.
- Flynn, M. (2002). *Title I Supplemental education services and after-school programs: Opportunities and challenges*. Washington, DC: The Finance Project.
- Flyvbjerg, B. (2001). The power of example. In Author (Ed.), *Making social science matter* (pp. 66-87, 178-183). Cambridge: University Press.
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(12), 219-245.
- Friesen, S., Clifford, P., & Jardine, D. W. (2002). Meditations on classroom community and the intergenerational character of mathematical truth. In D.W. Jardine, P. Clifford, & S. Friesen (Eds.), *Back to the basics of teaching and learning: Thinking the world together* (pp.115-131). Mahwah, NJ: Lawrence Erlbaum.
- Fuhrman, S.H., & Massell, D. (1992). *Issues and strategies in systemic reform*. New Brunswick, NJ: Consortium for Policy Research in Education.
- Fullan, M. G. (1991). *The new meaning of educational change*. New York: Teachers College Press.
- Furtak, E. M., & Ruiz-Primo, M.A. (2005). Questioning cycle: Making students' thinking explicit during scientific inquiry. *Science Scope*, 28(4), 22-25.
- Gabel, C. (2003). Linking science inquiry skills in a holistic approach. *Proceedings of 2003 Annual Meeting of the National Association for Research in Science Teaching*, Philadelphia.
- Genn, R. (Ed.). (2007). 'Competition' Art Quotations. Retrieved on July 7, 2007 from <http://quote.robertgenn.com/getquotes.php?catid=48&numcats=337>.
- Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. London: The Falmer Press.

- Giroux, H. A. (1988). *Teachers as intellectuals: Toward a critical pedagogy of learning*. Granby, MA: Bergin & Garvey.
- Giroux, H. A. (2002). Postmodernism and the discourse of educational criticism. In J. R. Grees (Ed.), *Curriculum: Frameworks, criticism, and theory* (pp. 351-378). Richmond, CA: McCutchan.
- Glasnapp, D. R., & Poggio, J.P. (1991). Mandated competency testing programs: Evaluating state and local initiatives. *Theory into Practice*, 30(1), 61-68.
- Goldberg, G. L., & Roswell, B. S. (2000). From perception to practice: The impact of teachers' scoring experience on performance-based instruction and classroom assessment. *Educational Assessment*, 6(4), 257-290.
- Gowen, A. (2002, February 12). Four Montgomery unions urge delay of next MSPAP. *The Washington Post*, p. B05.
- Grant, S.G. (2001). An uncertain lever: The influence of state-level testing in New York on teaching social studies. *Teachers College Record*, 60(3), 398-426.
- Greene, M. (1995). *Releasing the imagination: Essay on education, the arts, and social change*. San Francisco: Jossey-Bass.
- Grundy, S. (1987). *Curriculum: Product or praxis*. New York: The Falmer Press.
- Guba, E., & Lincoln, Y. (1982) Epistemological and methodological bases of naturalistic inquiry. *Educational Communications and technology Journal*, 30(4), 232-252.
- Guba, E.G., & Lincoln, Y.S. (1989). *Fourth generation evaluation*. Thousand Oaks, CA: SAGE.
- Haefner, L. A., & Zembal-Saul, C. (2004). Learning by doing? Prospective elementary teachers' developing understandings of scientific inquiry and science teaching and learning. *International Journal of Science Education*, 26(13), 1653-1674.
- Haigh, M., Frances, B., & Forret, M. (2005). Is "doing science" in New Zealand classrooms an expression of scientific inquiry? *International Journal of Science Education*, 27(2), 215-226.
- Hansen, D. T. (2001). *Exploring the moral heart of teaching: Towards a teacher's creed*. New York: Teachers College Press.
- Hamilton, L. (2003). Assessment as a policy tool. In R.E. Flooden (Ed.), *Review of Research in Education*, 27, 25-68.

- Hammer, D. (1997). Discovery learning and teaching. *Cognition and Instruction*, 15(4), 485-529.
- Hargreaves, A., & Fullan, M. (1998). *What's worth fighting for out there?* New York: Teachers College Press.
- Hargrove, T.Y., Jones, M. G., Jones, B. D., Hardin, B., Chapman, L., & Davis, M. (2000). Unintended consequences of high-stakes testing. *ERS Spectrum*, 18(4), 21-25.
- Harnden, J. (2005). Scientific inquiry scoring guides. *Science Scope*, 28(4), 52-54.
- Harper, D. (2001). *Online etymology dictionary*. Retrieved August 15, 2008 from <http://www.etymonline.com>.
- Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore: Paul H. Brookes.
- Harwood, W.S., Reiff, R.R., & Phillipson, T. (2002). Scientists' conceptions of scientific inquiry: Voices from the front. In P. A. Rubba, J. A. Rye, W. J. Di Biase, & B. A. (Eds.), *Proceedings of 2002 Annual International Conference of the Association for the Education of Teachers in Science*, Charlotte, NC.
- Haughey, M. (1993). *Toward teacher growth: A study of the impact of Alberta's teacher evaluation policy*. Edmonton: Alberta Department of Education.
- Haury, D. L. (1993). Teaching science through inquiry. ERIC/CSMEE Digest, ED59048.
- Haury, D.L. (2001). *The state of state proficiency testing in science*. Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.
- Herman, J. L. (1997). Large-scale assessment in pursuit of school reform: Lessons in the search for alternative measures. *CSE Technical Report 446*. Los Angeles: CREST/UCLA.
- Herman, J. L., Baker, E. L., & Linn, R. L. (2004). Accountability systems in pursuit of student learning: Moving to the next generation. *Newsletter of the National Center for research on Evaluation, Standards, and Student Testing*. CRESST Line.
- Herman, J. L., Webb, N. M., & Zuniga, S. (2003). Alignment and college admissions: The match of expectations, assessments, and educator perspectives. *CSE Technical Report 593*. Los Angeles: CREST/UCLA.

- Herriott, R.E., & Firestone, W.A (1983). Multi-site qualitative policy research: Optimizing description and generalizeability. *Educational Researcher*, 12(2), 14-19.
- Henderson, A. T., & Berla, N. (Eds.). (1996). *A new generation of evidence: A family is critical to student achievement*. Washington, DC: Center for Law and Education.
- Hickey, D. T, DeCuir, J., Hand, B., Kyser, B., Laprocina, S., & Mordica, J. (2002). *Technology-supported formative and summative assessment of collaborative scientific inquiry*. Arlington, VA: National Science Foundation.
- Hoffman, J. V., Assaf, L. C., & Paris, S. G. (2001). High-stakes testing in reading: Today in Texas, tomorrow? In S.J. Barrentine & S.M. Stokes (Eds.), *Reading assessment: Principles and practices for elementary teachers* (pp.108-120). Newark, DE: International Reading Association.
- hooks, b. (1994). *Teaching to transgress: Education as the practice of freedom*. New York: Routledge.
- Hirshy, A., & Wilson, M. (2002). The sociology of the classroom and its influence on student learning. *Peabody Journal of Education*, 77(3), 85-100.
- Horn, R.A., Jr., & Kincheloe, J. L. (2001). *American standards: Quality education in a complex world, the Texas case*. Oxford, England: Peter Lang.
- Hursh, D. (2005). The growth of high-stakes testing in the USA: Accountability, markets and the decline in educational equality. *British Educational Research Journal*, 31(5), 605-622.
- Ingersoll, R. M. (2003). *Is there really a teacher shortage?* Seattle, WA: Center for the Study of Teaching and Policy.
- Intrator, S. M., & Kunzam, R. (2006). Starting with the soul. *Educational Leadership* 63(6), 38-42.
- Invernizzi, M. A., Landrum, T. J., Howell, J. L., & Warley, H. P. (2005). Toward the peaceful coexistence of test developers, policymakers, and teachers in an era of accountability. *Reading Teacher*, 58 (7), 610-618.
- Irons, E. J., & Harris, S. (2007). *The challenges of No Child Left Behind: Understanding the issues of excellence, accountability, and choice*. Lanham, MD: Rowman & Littlefield Education.

- Jacob, B. A. (2002). *Accountability, incentives, and behavior: The impact of high stakes testing the Chicago Public Schools*. Retrieved October 8, 2006, from <http://www.nber.org/papers/w8968>.
- Jardine, D. W., Clifford, P., & Friesen, S. (2003). *Back to the basics of teaching and learning: Thinking the world together*. Mahwah, NJ: Lawrence Erlbaum.
- Jardine, D. W. & Rinehart, P. (2003). Relentless writing and the death of memory in the elementary classroom. In D. W. Jardine, P. Clifford, & S. Friesen (Eds.), *Back to the basics of teaching and learning: Thinking the world together* (pp. 73-88). Mahwah, NJ: Lawrence Erlbaum.
- Johnson, D.W., & Johnson, R.T. (1991). *Learning mathematics and cooperative learning: Lesson plans for teachers*. Minneapolis, MN: Interaction Book Company.
- Jones, B. D., & Egley, R. J. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing, *Education Policy Analysis Archives*, 12(39), Retrieved November 5, 2006 from <http://epaa.asu.edu/epaa/v12n39/>.
- Kantor, H. & Lowe, L. (2006). From new deal to no deal: *No Child Left Behind* and the devolution of responsibility for equal opportunity. *Harvard Educational Review*, 76(4), 474-502.
- Karplus, R., & Thier, H. D. (1967). *A new look at elementary school science*. Chicago: Rand McNally.
- Ketterlin-Geller, L. R., McCoy, J., Twyman, T., & Tindel, G. (2003). How do critical thinking measures fit within standards-based reform?. *Assessments for Effective Intervention*, 28(3-4), 37-48.
- Kingdon, J. W. (2003). *Agendas, alternatives, and public policies*. New York: Longman.
- Kinney, J. H. (1989). *Study of the effects of a cooperative learning program on the achievement of ninth grade multi-cultural general biology classes*. Paper presented to the Alexandria City, VA School Board, Alexandria, VA.
- Kinter, S & Schwadron, S. (2007). MSDE to eliminate HSA writing sections. *Silver Chips Online*, Retrieved November 8, 2007 from <http://silverchips.mbhs.edu/inside.php?sid=7744>.
- Kleinhenz, E., & Ingvarson, L. (2004). Teacher accountability in Australia: Current policies and practices and their relation to the improvement of teaching and learning. *Research Papers in Education*, 19(1), 31-49.
- Klopfer, L. E. (1969). Science education in 1991. *The School Review*, 77(3), 199-217.

- Koretz, D. (2005). *Alignment, high-stakes, and the inflation of test scores*. CSE Report 655.
- Koretz, D., Mitchell, K., Barron, S., & Keith, S. (1996). *Final report: Perceived effects of the Maryland School Performance Assessment Program*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Kozol, J., (1992). *Savage inequalities: Children in America's schools*. New York: Crown Publishers.
- Krug, E. A. (1964). *The shaping of the American high-school, 1880-1920*. Madison, WI: University of Wisconsin Press.
- Kulm, G., Wilson, D.L., & Kitchen, R. (2005). Alignment of content and effectiveness of assessment items. *Educational Assessment*, 10(4), 333-356.
- Lee, V. E., & Berkham, D. T., (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.
- Lageman, E. C. (2000). *An elusive science: The troubling history of education research*. Chicago: University of Chicago Press.
- Lederman, N. G. (1998). The state of science education: Subject matter without context. *Electronic Journal of Science Education*, 3(2), 1-11.
- Lederman, N. G., & Niess, M. L. (2000). Problem solving and solving problems: Inquiry about inquiry. *School Science and Mathematics*, 100(3), 113-116.
- Leffler, J. (2004). *Alignment of Idaho state standards with the Idaho Direct Mathematics Assessment in Grades 4, 6, and 8*. Boise: Idaho State Board of Education.
- Leonard, W. H., & Penick, J. E. (2005). Assessment of standards-based biology teaching. *American Biology Teacher*, 67(2), 73-76.
- Light, R. (1990). *The Harvard assessment seminars*. Cambridge, MA: Harvard University Press.
- Lin, W. (1998, April). *The effects of restructuring biology teaching by a constructivist teaching approach: An action research*. Paper presented at the 71st Annual Meeting of the National Association for Research in Science Teaching, San Diego, CA.
- Lincoln, Y. S., & Guba, E.G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: SAGE.

- Linn, R. (2005a). *Issues in the design of accountability systems*. CSE Technical Report 650.
- Linn, R. (2005b). *Test based educational accountability in the era of No Child Left Behind*. CSE Technical Report 651.
- Lipsky, M. (1980). *Street-level bureaucracy: Dilemmas of the individual in public service*. New York: Russell Sage Foundation.
- Lord, T. (2001). 101 reasons for using cooperative learning in biology teaching. *American Biology Teacher*, 63(1), 30-38.
- Loucks-Horsley, S., & Bybee, R. W. (1998). Implementing the national science education standards. *Science Teacher*, 65(6), 22-26.
- Louis, K. S., Febey, K., & Shroeder, R. (2005). State-mandated accountability in high schools: Teachers' interpretations of a new era. *Educational Evaluation and Policy Analysis*, 27(2), 177-204.
- Lucas, D., Broderick, N., Lehrer, R., & Bohanen, R. (2005). Making the grounds of scientific inquiry visible in the classroom. *Science Scope*, 29(3), 39-42.
- Malen, B., & Muncey, D. (2000). Creating "a new set of givens"? The impact of state activism on school autonomy. In N. D. Theobald & B. Malen (Eds.), *Balancing local control and state responsibility for K-12 education* (pp. 199-244). Washington, DC: American Education Finance Association.
- Marx, R. W., & Harris, C. J. (2006). No Child Left Behind and science education: Opportunities, challenges, and risks. *The Elementary School Journal*, 106(5), 455-466.
- Marshall, C., & Rossman, G. (1989). *Designing qualitative research*. Thousand Oaks, CA: SAGE.
- Maryland State Department of Education. (1999). *Science core learning goals*. Retrieved January 23, 2007 from <http://www.montgomeryschoolsmd.org/curriculum/science/forms/scicls99.pdf>.
- Maryland State Department of Education. (2000a). *Maryland School Performance Assessment Program (MSPAP)*. Retrieved July 27, 2006 from http://www.mdk12.org/mspp/mspap/how-scored/mspap_info/index.html.
- Maryland State Department of Education. (2000b). *Maryland science content standards*. Retrieved January 27, 2007 from http://www.unavco.org/edu_outreach/docs/maryland.pdf.

- Maryland State Department of Education. (2001). *The Maryland model for school readiness: Science modules*. Retrieved August 4, 2006 from http://www.mdk12.org/instruction/ensure/mmsr/mmsr992_science.html.
- Maryland State Department of Education. (2002a). *What is MSPAP?*. Retrieved July 27, 2006 from <http://www.mdk12.org/mspp/mspap/what-is-mspap/index.html>.
- Maryland State Department of Education. (2002b). *High school core learning goals: Biology*. Retrieved February 4, 2007 from http://hsaexam.org/img/CD%20Content/clg/CLG_Biology.pdf.
- Maryland State Department of Education. (2002c). *Maryland High School Assessment: Biology public release 2002*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2002d). *Answer key Biology High School Assessment*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2003a). *Maryland High School Assessment: Biology public release 2003*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2003b). *Answer key Biology High School Assessment*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2004a). *What is the MSA?*. Retrieved July 27, 2006 from http://www.mdk12.org/mspp/k_8/whatis.html.
- Maryland State Department of Education. (2004b). *Voluntary curriculum*. Downloaded on July 27, 2006 from <http://www.mdk12.org/instruction/curriculum/index.html>.
- Maryland State Department of Education. (2004c). *Maryland High School Assessment: Biology public release 2004*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2004d). *Answer key Biology High School Assessment*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2005a). *What is the High School Assessment program?* Retrieved July 27, 2006 from http://www.mdk12.org/mspp/high_school/what_is/index.html.
- Maryland State Department of Education. (2005b). *Maryland High School Assessment: Biology public release 2005*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2005c). *Answer key Biology High School Assessment*. Baltimore, MD: MSDE.

- Maryland State Department of Education. (2006a). *The fact book 2004-2005: A statistical handbook*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2006b). *Maryland High School Assessment: Biology public release 2006*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2006c). *Answer key Biology High School Assessment*. Baltimore, MD: MSDE.
- Maryland State Department of Education. (2007a). Class of 2009 on track to pass the High School Assessments: State superintendent recommends new plan for some students still struggling. Baltimore, MD: MSDE
- Maryland State Department of Education. (2007b). *Budget*. Retrieved September 8, 2007 from <http://www.msa.md.gov/msa/mdmanual/13sdoe/html/sdoeb.html>.
- Maslow, A. (1943). A theory of human motivation. *Psychological Review*, 50, 370-396.
- McDonnell, L. M. (1994). Assessment policy as persuasion and regulation. *American Journal of Education*, 102(4), 394-420.
- McDonnell, L. M. (2004). *Politics, persuasion, and educational testing*. Cambridge, MA: Harvard University Press.
- McDermott, L. C. (1990). A perspective on teaching preparation in physics and other sciences: The need for special science courses for teachers. *American Journal of Physics*, 58(7), 734-742.
- McFadden, C. (2001). Towards a new framework for science curriculum development. *Canadian Journal of Science, Mathematics and Technology Education*, 1(4), 459-462.
- McGinnis, R., & Stefanich, G. (2007). Special need and talents in science education. In N. Lederman & S. Albel (Eds.), *Handbook of research in Science Education* (pp. 319-344). Mahwah, NJ: Lawrence Erlbaum.
- McLaren, P. (1989). *Life in schools: An introduction to critical pedagogy in the foundations of education*. New York: Longman.
- McNeil, L. M. (2000). *Contradictions of school reform: The educational costs of standardized testing*. London: Routledge.
- Medley, D. M. (1982). Teacher effectiveness. In H.E. Mitzel (Ed.), *Encyclopedia of educational research* (pp. 1894-1903). New York: The Free Press.

- Mehrens, W. A. (1998). Consequences of assessment: What is the evidence? *Education Policy Analysis Archives*, 6(13). Retrieved November 5, 2006 from <http://epaa.asu.edu/epaa/v6n13.html>.
- Mentor, I., Muschamp, Y., Nicholls, P., Ozga, J., & Pollard, A. (1997). *Work and identity in the primary school*. Philadelphia: Open University Press.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.
- Mid South Regional Resource Center. (2004). *State standards and assessments: Critical elements checklist*. Lexington, KY: Mid South Regional Resource Center.
- Mintrop, H., & Trujillo, T. (2005). *Corrective action in low performing schools: Lessons for NCLB implementation from state and district strategies in first-generation accountability systems*. Los Angeles: Center for the Study of Evaluation.
- Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis* (2nd Edition). Thousand Oaks, CA: SAGE.
- Miller, J. L. (2004). *The sound of silence breaking and other essays: Working the tension in curriculum theory*. New York: Peter Lang.
- Mislevy, R.J. (2004). *The case for an integrated design framework for assessing science inquiry*. Los Angeles: Center for Research on Evaluation Standards and Student Testing.
- Mujis, D. (2006). Measuring teacher effectiveness: Some methodological reflections. *Educational Research and Evaluation*, 12(1), 53-74.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Retrieved July 24, 2006 from <http://www.ed.gov/pubs/NatAtRisk/index.html>.
- National Council on Education Standards and Testing. (1991). *Raising standards for American education*. Washington, DC: National Council on Education Standards and Testing.
- National Assessment of Educational Progress. (2005). *The nation's report card: Reading 2005*. Washington, DC: U.S. Department of Education Institute of Education Sciences.
- National Educational Association. (1893). *Report of the Committee on Secondary Education*. Washington, DC: Government Printing Office.

- National Educational Association. (1920). *Reorganization of science in secondary schools: A report of the Commission on the Reorganization of Secondary Education*. Washington, DC: U.S. Bureau of Education.
- National Educational Association. (2007a). *Stand up for children: Pontiac v. Spellings*. Retrieved on September 10, 2007 from <http://www.nea.org/lawsuit/index.html>.
- National Educational Association. (2007b). No Child Left Behind/ESEA: It's time for a change! Retrieved on September 10, 2007 from <http://www.nea.org/lac/esea/index.html>.
- National Institute of Child Health and Human Development–Early Child Care Research Network. (2005). A day in third grade: A large-scale study of classroom quality and teacher and student behavior. *Elementary School Journal*, 105, 305–323.
- National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
- National Science Teachers Association. (2004). *Scientific inquiry*. Arlington, VA: NSTA.
- Nesbit, C. R., & Rogers, C. A., (1997). Using cooperative learning to improve reading and writing in science. *Reading and Writing Quarterly*, 13(1), 53-70.
- Nieto, S. (2006). Solidarity, courage and heart: What teacher educators can learn from a new generation of teachers. *Intercultural Education*, 17(5), 457-473.
- Noddings, N. (2007). *When school reform goes wrong*. New York: Teachers College Press.
- Oakes, J. (1987). *Improving inner-city schools: Current directions in urban district reform*. Santa Monica, CA: RAND Corporation.
- O'Hear, P. (1994). An alternative national curriculum. In S. Tomlinson (Ed.), *Educational reform and its consequences* (pp. 55-72). London: IPPR/Rivers Oram Press.
- Olsen, B., & Kirtman, L. (2002). Teacher as mediator of school reform: An examination of teacher practice in 36 California restructuring schools. *Teachers College Record*, 104(2), 301-324.
- Olson, L. (2006). Progress report. *Teacher Magazine*, 17(5), 6.
- Osborne, J., Collins, S., Ratcliffe, M., Robin, M., & Duschl, R. (2003). What “ideas about-science” should be taught in school science?: A Delphi study of the expert community. *Journal of Research in Science Teaching*, 40(7), 692-720.

- Oser, F. K. (1992). Morality in professional action: A discourse approach to teaching. In F. K. Oser, A. Dick, & J. L. Patry (Eds.), *Effective and responsible teaching: A new synthesis* (pp. 109-125). San Francisco, CA: Jossey-Bass.
- Parke, C. S., Cerrillo, T. S., Levenson, J., O'Mara, J., Hansen, M. A., & Lane, S. (1999). *Impact of the Maryland School Performance Assessment Program (MSPAP): Evidence from classroom instruction and assessment activities (Reading, Writing)*. Washington, DC: Department of Education.
- Patton, M.Q. (1990). Designing qualitative studies: Critical trade-offs. In Author (Ed.), *Qualitative evaluation and research methods* (pp. 162-186). Newbury Park, CA: SAGE.
- Peterson, K. (2005). *No letup in unrest over Bush school law*. Washington, DC: Pew Research Center. Retrieved November 12, 2006 from <http://www.stateline.org/live/ViewPage.action?siteNodeId=136&languageId=1&contentId=41610>.
- Pinar, W. F. (2002). *Currere: Toward reconceptualization*. In J. R. Grees (Ed.), *Curriculum: Frameworks, criticism, and theory* (pp. 471-488). Richmond, CA: McCutchan Publishing Corporation.
- Pinar, W. F. (2004). *What is curriculum theory?* Mahwah, NJ: Lawrence Erlbaum.
- Pinar, W. F. (2005). "A lingering note": An introduction to the collected works of Ted T. Aoki. In W. F. Pinar & R. L. Irwin (Eds.), *Curriculum in a new key: The collected works of Ted T. Aoki* (pp.1-85). Mahwah, NJ: Lawrence Erlbaum Associates.
- Pinar, W. F., Reynolds, W. M., Slattery, P., & Taubman, P. M. (2000). *Understanding curriculum: An introduction to the study of historical and contemporary curriculum discourses*. New York: Peter Lang.
- Porter, A. C. (2002). Measuring the content of instruction: Uses in research and instruction. *Educational Researcher*, 31(7), 3-14.
- Posner, G. J. (2004). *Analyzing the curriculum* (3rd Edition). Boston: McGraw Hill.
- Pring, R. (2001). Education as a moral practice. *Journal of Moral Education*, 30(2), 101-112.
- Pringle, R. M., & Carrier-Martin, S. (2005). The potential impacts of upcoming high-stakes testing on the teaching of science in elementary classrooms. *Research in Science Education*, 35 (2-3), 347-361.

- Ranson, S. (1995). Theorizing education policy. *Journal of Education Policy*, 10(4), 427-448.
- Ravitch, D. (1995). *National standards in American education*. Washington, DC: Brookings Institution.
- Ravitch, D. (2007). Get Congress out of the classroom. *The New York Times*, 157(54,086), A25.
- Reay, D. (1998). Micro-politics in the 1990's: Staff relationships in secondary schooling. *Journal of Education Policy* 13(2), 179-196.
- Reiff, R., Harwood, W., & Phillipson, T. (2002). A scientific method based upon research scientists' conceptions of scientific inquiry. In P. A. Rubba, J. A. Rye, W. J. Di Biase, & B. A. Crawford (Eds.), *Proceedings of 2002 Annual International Conference of the Association for the Education of Teachers in Science*, Charlotte, NC.
- Resnick, L. B. (2006). Making accountability really count. *Educational Measurement: Issues and Practice*, 25(1), 3-37.
- Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. G. Gifford & M. C. O'Conner (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 37-75). Boston: Kluwer Academic.
- Resnick, L. B., Rothman, R., Slattery, J. B., & Vraneck, J. L. (2004). Benchmarking and alignment of standards and testing. *Educational assessment*, 9(1-2), 1-27.
- Rex, L.A., & Nelson, M.C. (2004). How teachers' professional identities position high-stakes test preparation in their classrooms. *Teachers College Record*, 106(6), 1288-1331.
- Reynolds, W. M., & Webber, J. A. (2004). Introduction: Curriculum dis/positions. In W. M. Reynolds & J. A. Webber (Eds.), *Expanding curriculum theory: Dis/positions and lines of flight* (pp. 1-18). Mahwah, NJ: Lawrence Erlbaum.
- Ripley, R. (1985). *Policy analysis in political science*. Chicago: Nelson Hall.
- Roehrig, G. H., & Luft, J. A., (2004). Inquiry teaching in high school chemistry classrooms: The role of knowledge and beliefs. *Journal of Chemical Education*, 81(10), 1510-1516.
- Rosenbaum, J. (2001). *Beyond college for all: Career paths for the forgotten half*. New York: Russell Sage Foundation.

- Rothstein, R. (2004). *Class and schools: Using social, economic, and educational reform to close the Black-White achievement gap*. Washington, DC: Economic Policy Institute.
- Rudelvidge, A. (2003). The politics of No Child Left Behind. *Education Next*, 3(4), 63-69.
- Rudman, H. C. (1989). *Integrating testing with teaching*. Washington, DC: Office of Educational Research and Improvement.
- Rudolph, J. L. (2005). Epistemology for the masses: The origin of “the scientific method” in American schools. *History of Education Quarterly*, 45(3), 341-376.
- Ruppert, S. (1994). *Charting higher education accountability*. Boulder, CO: Education Commission of the States.
- Ruiz-Primo, M. A., & Furtak, E. M. (2004). *Informal formative assessment of students’ understanding of scientific inquiry*. CSE Report 639. Los Angeles: Center for the Study of Evaluation.
- Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569-600.
- Ryan, P., & Walking-Woman, L. (2000). *Linking writing to the process of scientific inquiry: Strategies from writing teachers in the disciplines*. Minneapolis: Conference on College Composition and Communication.
- Sarason, S. B. (1990). *The predictable failure of educational reform: Can we change course before it's too late?* San Francisco: Jossey-Bass.
- Saul, J. R. (1992). *Voltaire’s bastards: The dictatorship of reason in the west*. New York: Free Press.
- Scanell, M. M., & Metcalf, P.L. (2000). Autonomous board and standards-based teacher development. *Educational Researcher*, 14(1), 61-76.
- Schmidt, P., Gillen, S., Zollo, T., & Stone, R. (2002). Literacy learning and scientific inquiry: Children respond. *The Reading Teacher*, 55(6), 534-548.
- Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1997). *Splintered vision: An investigation of U.S. science and mathematics education*. East Lansing, MI: Michigan State University.

- Scott, T. (2005). Consensus through accountability? The benefits and drawbacks of building community with accountability. *Journal of Adolescent & Adult Literacy*, 48(1), 48-59.
- Segal, S. (1998). The role of contingency and tension in the relationship between theory and practice in the classroom. *Journal of Curriculum Studies*, 30(2), 199-206.
- Selden, S. (1991, October). Conflicting paradigms in the research on college teaching: A source of creative tension in university life. *The Maryland Association for Higher Education Journal*, 14, 27-31.
- Seidman, R. H. (1996). National education "Goals 2000": Some disastrous unintended consequences. *Education Policy Analysis Archives*, 4(11), 1-35.
- Sheppard, K., & Robbins, D. M. (2006). Chemistry, the terminal science? *Journal of Chemical Education*, 83(11), 1617-1620.
- Shulman, L.S. (1988). Disciplines of inquiry in education: An overview. In R. M. Jaeger (Ed.), *Complementary methods for research in education* (pp. 3-17). Washington, DC: American Education Research Association.
- Shulte, B. (2002, February 13). Montgomery board wants Maryland to drop MSPAP testing. *The Washington Post*, p. B04.
- Sikes, P. J. (1992). Imposed changes and the experienced teacher. In M. Fullen & A. Hargreaves (Eds.), *Teacher development and educational change* (pp. 36-55). London: Falmer.
- Sinatra, G. M., & Pintrich, P. R. (Eds.). (2003). *Intentional conceptual change*. Mahwah, NJ: Lawrence Erlbaum.
- Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: SAGE.
- Stake, R. (1998). Some comments on assessment in U.S. education. *Education Policy Analysis Archives*, 6(14), Retrieved November 4, 2006 from <http://epaa.asu.edu/epaa/v6n14.html>.
- Stecher, B.M. (2002). Consequences of large scale, high-stakes testing on school and classroom practices. In L.S. Hamilton, B.M. Stecher, & S.P. Klein (Eds.), *Making sense of test-based accountability* (pp. 79-100). Santa Monica, CA: RAND.
- Stecher, B.M., & Hamilton, L.S. (2002). Putting theory to the test: System of accountability should be held accountable. *Rand Review*, 26(1), 16-23.

- Sternberg, B. (2006). Real improvement for real students: Test smarter, serve better. *Harvard Educational Review*, 76(4), 557-563.
- Stevenson, J. W. (2005). *Final 2004 report on Adequate Yearly Progress in Montgomery County Public Schools*. Rockville, MD: Department of Shared Accountability for Montgomery County Public Schools.
- Stone, C. A., & Lane, S. (2003). Consequences of a state accountability program: Examining relationships between school performance gains and teacher, student, and school variables. *Applied Measurement in Education*, 16(1), 1-26.
- Strike, K. A. (1997). Centralized goal formation and systemic reform: Reflections on liberty, localism and pluralism. *Education Policy Analysis Archives*, 5(11), Retrieved November 5, 2006 from <http://epaa.asu.edu/epaa/v5n11.html>.
- Sutton, C. (1992). *Words, science and learning*. Philadelphia: Open University Press.
- Tamir, P. (1998). Assessment and evaluation in science education: Opportunities to learn and outcomes. In K. Tobin (Ed.), *International handbook of science education* (pp. 761-789). London: Kluwer Academic.
- Tannen, D. (1995). The power of talk: Who gets heard and why. *Harvard Business Review*, 73(5), 138-148.
- Tobin, K., Tippins, D. J., & Gallard, A. J. (1994). Research on instructional strategies for teaching science. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 45-93). New York: McMillan.
- Townsend, T., Lynch, A., Gagnon, S., Kraemer, D., Szesze, M., & Morse, G. (2003). *Draft Biology framework*. Rockville, MD: MCPS Science Office.
- Treagust, D. F., Jacobowitz, R., Gallagher, J. L., & Parker, J. (2001). Using assessment as a guide in teaching for understanding: A case study of a middle school science class learning about sound. *Science Education*, 85(2), 137-157.
- Trump, J. L., & Miller, D. F. (1973). *Secondary school curriculum and improvement: Challenges, humanism, and accountability*. Boston: Allyn & Bacon.
- Tyler, R. (1949). *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press.
- United States Congress. (1994a). *Goals 2000: Educate America Act Public Law 103-227*. Washington, DC: The Government Printing Office.
- United States Congress. (1994b). *Improving America's Schools Act of 1994 Public Law 103-382*. Washington, DC: The Government Printing Office.

- United States Congress. (2002a). *No Child Left Behind Act of 2001 Public law 107-110*. Washington, DC: The Government Printing Office.
- United States Congress. (2002b). Executive Summary of *No Child Left Behind Act of 2001*. Retrieved July 2, 2007 from <http://www.ed.gov/nclb/overview/intro/execsumm.pdf>.
- United States Department of Education. (2004a). *Overview: Four pillars of NCLB*. Retrieved June 26, 2007 from <http://www.ed.gov/nclb/overview/intro/4pillars.html>.
- United States Department of Education. (2004b). *Glossary of Terms*. Retrieved July 2, 2007 from <http://www.ed.gov/nclb/index/az/glossary.html?src=az>.
- United States Department of Education. (2007). Secretary Spellings highlights *No Child Left Behind* reauthorization in remarks to members of the business coalition for student achievement. Retrieved on September 12, 2007 from <http://www.ed.gov/news/pressreleases/2007/09/09052007.html>.
- United States Department of Labor. (1992). *Learning a living: A blueprint for high performance*. Washington, DC: The Government Printing Office.
- United States National Research Council for TIMMS. (1997). *Splintered vision: An investigation of U.S. science and mathematics education*. London: OECD.
- 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.
- Valli, L., & Chambliss, M. (2007). Creating classroom cultures: One teacher, two lessons, and a high-stakes test. *Anthropology and Education Quarterly*, 38(1), 57-75.
- Viadero, D. (2006) Talk of US crisis in math and science is largely misplaced, skeptics say. *Education Week*, 25(28), 21.
- Vu, P. (2007). No Child law faces medley of changes. *Stateline.org*. Retrieved on September 12, 2007 from <http://www.stateline.org/live/details/story?contentId=193830>.
- Waldrop, B. (2003, September 24). Corvallis schools resist No Child Left Behind Act. *Corvallis Gazette-Times*. Retrieved November, 12, 2006 from <http://gazettetimes.com/articles/2003/09/25/news/community/thu02.txt>.
- Waterman, M.A. (1998). Investigative case study approach to biology learning. *Bioscene*, 24(1), 3-10.

- Waters, G. A. (1998). Critical evaluation for education reform. *Education Policy Analysis Archives*, 6(20), downloaded November 4, 2006 from <http://epaa.asu.edu/epaa/v6n20.html>.
- Walberg, H. J. (2003) Accountability helps students at risk. *Education Week*, 22(33), 42-44.
- Watson, S. (1991). Cooperative learning and group educational modules: effects on cognitive achievement of high school biology students. *Journal of Research in Science Teaching*, 28(2), 141-146.
- Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states*. Washington, DC: Council of chief state school officers.
- Weiss, I. R., Banilower, E. R., McMahon, K. C., & Smith, P. S. (2001). *Report of the 2000 National Survey of Science and Mathematics Education*. Chapel Hill, NC: Horizon Research, Inc.
- Wideen, M. F., O'Shea, T., Pye, I., & Ivany, G. (1997). High-stakes testing and the teaching of science. *Canadian Journal of Education*, 22(4), 428-444.
- Widlack, H. (1985). Real-life shock: The inapplicability of theoretical knowledge to classroom practice. *Education*, 30(2), 100-111.
- Wilson, M. R., & Bertenthal, M. W. (Eds.). (2005). *Systems for state science assessment*. Washington, DC: National Academies Press.
- Wood, T. (1988). State mandated accountability as a constraint on teaching and learning science. *Journal of Research in Science Teaching*, 25(8), 631-641.
- Wraga, W. G. (1999). The educational and political implications of curriculum alignment and standards based reform. *Journal of Curriculum and Supervision*, 15(1), 4-25.
- Wright, W. E. (2006). A catch-22 for language learners. *Educational Leadership*, 64(3), 22-27.
- Yin, R. K. (1981). The case study crisis: Some answers. *Administrative Science Quarterly*, 26 (1), 58-65.
- Yin, R. K. (2003). *Case study research: Design and methods*. Thousand Oaks, CA: SAGE.
- Yore, L. D., Bisanz, G. L., & Hand, B. M. (2003). Examining the literacy component of scientific literacy: 25 years of language arts and science teaching. *International Journal of Science Education*, 25(6), 689-725.

- Young, T. R. (1986). The sociology of sport: Structural Marxist and cultural Marxist approaches. *Sociological Approaches*, 29(1), 3-28.
- Zachos, P., Hick, T. L., Doane, W. E. J., & Sargent, C. (2000). Setting theoretical and empirical foundations for assessing scientific inquiry and discovery in educational programs. *Journal of Research in Science Teaching*, 37(9), 938-962.
- Zacanella, D. (1992). The influence of state-mandated testing on teachers of literature. *Educational Evaluation and Policy Analysis*, 14(3), 283-295.
- Zion, M., Michalsky, T., & Maverich, Z. R. (2005). The effects of metacognitive instruction embedded within an asynchronous learning network on scientific inquiry skills. *International Journal of Science Education*, 27(8), 957-983.