Understanding the Challenges to the Implementation of Assessment Reform in Science Classrooms: A Case Study of Science Teachers' Conceptions and Practices of Assessment

Mehmet Aydeniz
THE FLORIDA STATE UNIVERSITY
COLLEGE OF EDUCATION

UNDERSTANDING THE CHALLENGES TO THE IMPLEMENTATION OF
ASSESSMENT REFORM IN SCIENCE CLASSROOMS:
A CASE STUDY OF SCIENCE TEACHERS’ CONCEPTIONS and PRACTICES OF
ASSESSMENT

By

MEHMET AYDENIZ

A Dissertation submitted to the
Department of Middle and Secondary Education
In partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Spring Semester, 2007
The members of the committee approve the dissertation of Mehmet Aydeniz defended on June 6, 2006.

________________________
Nancy T. Davis
Professor Directing Dissertation

________________________
John Dorsey
Outside Committee Member

________________________
Sherry Southerland
Committee Member

________________________
Penny J. Gilmer
Committee Member

Approved:

Pamela Carroll, Chair, Department of Middle and Secondary Education

The Office of Graduate Studies has verified and approved the above named committee members.
This dissertation is dedicated to my family; my mother, my father, my sisters and my only brother. Without your love, compassion, and support I would not have been able to be here.

To my students at Seminole High School who lifted my spirit up every single day and challenged me to grow as a science educator.

To name few:
Sonoma Lious, Nicole Battoon, Haley Nabergale, Courtney Gilmore, Sara Simons, Britney Schniener, Nicole Seminova, Lena Brake, Kelsey Quanz
Jordan Lious, Ryan Weiss, Joe Guerra, Jessica, Katherine Rickards, James, Stephanie Escobar, Jessica, Thomas, Kristen, Alyssaa, Monica, John, Kayla Heflin, Tiffany Kirsten Kaflowic, Nicole, Reena, Ally D,
and
all others who I will always remember.

Thanks to you all.
ACKNOWLEDGEMENTS

It has been an exceptional journey -one that I had never dreamed of. I started this journey by earning a scholarship from the Turkish Government in 1999.

The work with this dissertation has been extensive, laborious and trying, but in the first place exciting, instructive, and fun. Without help, support, and encouragement from several people I would have never been able to finish this work.

I would like to express my gratitude to my advisor, Dr. Nancy Davis, for her continuous support, patience, and encouragement throughout my graduate studies. Dr. Davis has guided me to a deeper understanding of the theoretical and practical issues in science education through her invaluable comments during the writing process of this dissertation.

My thanks also go to the members of my committee, Dr. Sherry Southerland, Dr. Penny Gilmer and Dr. John Dorsey for reading my prospectus and providing many valuable comments that improved the content and presentation of this dissertation. I deeply thank Dr. Sherry Southerland for her critical feedback on my writing.

My special thanks go to Dr. Penny Gilmer for her continuous support and encouragement for me to write and present at various national and international conferences. Without her support and encouragement I would not have been able to be where I am today. I also thank Dr. Pamela Carroll for her support and encouragement.

I am fortunate to have the opportunity to work with a group of energetic, fun and hardworking people in Science Education Program. I appreciate all their friendships and their collective insights and support throughout my graduate studies. These people are: Dr. Randy Spaid, Dr. Felicia Moore, Dr. Scott Sowell, Dr. Martin Balinsky, Dr. Ahmad Qablan, Dr. Margaret Blanchard, Dr. Evrim Genc, and Mr. Yavuz Saka.

I express my thankfulness to all the subjects in my research, who have all dedicated part of their valuable time for my curiosity.

Finally, I express my gratitude to the Turkish Government for financially supporting my graduate studies. I extend my gratitude to officers working in Ankara and New York offices for taking care of the administrative tasks.
TABLE OF CONTENTS

LIST OF TABLES ......................................................................................................................... ix
ABSTRACT ..................................................................................................................................... x
CHAPTER ONE .................................................................................................................................. 1
INTRODUCTION TO THE STUDY ................................................................................................. 1
Statement of Problem and Purpose ............................................................................................... 3
Calls for Reform in Science Education ....................................................................................... 5
Significance of Teachers’ Conceptions of Assessment ................................................................. 8
Significance of Assessment in Science Education Reform ............................................................. 10
CHAPTER TWO ............................................................................................................................. 13
LITERATURE REVIEW .................................................................................................................. 13
National Reports and Standards Movement ............................................................................. 13
   Project 2061 ............................................................................................................................. 13
   National Science Education Standards ..................................................................................... 15
How Do Reform Documents Inform Science Educators? ........................................................... 16
   Purposes of Assessment .......................................................................................................... 17
   The Need for Alternative Forms of Assessment ...................................................................... 18
Science Education Reform and Assessment ............................................................................. 19
   A Historical Perspective on Assessment ............................................................................... 22
   Social Efficiency, Behaviorism and Standardized Testing ...................................................... 23
Pedagogical Shift on Learning and Assessment ....................................................................... 24
   Social Constructivist Pedagogy ............................................................................................... 24
   Social Constructivism and Assessment .................................................................................. 26
   Formative Assessment and Student Learning ........................................................................ 26
National Reports and Standardized Testing ............................................................................ 28
   A Nation at Risk ....................................................................................................................... 28
   The No Child Left Behind Act ............................................................................................... 30
   Critics of the No Child Left Behind Act ............................................................................... 31
Florida Comprehensive Assessment Test (FCAT) .................................................................... 32
   Equity and Access to Education For All .............................................................................. 33
   The Influence of Standardized Testing on Teaching .............................................................. 35
Concluding Remarks ................................................................................................................. 36
   Understanding Teachers’ Beliefs on Assessment .................................................................. 38
CHAPTER THREE ......................................................................................................................... 41
METHODOLOGY ......................................................................................................................... 41
Theoretical Foundations .............................................................................................................. 41
   Interpretive Research .............................................................................................................. 41
   Interpretive Inquiry and the Focus of My Study .................................................................... 42
Symbolic Interactionism ............................................................................................................. 44
Research Questions .................................................................................................................... 45
Design of the Study ..................................................................................................................... 46
Research Settings ....................................................................................................................... 46
Selection of Participants ............................................................................................................ 47
My Involvement in the Context .................................................................................................. 49
Ethical Considerations .............................................................................................................. 51
Data Collection Methods .......................................................................................................... 51
   Classroom Observations ........................................................................................................ 51
   Interviews ............................................................................................................................... 52
   Document Analysis ................................................................................................................ 53
Data Analysis ............................................................................................................................... 54
Quality Criteria: Establishing the Rigor of the Research .......................................................... 56
   Trustworthiness ...................................................................................................................... 56
   Hermeneutic Process .............................................................................................................. 57
LIST OF TABLES

Table 1. Research Questions ........................................................................................................ 12
Table 2. Data Collection Matrix .................................................................................................. 45
Table 3. The thematic content of the findings .............................................................................. 62
Table 4. Teachers’ Essential Epistemic Views of Science ............................................................. 70
Table 5. Teachers’ Essential Views of School Science ................................................................. 74
Table 6. Teachers’ Essential Pedagogical Conceptions ................................................................. 105
Table 7. Teachers’ Essential Conceptions of Assessment ............................................................. 125
Table 8. Sample Test Questions .................................................................................................. 131
Table 9. Learning and Assessment Tools ..................................................................................... 141
Table 10. Teachers’ Essential Practices of Assessment ................................................................. 144
Table 11. Teachers’ Essential Epistemic Views of Science ........................................................... 165
Table 12. Teachers’ Essential Views of School Science ................................................................. 165
Table 13. Teachers’ Essential Pedagogical Conceptions ............................................................... 166
Table 14. Teachers’ Essential Conceptions of Assessment ........................................................... 167
Table 15. Teachers’ Essential Practices of Assessment ................................................................. 170
ABSTRACT

The purpose of this study is to understand the professional and structural, political and cultural factors that present challenges to the implementation of assessment reform in science classrooms.

An analysis of recent science education literature and national science education reform documents suggests that change in schools is a complex process that does not happen overnight and is subject to the influence of many professional and structural attributes (Berliner, 2006; Brooks, 2005; Cuban, 1990; Duschl, 1990; Gess-Newsome, Southerland, Johnston & Woodbury, 2003; Southerland & Hutner, in press; Tyack & Cuban, 1995). Structural components include school culture; bell schedule, administration policies and mandates, standards, curriculum and accountability measures. Professional components involve teachers’ epistemic views of science, their pedagogical conceptions and their conceptions of assessment along with knowledge necessary to translate these conceptions into practice (Barnett & Hodson, 2001; Gess-Newsome & Lederman, 1999; Gess-Newsome et al., 2003; Mortimer & Scott, 2003; Shulman, 1986). Education literature suggests both of these components, professional and structural components, shape how teaching takes place, thus, what students learn in science classrooms (Brickhouse, 2006; Duschl, 1990; Gallagher, 2006; Gess-Newsome & Lederman, 1999; Gess-Newsome et al., 2003; Southerland & Hutner, in press).

Assessment plays a significant role in efforts to bring about improvements in the educational system (Brickhouse, 2006; Davis, Genc & Aydeniz, in press; NAS, 2006; NRC, 2001; 2005; Southerland & Hutner, in press). Assessment serves multiple purposes. Assessment can be used to support learning (Abell & Volkmann, 2006; Bell & Cowie, 2001; Black & William, 1998; Brookhart, 2006; Davis et al., in press; Klassen, 2006; Shepard, 2000), to monitor the effectiveness of a particular curriculum (NAS, 2006; NRC, 2005), to evaluate the quality and effectiveness of instruction (Bell & Cowie, 2001; NAS, 2006; NRC, 2001; Shepard, 2000), and to evaluate the efficiency of the school system (Davis et al., in press; Linn, 2000; Popkewitz, 2000). Although assessment can serve multiple purposes, the confusion over learning and achievement as manifested in political initiatives that aim to bring about improvements to the educational system
through standardized testing has reduced the role of assessment in educational reform to accountability (Abell & Volkmann, 2006; Brickhouse, 2006; Darling-Hammond, 2003; Davis et al., in press; DeBoer, 2002; Delandshere, 2002; Southerland & Hutner, in press; Stiggins, 2004). Several science educators, who view learning as more than just what is revealed through a single standardized test, find this approach to educational reform problematic (Abell & Volkmann, 2006; Brickhouse, 2006; Davis et al., in press; DeBoer, 2002; Southerland & Hutner, in press). These scholars maintain that the increasing focus on ensuring high test scores has pressured teachers to reduce the role of assessment to the accountability purposes and the content of science teaching to students’ acquisition of only the knowledge and skills necessary for passing the test.

Although the pressure that the standardized tests generate influences what teachers assess in students’ learning and how they go about assessing what they teach (Brickhouse, 2006; Darling-Hammond, 2003; Popham, 2006; Stiggins, 2004), other factors may also influence how teachers come to assess students’ learning. For instance, teachers’ conceptions of assessment may fail to reinforce the goals of science education reform documents. Furthermore, teachers’ epistemic views of science and their pedagogical conceptions may also have an impact on what teachers assess and how they go about assessing students’ learning in science. Finally, research indicates that the political and cultural structures have an impact on teachers’ conceptions and practices of assessment (Berliner, 2006; Darling-Hammond, 2003; Southerland & Hutner, in press; Stiggins, 2004). It follows that characterizing the challenges of enacting assessment reform in science classroom includes exploring science teachers’ professional knowledge base (epistemic views of science, pedagogical conceptions and their conceptions of assessment), and influences of the cultural and political structures.

In this study, I focus on characterizing three high school science teachers’ conceptions and practices of assessment to develop an in-depth understanding into the problems with the implementation of assessment reform in science classroom. In addition to three teachers’ conceptions and practices of assessment I analyzed the major science education reform document such as NSES (NRC, 1996) as well as policy initiatives such as the No Child Left Behind Act (NCLB) and the Florida Comprehensive Assessment Test (FCAT).
My dissertation reveals that although the political and cultural structures of the school system influence what teachers teach in science classrooms and how they go about assessing students’ learning in science, the fundamental challenge to the implementation of assessment reform in science classroom is teachers’ naïve pedagogical content knowledge (PCK) base. Furthermore, although my findings suggest a nested relationship between science teachers’ epistemic views of science, their pedagogical conceptions and their conceptions of assessment, there is no relationship between science teachers’ teachers’ epistemic views of science, their pedagogical conceptions, their conceptions of assessment and their assessment practices. This is partly due to the influence of the political and cultural structures of the school system and partly due to teachers’ naïve understanding of the nested relationships between various constructs (epistemic beliefs, pedagogical beliefs, assessment beliefs), that make up their conceptual ecology (Southerland, Johnston & Sowell, 2006).

In my conclusions I argue that in order for the assessment reform to take place in science classrooms, teacher educators must help pre-service and in-service teachers to develop a sophisticated PCK base. This involves helping teachers to develop sophisticated epistemic views of science, sophisticated pedagogical conceptions, and sophisticated conceptions of assessment.
CHAPTER ONE

INTRODUCTION TO THE STUDY

As the world becomes increasingly dependent on science and technology, our health, the future of this planet and the growth of economy depend on how wisely we understand and consume the scientific knowledge (American Association for the Advancement of Science [AAAS], 1993; Gore, 2006). The ability to consume scientific knowledge for economical benefits and protecting the environment and human health from the adverse effects of the products of science and technology depends on how much citizens know about science (National Academy of Sciences [NAS], 2006). Therefore, producing scientifically literate citizens has been a core goal of science education reform initiatives (DeBoer, 1991; Gallagher, 2006; Hurd, 1998). Recent reports (AAAS, 1990; Gore, 2006; National Research Council [NRC], 1996; National Commission on Mathematics and Science Teaching for the 21st Century, 2000; NAS, 2006) highlight both the problems endemic to science education and put forward recommendations for addressing the problems with American students’ learning in science (Brickhouse, 2006; Collins, 1998; Southerland & Hutner, in press). These recommendations deal with the definition of scientific literacy, content standards, pedagogy and assessment (Blanchard, 2006; Brickhouse, 2006; DeBoer, 2002; Gallagher, 2006; Radford, 1998; Southerland & Hutner, in press).

Prominent science education reform documents such as *Science for All Americans* (AAAS, 1990), *Benchmarks for Scientific Literacy* (AAAS, 1993) and *The National Science Education Standards (NSES)* (NRC, 1996) provide a new definition for scientific literacy. This definition reflects the demands of the 21st century citizenship (Brickhouse, 2006). The *NSES* (NRC, 1996) defines scientific literacy as, “the knowledge and understanding of scientific concepts and processes required for personal decision making, participation in civic and cultural affairs, and economic productivity” (p. 22). *Science for All Americans* (AAAS, 1990) defines a scientifically literate person as one who is aware that science and technology are human enterprises with associated strengths and limitations, understands key scientific concepts and principles, is familiar with how the
the natural world works and applies scientific knowledge and skills for individual and social purposes.

These definitions place new demands on what we teach in science classrooms, how we teach science, and how we go about assessing students’ learning in science. Therefore, prominent national science education reform documents (AAAS, 1990; 1993; NRC, 1996; 2001; NAS, 2006) put forward recommendations and provide guidelines for content standards, pedagogy and assessment to ensure students’ acquisition science literacy goals outlined by the *Benchmarks for Scientific Literacy* (AAAS, 1993) and *NSES* (NRC, 1996). Science educators argue that in spite of these recommendations, scant changes have been observed in what science teachers teach in science classrooms, how they teach science, and how they go about assessing students’ learning in science (DeBoer, 1991; 2002; Firestone, 2001; Gallagher, 2006; Gess-Newsome, Southerland, Johnston & Woodbury, 2003; Kennedy, 2005; NAS, 2006; NRC, 1996; NRC, 2001; Settlage & Meadows, 2002; Southerland & Hutner, in press). These scholars maintain that although improvements are being made in content standards, traditional pedagogies and assessment methods still dominate science classrooms across the nation (Brickhouse, 2006; Gallagher, 2006; Genc, 2005; Haney, Lumpe & Czerniak, 2002; Southerland & Hutner, in press). Furthermore, these scholars argue the traditional pedagogical and assessment practices are to blame for students’ low achievement scores in science on national and international tests.

Several national and international test results reveal American students’ poor performance in science. For instance, The *Nation’s Report Card* (National Assessment of Educational Progress [NAEP], 2006) indicates that despite increased funding and attention to science education, the achievement of 12th graders in science declined since 1996 and the achievement of 8th graders in science has not shown significant change since 1996. Although a significant number of scholars question whether the test results reflect students’ learning in classroom or not (Berliner, 2006; Brickhouse, 2006; Popham, 2006), the results of these tests have encouraged many parents, scientists, politicians and business executives to question the quality of science education that students receive in public schools. These concerned citizens maintain that public school science classrooms fail to help students achieve the knowledge and skills that prominent
science education reform documents promote (Collins, 1998; Gallagher, 2006; Haney, Lumpe, Czerniak, 2002; Jeanpierre, Oberhauser & Freeman, 2005; NAEP, 2006; NAS, 2006; Perkins, 1992; Radford, 1998; Southerland & Hutner, in press). Furthermore, several reports reveal that only a small number of students are in science classrooms that cultivate the habits of mind necessary for scientific literacy advocated by the prominent science education reform documents (NAEP, 2006; NAS, 2006). These reports emphasize that a great number of students have teachers who do not have certification in science and that the science curriculum is not challenging enough for preparing students for the 21st century economy (NAS, 2006). Moreover, students participate in laboratory activities that require them to follow “cook-book” procedures to conduct experiments, which, in turn does not help them develop scientific habits of mind necessary for fully understanding science (AAAS, 2005; NAS, 2006).

Several science educators maintain that students’ learning in science fails to meet nation’s expectations partly because the implementation of science education reform has been reduced to formulating rigorous content standards and developing standardized tests (Brickhouse, 2006; DeBoer, 2002; Settlage & Meadows, 2002; Southerland & Hutner, in press). They argue that although major improvements have taken place in the quality of content standards both at national and local levels (Brickhouse, 2006; Delandshere, 2002; Gallagher, 2006), the quality of science teaching and the nature of assessments in science classrooms does not measure up to the recommendations put forward by national reform documents (AAAS, 1993; Blanchard, 2006; Brickhouse, 2006; Carlone, 2003; DeBoer, 2002; NRC, 1996; Radford, 1998). This discrepancy is due, in part, to the pressures of standardized tests (Abell & Volkmann, 2006; Brickhouse, 2006; DeBoer, 2002; Delandshere, 2002; Settlage & Meadows, 2002; Southerland & Hutner, in press). These observations suggest that science education reform is failing in its goal of ensuring scientific literacy for all students. In this study, I focus on science teachers’ conceptions and practices of assessment to understand why reform efforts in science education fail to produce envisioned outcomes.

Statement of Problem and Purpose

Science educators maintain that science education reform fails partly because the
assessment methods that teachers use do not serve the purposes of science education reform (Abell & Volkmann, 2006; Brickhouse, 2006; Davis, Genc & Aydeniz, in press; Southerland & Hutner, in press). These scholars argue that recent federal and state policies that aim to bring about improvements to academic performance of public schools through standardized testing pressures teachers to maintain status quo in assessment, thus, further complicates the enactment of reform in science classrooms (Collins, 1998; Delandshere, 2002; Southerland & Hutner, in press).

Assessment plays a central role in efforts to bring about improvements in the educational system (Brickhouse, 2006; Davis et al., in press; NAS, 2006; NRC, 2001; 2005; Southerland & Hutner, in press). For instance, assessment can support learning (Abell & Volkmann, 2006; Bell & Cowie, 2001; Black & William, 1998; Brookhart, 2006; Davis et al., in press; Klassen, 2006; Shepard, 2000), help monitor the effectiveness of a particular curriculum (NAS, 2006; NRC, 2005), evaluate the quality and effectiveness of instruction (Bell & Cowie, 2001; NAS, 2006; NRC, 2001; Shepard, 2000), and enhance the efficiency of the school system (Davis et al., in press; Linn, 2000; Popkewitz, 2000). Although assessment can serve many purposes, the confusion over learning and achievement, as manifested in political initiatives that aim to bring about improvements to the educational system through standardized testing reduces the role of assessment in educational reform to accountability (Abell & Volkmann, 2006; Brickhouse, 2006; Darling-Hammond, 2003; Davis et al., in press; DeBoer, 2002; Delandshere, 2002; Southerland & Hutner, in press; Stiggins, 2004). Several science educators, who view learning as more than just what is revealed through a single standardized test, find this approach to educational reform problematic (Abell & Volkmann, 2006; Brickhouse, 2006; Darling-Hammond, 2003; Davis et al., in press; DeBoer, 2002; Southerland & Hutner, in press). These scholars maintain that increasing focus on ensuring high test scores pressures teachers to reduce the role of assessment to the accountability purposes and the content of science teaching to students’ acquisition of only the knowledge and skills necessary for passing the test.

Although the pressure that the standardized tests generate influences what teachers assess in students’ learning and how they go about assessing what they teach (Brickhouse, 2006; Darling-Hammond, 2003; Popham, 2006; Stiggins, 2004), other
factors may also influence how teachers come to assess students’ learning. For instance, teachers’ conceptions of assessment may fail to reinforce the goals of science education reform documents. Furthermore, teachers’ epistemic views of science and their pedagogical conceptions may also have an impact on what teachers assess and how they go about assessing students’ learning in science. Finally, research indicates that the political and cultural structures have an impact on science teachers’ conceptions and practices of assessment (Berliner, 2006; Darling-Hammond, 2003; Southerland & Hutner, in press; Stiggins, 2004). It follows that characterizing the challenges of enacting assessment reform in science classroom includes exploring teachers’ professional knowledge base (epistemic views of science, pedagogical conceptions, conceptions of assessment), and the cultural and political influences on their conceptions and practices of assessment.

In this study, I focus on characterizing three high school science teachers’ conceptions and practices of assessment to develop an in-depth understanding of the problem in science education reform implementation. However, I find it useful to introduce the calls for reform in science education prior to discussing the importance of science teachers’ conceptions and practices of assessment for the enactment of assessment reform in science classrooms.

**Calls for Reform in Science Education**

An analysis of recent science education literature and national science education reform documents suggests that change in schools is a complex process that does not happen overnight and is subject to the influence of many professional and structural attributes (Berliner, 2006; Brooks, 2005; Cuban, 1990; Duschl, 1990; Gess-Newsome, Southerland, Johnston & Woodbury, 2003; Southerland & Hutner, in press; Tyack & Cuban, 1995). Structural components may include school culture, bell schedule, administration policies and mandates, standards, curriculum and accountability measures. Professional components involve teachers’ epistemic views of science, their pedagogical conceptions and their conceptions of assessment along with knowledge necessary to translate these conceptions into practice (Barnett & Hodson, 2001; Gess-Newsome & Lederman, 1999; Gess-Newsome et al., 2003; Mortimer & Scott, 2003; Shulman, 1986).
Literature suggests both of these components shape how teaching takes place, therefore, what students learn in science classrooms (Brickhouse, 2006; Duschl, 1990; Gallagher, 2006; Gess-Newsome & Lederman, 1999; Gess-Newsome et al., 2003; Southerland & Hutner, in press). Assessment serves as a special context in which we can examine the ways in which structural and professional aspects of science teaching profession interact and influence students’ learning. Assessment serves as a special context partly because it enables us to understand how teachers entertain both structural and professional calls for reform in science education to improve students’ learning in science. I describe the nature of the calls for improvement in students’ learning in science through two different perspectives; structural and professional, in the following passages.

Major science education reform documents; *Benchmarks for Scientific Literacy* (AAAS, 1993) and NSES (NRC, 1996) and policy initiatives; *No Child Left Behind Act (NCLB)* (U.S. Department of Education, 2002), *Florida Comprehensive Assessment Test (FCAT)* (Florida Department of Education, 2004) aim to bring about change in how public schools function. While science education reform documents focus on improving the professional quality of teachers to better serve students’ learning needs in science, policy initiatives aim to bring about improvements in students’ learning of science through accountability measures in the form of standardized testing. The aim of the national science education reform documents is to bring about change in that we teach, along with change in pedagogies and assessment methods, to promote science literacy goals (AAAS, 1993; NRC, 1996; 2001). The purpose of the national and local policy initiatives, on the other hand, is improving the efficiency of the educational system through accountability (Delandshere, 2002; Linn, 2000; William, 2001).

Accountability addresses structural aspects of change pertinent to assessment (Brickhouse, 2006, Collins, 1998; Delandshere, 2002; Southerland & Hutner, in press). By structure, I mean exterior political influences such as NCLB and Florida Comprehensive Assessment Test (FCAT) aiming to bring about change in schools. The purpose of the exterior structures is to bring about improvements in educational system by focusing on the efficiency of the system (Brickhouse, 2006, Collins, 1998; Davis et al., in press; Linn, 2000; Popkewitz, 2000; Southerland & Hutner, in press). To meet structural purposes, local administrators design assessments (often in the form of
standardized tests) to give policy makers a means to evaluate the effectiveness of curriculum and instruction (Brickhouse, 2006, Davis et al., in press; Delandshere, 2002; Southerland & Hutner, in press). Although the politicians put substantial efforts into designing accountability measures such as NCLB Act to address structural component of educational reform, they invest limited effort into improving the professional aspect of science education reform agenda (Duschl, 1990; Gallagher, 2006; Southerland & Hutner, in press). Professional aspects of the reform agenda involve teachers’ knowledge of subject matter, the nature of their epistemic views of science, their pedagogical conceptions and their conceptions of assessment (Gess-Newsome & Lederman, 1999; Shulman, 1986; Tsai, 2002). Several scholars maintain that the quality of professional aspects of reform, teachers’ internally held conceptions, have far greater influence on students’ learning in science than the external measures such as testing (Darling & Hammond, 2003; Davis et al., in press; Delandshere, 2002; Shepard, 2000; Southerland & Hutner, in press; Tsai, 2002). These scholars argue that although standardized tests may be able to provide useful information on the efficiency of schools, they fail to create conditions necessary to bring about improvements in students’ learning in science. Many argue that in order for assessment to improve students’ learning in science and serve the purposes of science education reform documents, teachers’ conceptions of assessment must be compatible with social constructivist pedagogy that underpins the reforms (Beck, 1997; Delandshere, 2002; NAS, 2006; NRC, 2001; Shepard, 2000; Southerland & Hutner, in press).

An understanding of assessment that is informed by the social constructivist pedagogy acknowledges that learners construct knowledge through communication by using cultural tools in social and technological systems (Marshall, 2006; Shepard, 2000). Social constructivist pedagogy focuses on students and views learning as the object of inquiry (NRC, 2001; Shepard, 2000). The fundamental premise of social constructivist accounts of assessment is that teachers use assessments to create fruitful conditions for students’ learning rather than as a punishment or reward system (Davis et al., in press; Marshall, 2006; NRC, 2001; Shepard, 2000). Teachers can structure assessment so that students feel responsible for their learning, feedback is immediate, challenging and yet encouraging (Bell & Cowie, 1997; 2001; Brookhart, 2006; Davis et al., in press; NRC,
My discussion so far highlights two perspectives of the role of assessment in bringing about improvements in students’ learning in science. Although structural understanding of assessment as an external measure focuses on system improvement in terms of efficiency, the professional understanding of assessment as an internally held construct focuses on a student’s growth in a particular learning domain (Davis et al., in press; Delandshere, 2002; Popkewitz, 2000; Shepard, 2000). Furthermore, while both of these views of assessment are necessary for students’ acquisition of knowledge and skills advocated by science education reform (Collins, 1998; Darling-Hammond, 2003; Davis et al., in press; Delandshere, 2002; Linn, 2000; Southerland & Hutner, in press), internal perspectives on learning has been ignored due to increasing emphasis on exterior accountability through standardized testing (Abell & Volkmann, 2006; Brickhouse, 2006; Davis et al., in press; Delandshere, 2002; Linn, 2000; NAS, 2006; NRC, 2005; Southerland & Hutner, in press).

Scholars argue that limits for success or failure of any educational reform are due to the way teachers internalize and respond to the calls for improvement in pedagogy and assessment (Carlone, 2003; Cuban, 1990; Duschl, 1990; Gess-Newsome et al., 2003). It follows that without explicit attention to teachers’ internally held conceptions, the data gathered through external measures such as standardized tests provide limited information on the sources of the failing state of science education reform (Borko, Mayfield, Marion, Flexer, & Cumbo, 1997; Brickhouse & Bodner, 1992; Pajares, 1992). In this study, I aim to develop an in-depth understanding into teachers’ beliefs, internally held conceptions that guide their instructional and assessment decisions in light of the influences of political structures.

Significance of Teachers’ Conceptions of Assessment

Richardson (1996) defines beliefs as "psychologically held understandings, premises, or propositions about the world that are felt to be true” (p. 103). Thompson (1992) introduces the term “conceptions” as a general mental structure defining teachers’ conceptual ecology which encompasses teachers’ beliefs, knowledge and dispositions (Tsai, 2002). Pratt (1992) maintains that beliefs are a culturally defined lens through which people make sense of events, symbols and interactions. Research on beliefs
indicates that teachers’ beliefs play a significant role in how they interpret pedagogical knowledge, conceptualize teaching, and scaffold instruction (Brickhouse & Bodner, 1992; Bryan, 2003; Haney, Czerniak, & Lumpe, 1996; Pintrich, 1990). It follows that beliefs serve as filters in the individual’s processing of new knowledge, allowing some information to become part of existing conceptual structures, which reinforces previous beliefs; or discarding information that may not match previously held pedagogical conceptions (Gess-Newsome & Lederman, 1999; Yalaki, 2004).

Previous research suggests that teachers’ conceptions of teaching and learning significantly influence how they go about scaffolding instruction (Borko, Mayfield, Marion, Flexer, & Cumbo, 1997; Calderhead, 1996; Kahn, 2000; Pajares, 1992; Tittle, 1994; Thompson, 1992). It follows that conceptions act as a framework through which a teacher views, interprets, and interacts with various aspects of classroom practice (Marton, 1981). Although a number of studies look at teachers’ conceptions of teaching and learning (Eslinger & Metz, 2005; Ly & Metz, 2006; Tal, Geier & Krajcik, 2000) and the relation of these two conceptions to their epistemic view of science (Tsai, 2002), teachers’ conceptions of assessment have received limited attention from science educators (Dahlin, Watkins & Ekholm, 2001; Genc, 2006). This omission undermines the integrity of science education reform and the broader research agenda into those reforms.

Science education reform has three components: curriculum, pedagogy and assessment (Brickhouse, 2006; DeBoer, 2002; Southerland & Hutner, in press). Therefore, an attempt to understand the challenges of enacting reform in science classrooms mandates that we examine teachers’ epistemic conceptions of science, their pedagogical conceptions and their conceptions of assessment. Such an approach will allow us to understand the coherences or contradictions between these different conceptions that influence how teachers teach science and assess students’ learning in science (Tsai, 2002). In this study, I attempt to understand how three high school science teachers’ conceptions and practices of assessment 1) are consistent with social constructivist pedagogy and reinforce or contest the science literacy goals outlined in reform documents.

My argument is that a comprehensive understanding of these constructs that make up a teacher’s conceptual ecology for assessment holds promise for identifying
contradictions and finding ways to resolve them.

Significance of Assessment in Science Education Reform

As has been discussed assessment has been an integral component of calls for reform in science education for the last seven to eight decades (Delandshere, 2002; Klassen, 2006; NRC, 2006; Rothman, 1995). Assessing student learning through standardized forms of testing frequently fulfills the role of assessment in political efforts to reform education (Berliner, 2006; Firestone, 2001; Sacks, 1999; Shepard, 2000; Stiggins, 2004). These external measures limit the role of assessment in educational reform to accountability purposes (Brickhouse, 2006; Davis et al., in press; Popkewitz, 2000). However, we know that assessment can serve multiple purposes. Several national science education reform documents maintain that in order to bring about improvements in students’ learning in science, assessment should be part of learning (NAS, 2006; NRC, 2001; 2005). This argument has been supported by empirical evidence, which suggests that when teachers use assessment to support learning, their students’ academic performance significantly increases (Abell & Volkmann, 2006; Black & William, 1998; NRC, 2001; Pringle, 2000). However, recent national and local policy initiatives undermine the role that assessment may play in improving students’ learning in science.

Firestone (2001) maintains that schools are social places, structured through political rules and regulations, and surrounded by cultural values, norms and expectations that both teachers and students are expected to abide and understand. Therefore, teachers are strategically sensitive and vulnerable to how power, politics (as it appears in the form of accountability) and culture surrounding their practice function to influence their instructional and assessment decisions (Delandshere, 2002; Little & McLaughlin, 1993; Moore, 2004). In other words, teachers may react to the pressure from political influences such as administrative demands for teaching to the test, and ignore the recommendations from documents such as the National Science Education Standards (NSES) (NRC, 1996) that promote social constructivist accounts of assessment, just as they have failed to “take up” constructivist accounts of pedagogy (Abell & Volkmann, 2006; Arter, 2003; Black & William, 1998; Davis et al., in press; Firestone, 2001; Klassen, 2006; Shepard, 2000; Southerland & Hutner, in press). An understanding of assessment informed by social constructivism focuses on the quality of student work and their learning rather than their
grades. Social constructivist based understanding of assessment focuses on how students make judgments as they learn and how teachers may be able to structure those judgments in a way to guide their students’ learning (NRC, 2001; Shepard, 2000). It follows that social constructivist accounts of assessment are more likely to ensure students’ learning in science in ways that mesh with the goals of science education reform documents than standardized tests that measure students’ acquisition of canonical knowledge. Therefore, if reforms are to occur, science teachers must construct the understanding and knowledge necessary for frequent use of assessment practices that are consistent with social constructivist pedagogy.

Several scholars argue the NCLB Act (U.S. Department of Education, 2002), which promotes use of standardized testing to measure student learning, leads teachers to abandon assessment practices that address students’ varying learning needs in favor of those practices that ensure high test scores (Brickhouse, 2006; DeBoer, 2002; Shepard, 2000; Southerland & Hutner, in press; Stiggins, 2004). This argument assumes that most science teachers have the knowledge, understanding, skills and motivation to enact assessment practices that attend to students’ varying learning needs and ensures students’ acquisition of science literacy goals defined by reform documents such as NSES (NRC, 1996). However, scarce evidence exists to suggest that teachers indeed have knowledge and skills necessary to scaffold instruction and design assessment around the goals of science education reform documents (Genc, 2006; NRC, 2001; 2005; Southerland & Hutner, in press). Several scholars maintain that even when teachers have necessary knowledge of assessment to reinforce the science literacy goals outlined in major science education reform documents, they may not be able to use that knowledge to support students’ learning due to political and cultural influences (Barnett & Hodson, 2001; Darling-Hammond, 2003; Matase, Griesdorn & Borko, 2002). It follows that understanding the challenges of implementing assessment reform in science classrooms requires understanding teachers’ professional capacity (e.g., beliefs and knowledge) and exploring the perceived structural, cultural and political influences, on teachers’ conceptions and practices of assessment. My research questions reflect this awareness.

The following questions will guide my inquiry into professional and structural challenges to the implementation of assessment reform in science classrooms.
Table 1. Research Questions

Main Research Question:

What are the perceived challenges associated with changing the culture of assessment practices in science classrooms at a public high school?

Sub Questions:

Question 1: How do science teachers’ professional knowledge and understanding of assessment influence enactment of assessment reform?

Question 2: How does school culture influence science teachers’ assessment practices?

Question 3: How do policies and political structures affect science teachers’ assessment practices?
CHAPTER TWO
LITERATURE REVIEW

I start this chapter by introducing prominent national science education reform documents; *Project 2061* (AAAS, 1985) and *NSES* (NRC, 1996) followed by the implications of these reform documents for the assessment of students’ learning in science. Then, I discuss how national and local accountability initiatives conceptualize assessment in an effort to bring about improvements in the educational system. In my discussion, I focus on exploring the tensions between two national policy trends; one establishing standards for content, pedagogy and assessment to improve students’ learning in science and the other focusing on accountability through standardized testing to bring about perceived improvements in the educational system.

National Reports and Standards Movement

National reports in the mid 1980s through 2000 reflect a state of crisis in science education in American schools: declaring the failure of the educational system to produce scientifically literate citizens (Brickhouse, 2006; Collins, 1998; DeBoer, 2002; Delandshere, 2002; National Commission on Excellence in Education, 1983; National Assessment of Educational Progress [NAEP], 1983; 2006; NRC, 1996; Yager, 2005). Although several scholars maintain that the policy makers use a language of crisis for political reasons (Collins, 1998; Popham, 2001; Sacks, 1999), these reports led to an array of reform initiatives in science education (Berliner, 2006; Brickhouse, 2006; Collins, 1998; Linn, 2000; 2003; Stiggins, 2004). In the following passages, I discuss the most prominent science education reform initiatives of the last twenty years, *Project 2061* (AAAS, 1985) and *NSES* (NRC, 1996), and the implications of these initiatives for assessment reform in science in light of wider assessment literature.

*Project 2061*

*Project 2061* started as a long term initiative of the American Association for the Advancement of Science (AAAS) to help improve K-12 education so that all high school graduates in America are literate in science, mathematics, and technology (AAAS, 2005). In its 1989 publication *Science for All Americans* (AAAS, 1989), *Project 2061* set out
recommendations for what all adult citizens should know and be able to do in science, mathematics and technology (Collins, 1998; Yager, 2005). *Science for All Americans* (AAAS, 1989) describes the nature of scientific knowledge along with the specific knowledge and skills necessary for students to develop a scientific habit of mind.

To reinforce the goals of *Science for All Americans* (AAAS, 1989), AAAS published a document called *Benchmarks for Scientific Literacy* (AAAS, 1993). *Benchmarks* serve as a curriculum framework, which educators can use to design their own curriculum to help students achieve the basic science literacy goals outlined in *Science for All Americans* (AAAS, 1989). To help science educators design their own curriculum (AAAS, 2005), *Benchmarks*:

- describes the specific levels of understanding, critical thinking and inquiry skills that we expect all students to acquire on their way to becoming scientifically literate.
- focuses on the use of teaching methods that contribute to students’ acquisition of the basic science literacy goals outlined in *Science for All Americans* (AAAS, 1989).
- calls for improvements in the assessment methods used to evaluate students’ learning in science.

Because research on how students learn informs *Benchmarks*, many science educators think highly of the changes that this document recommends (Brickhouse, 2006; DeBoer, 2002; Gallagher, 2006; Southerland & Hutner, in press). State departments of education used *Benchmarks* for developing state curriculum standards. Educators use *Benchmarks* to make informed decisions on adopting new curriculum materials, curriculum developers use *Benchmarks* to improve their existing instructional materials and finally colleges of education across the country incorporate *Benchmarks* in their science method courses as they prepare teachers for science classrooms (AAAS, 2005; Brickhouse, 2006; Collins, 1998; Southerland & Hutner, in press). Since publication of the *Benchmarks*, science educators demonstrate heightened interest in improving pedagogical practices in science classrooms (AAAS, 2005). Several scholars maintain that assessment has received the least attention in efforts to reform science education since publication of the *Benchmarks* (Abell & Volkmann, 2006; Brickhouse, 2006;
Southerland & Hutner, in press). In fact, to address concerns about the alignment between curriculum goals, pedagogy and assessment, AAAS recently formed a committee on assessment to make sure that science assessments support students’ acquisition of specific learning goals outlined in the *Benchmarks for Scientific Literacy* (AAAS, 2005).

**National Science Education Standards**

Following the *Benchmarks for Scientific Literacy* (AAAS, 1993), the National Research Council (NRC) released *National Science Education Standards (NSES)* in 1996, which set the foundation for the nationwide science education reform efforts. Although *Benchmarks* (AAAS, 1993) outlines specific learning goals in science so that educators can determine what is important for students to know, the *NSES* (NRC, 1996) builds on recent research about how students learn makes specific recommendations for how to teach science and assess students’ learning in science (NRC, 1996). The *NSES* (also referred to as the *Standards*) envision the following improvements in science education.

1. To make scientific literacy a reality through reflecting a broad and integrated body of scientific knowledge in content area specifications.

2. To foster students’ acquisition of process learning skills in science by focusing on teaching science through an inquiry framework.

3. To foster use of alternative assessment strategies with “assessment for learning” notion in mind

The *Standards* provide a new definition for the scientific literacy, thus, for what the American students need to learn in science. The *NSES* defines scientific literacy as “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). The *Standards* recommend changes in pedagogy and assessment as well. The *NSES* (NRC, 1996) highlight the importance of teaching science through an inquiry framework and frequent use of collaborative learning methods for improved student learning in science. The *NSES* also focus on the shift of power relationships between the teacher and students (Davis & Helly, 2004). The *Standards* advocate students’ autonomy by encouraging teachers to design learning activities that will engage students in open-ended inquiry and a search for meaning instead of requiring
them to match their learning with teachers’ expert knowledge (Blanchard, 2006; Davis & Helly, 2004). Since the Standards introduce a set of new science literacy goals and advocates alternative pedagogies to support students’ learning in science, alternative assessment formats must accommodate these changes. The Standards (NRC, 1996), in fact, advocate for alternative forms of assessment to ensure students’ acquisition of science literacy goals.

How Do Reform Documents Inform Science Educators?

Both the NSES (NRC, 1996) and the Project 2061 (AAAS, 1985) provide a vision for teaching and learning important content and skills in science along with the guidelines for effective pedagogy, and the criteria to measure how students progress towards that vision (Collins, 1998). As Collins (1998) argues, because of the call for alternative pedagogies and assessment methods, reform documents open doors to a number of research questions about:

…how and how much students are able to achieve, about teachers and how they conduct and learn to conduct their practice, about how to assess what students have come to understand and be able to do, about materials that promote high-quality scientific literacy, and about the supports and constraints to classrooms and schools as communities of learners. (Collins, 1998, p. 725).

Although research focusing on content standards and pedagogy has received visible attention since the publication of the reform documents, there has been scarce interest in assessment from science education community. This has significant implications for the integrity and success of science education reform. For instance, narrowing the focus of science education research merely to pedagogy is unlikely to bring about improvements to science education in public schools (Southerland & Hutner, in press).

Brickhouse (2006) maintains that the path that we take in our research shapes the understandings and knowledge that we make accessible to policy makers, curriculum developers and the practicing community. Brickhouse’s (2006) argument has several implications for various stakeholders involved in reforming science education. First, it urges science educators to form the theoretical foundations of alternative forms of assessment as well as educate teachers to use alternative assessment to support students’
learning in science. Secondly, she urges curriculum developers to use the theoretical knowledge that science educators produce to design assessment materials that will ascertain students’ acquisition of skills and knowledge that the Standards promote. Finally, it urges policy makers to invest money into purchasing curriculum products that assess students’ learning in science through alternative means and fund professional development programs that will provide teachers with the knowledge and skills necessary to use alternative methods for assessing students’ learning in science. In the following passages I discuss the purposes that assessment can serve and focus on the need for and importance of using alternative methods for assessing students’ learning in science.

**Purposes of Assessment**

In efforts to improve students’ learning in science across grade levels, assessment can serve multiple purposes (NRC, 2001). Assessment can be used:

1. To monitor student progress, identify misconceptions and provide support for subsequent learning.
2. To diagnose students’ performance and make decisions about students’ future learning.
3. To plan lessons and scaffold instruction.
4. To improve the quality of curriculum materials and assessment instruments.
5. To formulate educational policy at the local and national level.

It follows that because assessment has multiple functions and has the potential to serve the information needs of different stakeholders, multiple forms of assessments must exist to bring about improvements to students’ learning. For instance, assessment forms used for institutional accountability purposes must be different from the assessment formats used for identifying students’ learning needs (Davis et al., in press; Delandshere, 2002). While assessment can serve different purposes, scholars argue that recent policies aiming to bring about improvements to the educational system through standardized testing has reduced the role of assessment in students’ learning to accountability purposes (Darling-Hammond, 2003; Delandshere, 2002; Linn, 2000; NRC, 2005; Southerland & Hutner, in press; Stiggins, 2004). Although accountability measures in the form of standardized testing may ensure system efficiency (Brickhouse, 2006), assessment of
students’ learning through standardizing tests does not help students to continuously
improve as they learn science (Davis et al., in press). Furthermore, several scholars
maintain the type of knowledge and skills tested on standardized tests do not reinforce the
science literacy goals of science education reform (Brickhouse, 2006; Southerland &
Hutner, in press). This is partly because standardized tests reduce students’ learning to
acquisition of prescribed scientific facts and laws (Abell & Volkmann, 2006; NAS, 2006;
Southerland & Hutner, in press). If the assessment methods used do not reinforce science
literacy goals outlined in reform documents, the standards alone are unlikely to result in
significant increases in students’ learning in science (AAAS, 2005; Abell & Volkmann,
2006; Delandshere, 2002; Southerland & Hutner, in press; Stiggins, 2004). Furthermore,
the lack of alignment between the learning goals that the reform documents promote and
the assessments used to measure students’ learning calls for finding ways to encourage
teachers to use new forms of assessments to ensure students’ acquisition of science
literacy goals defined in NSES (NRC, 1996).

The Need for Alternative Forms of Assessment

Assessment of student learning in science must change for several reasons. First,
an analysis of recent science education research and of prominent national science
education reform documents, Benchmarks for Scientific Literacy (AAAS, 1993) and
NSES (NRC, 1996), reveals that the rote memorization of facts once deemed important
and prerequisite to the learning of science starts to give way to the acquisition of inquiry
skills, ability to collect and analyze data and ability to make connections between science
content learned in the classroom and its applications in real life (Brickhouse, 2006;
Gallagher, 2006; Linn, Clark, & Slotta, 2003; NRC, 1996; NAS, 2006; Roth, 1995;
Southerland & Hutner, in press). Secondly, today students face a job market that demands
new knowledge and skills that cannot be reinforced through traditional means of
assessment (Abell & Volkmann, 2006; Brickhouse, 2006; Collins, 1998; Matese,
Griesdorn & Borko, 2002; NAS, 2006; NRC, 2005). Although students still need an
understanding of the fundamental principals of science, workers need to think critically,
weigh the validity and viability of information, analyze, and make inferences (NAS,
2006). Helping students develop these skills requires significant changes in science
assessments- both summative and formative (Southerland & Hutner, in press).
Furthermore, although collaborative learning experiences (as highly advocated by Standards) are seen as a significant context for improved learning, traditional accounts of assessment do not provide opportunities for students to collaborate because students and parents view assessment as a system of reward and punishment for academic performance (Bell & Cowie, 2001; Shepard, 2000). Finally, while reform-based understanding of assessment strongly places an emphasis on students’ autonomy, traditional assessments encourage students to submit to the teacher authority for learning science (Black & William, 1998; Cook-Sather, 2002; Davis et al., in press).

Although the need for change in assessment of students’ learning is clear, political will and teachers’ professional capacity to implement reform-based assessment ideas are not present (Abell & Volkmann, 2006; Delandshere, 2002; Gallagher, 2006; Stiggins, 2004). Science educators maintain that in order for real change to take place in the culture of assessment in science classrooms, teachers’ internally held conceptions must be consistent with the most current educational research and supportive of the goals that the reform documents promote (Gallagher, 2006; Southerland & Hutner, in press). Teachers’ internally held conceptions involve their epistemic views of science, their pedagogical conceptions and their conceptions of assessment (NRC, 1996). However, in order for teachers to hold conceptions that are consistent with the most current research so they can help students acquire the skills and knowledge that science education reform promote, teachers need to be provided with ample educational opportunities to take a critical look at their internal conceptions (Davis, 1996). According to Davis (2006) unless we provide teachers with the educational opportunity to critically reflect on their beliefs and practices, the external measures such as standards and testing are unlikely to help teachers hold conceptions that are supportive for students to learn science across the goals of science education reform (personal communication, September, 4, 2006).

Science Education Reform and Assessment

Embedded in any forms of classroom assessment are the values that we consider as important for students to know (Abell & Volkmann, 2006; Brickhouse, 2006; Brookhart, 2006; Klassen, 2006; Linn, 2000; NAS, 2006; NRC, 2005; Southerland & Hutner, in press). How we view schools heavily influences what we want our schools to accomplish. Labaree (2000) (as cited in Southerland & Hutner, in press) maintains that
the American educational system functions to serve three competing educational goals:

a) *social efficiency*: schools function to produce knowledgeable experts, prepare competent talents for the advancement of economical growth.

b) *democratic equality*: schools are to produce responsible citizens with adequate knowledge and skills to make informed judgments.

c) *social mobility*: schools are to graduate individuals with competitive skills in the struggle for advantageous social status. From the social mobility perspective, the achievement of academic credentials provides an array of opportunities for children to climb the ladder in the job market and maintain or improve social status (Labaree, 2000).

An analysis of recent policy and reform documents conveys mix messages on the purposes of science education (Collins, 1998). For instance, teaching science to produce expert skills that industry needs for economical productivity is one of the core arguments of educational policies and science education reform initiatives (Brickhouse, 2006; Gallagher, 2006; Marshall, 2006; NAS, 2006). Several scholars maintain that in addition to their role to prepare expert skills for industry, politicians pressure schools to address social and racial inequality (Collins, 1998; Harrington & Harris, 2006). For instance, the *NCLB* Act (U.S. Department of Education, 2002) maintains that all students regardless of their ethnic, racial or socio-economic backgrounds have the right to sufficient resources and support for developing scientific literacy, and are capable of gaining knowledge and strength to pursue careers in science.

Several scholars maintain by using science education for political interests, the *NCLB* Act (U.S. Department of Education, 2002) reduces the role of science education to social mobility (Collins, 1998). For instance, instead of focusing on fostering a scientific habit of mind in science classrooms, the law focuses on closing the achievement gap between different sects of society. However, even then the *NCLB* Act widens the achievement gap between different sects of society and fails to serve the purpose for which it was designed (Berliner, 2006; Darling-Hammond, 2003). For instance, research shows that increasing focus on standardized testing negatively influences the learning of minority and socio-economically disadvantaged student populations (Popham, 2006; Sacks, 1999; Stiggins, 2004).
Furthermore, critics argue that reducing educational reform to social mobility is in contrast to what science education reform initiatives are trying to accomplish (Brickhouse, 2006; Southerland & Hutner, in press). They argue that science’s ability to empower citizens for making informed decisions about controversial issues justifies the teaching of science in schools (AAAS, 1993; NRC, 1996). Brickhouse (2006) argues, “Contemporary democracies depend on an educated citizenry that can make reasoned decisions about the products of the sciences and can influence the sciences in productive ways” (p. 3). Brickhouse’s (2006) argument suggests that education for democratic participation is the fundamental focus of recent science education reform initiatives (Gallagher, 2006; Southerland & Hutner, in press). However, the policy initiatives that pressure schools to address social inequities between different sects of society create tension that acts against the achievement of that goal. For instance, some policy initiatives focus on test scores to bring about social equity instead of producing critical thinkers who can truly participate in a democratic society and demand power from the political authority and influence the future of the nation.

If a purpose of science education is to teach for democratic participation, the pedagogy and assessment methods used must serve that purpose (Southerland & Hutner, in press). A significant number of science educators are active in changing the traditional pedagogies dominating science classrooms across the nation. However, research on assessment of students’ learning in science has gained scarce interest from science educators. Southerland and Hutner (in press) maintain, “To complement this pedagogy and educational goals, assessments that measure students’ knowledge and their abilities to apply that knowledge would be a visceral part of such inquiries, both on a formative and summative level” (p. 9). Summative assessment is an indication of what students have learned throughout the year or throughout a lesson unit (Brickhouse, 2006). Formative assessment, on the other hand, is ongoing and serves as a means for shaping the nature of student learning as the learner makes progress towards accomplishing their learning goals (Bell & Cowie, 2001; Black & William, 1998, Delandshere, 2002; Shepard, 2000). Both formative and summative assessments can play significant roles in making judgments about student learning and about the course of instruction (Davis et al, in press; NAS, 2006; NRC, 2001; Shepard, 2000; Southerland & Hutner, in press).
However, to ensure science literacy for all, assessments must measure students’ learning against the same standards (Southerland & Hutner, in press). Southerland and Hutner (in press) maintain that if classroom assessments are different in content and format than the content of the summative assessments, the results of the summative assessments are likely to give us false readings about students’ learning and about the academic performance of schools.

A number of researchers maintain that how teachers design their assessments and what they do with the information they collect from their assessments informs us about what they value in their teaching (Brookhart, 2006; Davis et al., in press; Klassen, 2006). Many argue that in spite of the changes in the goals for scientific literacy, both classroom assessments and summative assessments have remained traditional (Abell & Volkmann, 2006; Delandshere, 2002; Gallagher, 2006; NAS, 2006). They have remained traditional because both classroom and summative assessments heavily focus on measuring students’ content proficiency rather than their acquisition of knowledge and skills deemed important by the national documents advocating reform in science education (Brickhouse, 2006; Davis et al., in press; Southerland & Hutner, in press). If our assessments, both summative and formative, fail to measure what we want our students to learn, the progress of science education reform is under a serious threat. In order to make sure that our assessment methods reinforce the goals of science education reform documents (AAAS, 1993; NRC, 1996) and ensure students’ continuous improvement (Davis et al., in press), science educators must invest serious effort into understanding the professional and structural dynamics that present challenges to the enactment of assessment reform in science classrooms. However, in order to develop such understanding we need to look at the role of assessment in students’ learning and in the operations of schools from a historical perspective.

**A Historical Perspective on Assessment**

Assessment has played a significant role in efforts to bring about improvements in the educational system in the U.S. throughout the history of the U.S education (Darling-Hammond, 2003; Delandshere, 2002; Herman, 1992; Linn, 2000; NRC, 2001; Popkewitz, 2000; Stiggins, 2004). However, most of these efforts focus on bringing about improvements in the educational system through standardized testing (Education Week,
The strong emphasis on the value of testing is rooted in the political ideals and educational pedagogy that shaped American education for the first half of the 20th century (Brookhart, 2006; Davis et al., in press; Pinar et al., 1996; Popkewitz, 2000; Shepard, 2000; Southerland & Hutner, in press).

**Social Efficiency, Behaviorism and Standardized Testing**

Social efficiency and behaviorist pedagogy played a crucial role in the operation of American schools during the first half of the 20th century (Pinar et al., 1996). The behaviorist theorists like Gagne (1965), Skinner (1938) and Thorndike (1932) interpreted learning through a stimulus-response approach and considered learning as an accumulation of facts that must be mastered sequentially (Linn, 2000; Popkewitz, 2000; Shepard, 2000; Sloane & Kelly, 2003). They argued that a specific stimulus generates a particular response and through quantifying that response one can measure the quality of students’ learning (Pinar et al., 1996; Shepard, 2000). Consistent with these beliefs, behaviorists developed and fostered objective achievement tests to measure students’ learning outcomes (Pinar et al, 1996; Shepard, 2000). Moreover, they convinced the educational community of the merit of objective tests as an accurate measure of learning gains (Education Week, 2004; Pinar et al., 1996). An objective test refers to an examination in which each question is set in such a way that each question can only have one correct answer (NRC, 2001; Shepard, 2000).

The norms established by Thorndike gave school politicians their first chance to compare the knowledge and abilities of their students against an established set of norms (Education Week, 2004; Pinar et al., 1996). For instance, school officials assigned students to academic tracks based on their performance on achievement tests (Pinar et al., 1996). The proponents of social efficiency considered schooling to produce model citizens with perfect knowledge and skills who will reproduce culture (Pinar et al., 1996). They viewed testing as a perfect instrument in tracing how individuals progressed towards that goal (Pinar et al., 1996). Several scholars argue because today’s schools still function based on social efficiency of model of education (Collins, 1998; Marshall, 2006; Sacks, 1999; Southerland & Hutner, in press), testing plays a central role in political efforts to bring about improvements in schools.
Pedagogical Shift on Learning and Assessment

A growing number of scholars started to challenge the behaviorist pedagogy based on epistemological, political and moral reasons around the second half of the 20th century (Cook-Sather, 2002; Driscoll, 1994; Pinar et al., 1996). With the rise of constructivist learning theories, critical pedagogy and feminism, educators started to voice that standard methods of scientific measurement are grounded in assumptions that are unlikely to hold for learning (Cook-Sather, 2002; Driscoll, 1994; Feiman-Nemser, 2003; Fosnot, 1996; Pinar et al., 1996; Shepard, 2000; Tobin, Tippins & Gallard, 1996). In the light of the new insights about how students learn, critics started to argue that the behaviorist pedagogy does not foster students’ creativity, instead it focuses on classroom control, and as a result, structures curriculum in a way that limits science learning to the mastery of basic skills (Collins, 1998; NRC, 2001; Shepard, 2000; Southerland & Hutner, in press). Furthermore, the growing recognition for alternative theories of how student learn (Gardner, 1983; Piaget, 1975; Vygotsky, 1978) leads to a shift in science educators’ thinking about teaching and learning (Cook-Sather, 2002; Wagner, 2002). This paradigm shift gave rise to the constructivist pedagogy.

Although there are several forms of constructivist pedagogy, I limit my discussion to social constructivism. That is because social constructivist pedagogy is consistent with the call for assessment reform in science education (Delandshere, 2002; NRC, 2001; NRC, 2001; Shepard, 2000).

Social Constructivist Pedagogy

Social constructivism suggests that student learning takes place through a social and communicative process, whereby learners share knowledge and construct understandings in a social and historical context (Aldridge, Fraser & Taylor, 2000; Mercer & Jordan, 1994; Tobin & Tippins, 1993; Tobin et al., 1996; Vygotsky, 1978). A pedagogy informed by social constructivism assumes that teachers are prepared for contextual uncertainty and the conflicting nature of teaching with its associated demand for creativity and adaptability (Feiman-Nemser, 2003; Fosnot, 1996; Shepard, 2000; Tobin et al., 1996). It follows that teaching is less a matter of transmitting knowledge, classroom management, and testing and more a matter of engaging and nurturing students’ inner dispositions, and knowledge construction through constant challenge,
collaboration, assessment and feedback (Arter, 2003; Davis et al., in press; Fosnot, 1996; Shepard, 2000).

The foundation of social constructivist pedagogy comes from the work of Vygotsky, the leading psychologist, who challenges the behaviorist assumptions about how students learn and advocates for alternative approaches to learning (Cook-Sather, 2002; Driscoll, 1994; Shepard, 2000). Dewey’s earlier work on inquiry and democratization of schools served as a fertile ground for the widespread acceptance of social constructivism (Oakes & Lipton, 2002). John Dewey, one of the prominent leaders of education in early 20th century, suggested that citizens in a democratic society should be active participants in the construction of society. To maintain the existence of a democratic society, he argues, citizens should be able to ask questions about their physical and social environments (Dewey, 1916). According to Dewey (1916), “the purpose of education should be about providing resources for these engaged citizens to find answers to their questions independent of external authority” (DeBoer, 2002, p. 406). One of the arguments is that standardized curriculum and standardized testing “makes students exclusively dependent on authority and ultimately produces passive and unengaged citizens” who are unlikely to transform society in Dewey’s terms (DeBoer, 2002, p. 406). Several scholars maintain that because the politicians are concerned about creating an efficient society who will economically and technologically outperform other nations in the global competition for leadership, they promote competition rather than creativity in schools (Cook-Sather, 2002; DeBoer, 2002; Marshall, 2006). They further argue that politicians focus on testing partly because testing serves as a perfect instrument of sorting the best skills who will meet the demands of industry (Marshall, 2006).

Although politicians focus on the efficiency of schools, science educators advocate for social transformation of society, thus, focus on fostering students’ intellectual inquiry skills (NRC, 1996; Gallagher, 2006; Southerland & Hutner, in press). These scholars view social constructivist pedagogical practices as effective means for producing citizens who will transform society in Dewey’s (1916) terms. Science education reformists promote social constructivist pedagogical practices partly because they 1) enable students to develop a critical habit of mind, a fundamental goal of science education reform, and 2) provide the challenge and opportunities for students to learn continuously (Driver,
Social Constructivism and Assessment

Social constructivist theorists make the following assumptions about assessment. They argue that assessment should be collaborative, continual, and less formal and embedded in real world tasks (Arter, 2003; Bell & Cowie, 1997; 2001; Davis et al., in press; Klassen, 2006; Marzano, Pickering & McTighe, 1993; NRC, 1996; Pringle, 2000; Shepard, 2000). Social constructivism views assessment as a formative process that encourages students to learn continuously (Davis et al., in press; NAS, 2006; NRC, 2001; Shepard, 2000). Proponents argue that the fundamental strength of a social constructivist assessment format lies in its ability to individualize assessment, to engage teachers more deeply in the assessment process and to provide more rigorous and meaningful feedback (Abell & Volkmann, 2006; Arter, 2003; Black & William, 1998; Crooks, 2001; NAS, 2006; Shepard, 2000; William & Lee, 2001). Furthermore, such assessment practices involve students in reflective activities in which teachers encourage students to consider the strengths and weaknesses of their learning and make plans for subsequent actions (Abell & Volkmann, 2006; Marzano et al., 1993; NRC, 2001; Shepard, 2000; Stiggins, 1997; Wiggins, 1998).

Shepard (2000) points out that assessment has fallen behind instruction in terms of incorporating recent research on student learning. Arter (2003) maintains educators appear to have a naïve understanding of the role of assessment in their instructions. For instance, they appear to be investing a lot of time and energy into “assessment of learning” at the expense of “assessment for learning” (p. 472). Shepard (2000) argues that to be consistent with the social constructivist pedagogy, assessment must be redefined to reflect recent understanding about how students learn. First, teachers need to re-conceptualize their understanding and knowledge of assessment: moving from assessment for record keeping to assessment for improvement in students’ learning (NRC, 2001). The name “formative assessment” describes assessment formats that help students learn continuously (NRC, 2001).

Formative Assessment and Student Learning

Assessment becomes formative when the information is used to inform the course
Bell and Cowie (1997) maintain that formative assessment generates information for two different purposes: to inform instruction and to acknowledge students’ about their strengths and weaknesses. This suggests that teachers must act on the information that they collect through their assessments either to craft their instruction when it is deemed necessary or to provide feedback to the students when they are in need of help to pursue their learning goals. Furthermore, many researchers argue that formative assessment techniques promote teaching to higher standards, motivate students to take more responsibility for their own learning, and provide a more credible form of accountability (Arter, 2003; Black & William, 1998; Boston, 2002; Gipps, 1994; 2000; Shepard, 2000).

Several scholars maintain that students should no longer be seen as the passive recipients of the teachers’ value judgments (Abell & Volkmann, 2006; Bell & Cowie, 2001; Cook, 1998; Davis et al., in press; Shepard, 2000) rather “they are active, engaged, and challenged contributors to their own learning”, when formative assessments are used to support students’ learning (Hargreaves, Earl & Schmidt, 2002, p. 77). According to Shepard (2000), since students are often called upon to make their understandings explicit, formative assessments permit and encourage students to become articulate advocates of their learning (Abell & Volkmann, 2006; Bell & Cowie, 2001; Davis et al., in press; Shepard, 2000). In fact, Black and William (1998) find that students’ learning gains are greater when teachers use formative assessments to assess students’ learning than the traditional assessment formats partly because formative assessments places much of accountability on students by making learning student centered. Furthermore, formative assessments are promising in ascertaining students’ acquisition of knowledge and skills deemed important by national science education reform documents (Abell & Volkmann, 2006; NAS, 2006; NRC, 2001).

In spite of the reported benefits to the students’ learning, most prospective teachers graduate from the teacher education programs without having experienced social constructivist assessment practices (Davis, et al., in press; Shepard, 2000). Davis et al (in press) maintain that prospective teachers need to have such experiences in their own repertoire before they can start teaching assessment practices that are consistent with
social constructivist pedagogy. In addition to teachers’ limited knowledge of and experiences with social constructivist assessment (Shepard, 2000), policy initiatives such as NCLB Act and FCAT present major challenges to the use of social constructivist-based assessments in science classrooms (Southerland & Hutner, in press).

In the following section, I discuss how recent federal and local policy initiatives aiming to improve the state of public school education through standardized testing undermines the promises of social constructivist accounts of assessment from a historical perspective.

National Reports and Standardized Testing

Policy makers’ recent commitment to changing the course of students’ learning through standardized testing dates back to the well-known report A Nation At Risk (National Commission on Excellence in Education, 1983). In the following passages I discuss A Nation at Risk and the implications of this report for science education reform and assessment of students’ learning in science.

A Nation at Risk

Twenty years ago in an open letter to the American citizens, The National Commission on Excellence in Education published a well known report, A Nation at Risk (National Commission on Excellence in Education, 1983) that declared a state of crisis in the educational system evident in the following quote, “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" (National Commission on Excellence in Education, 1983, p. 5). The main argument of the report is the following: the American schools are not producing the educational excellence required to compete in the competitive global market (Brickhouse, 2006; Collins, 1998; Sloane & Kelly, 2003; The National Commission on Excellence in Education, 1983). The strong language of fear that the report used in its findings and recommendations caused a stir among policymakers that provided much of the impetus for the standards-driven curriculum and standardized testing cross the nation (Brickhouse, 2006; Collins, 1998; Grossman, 2003; Settlage & Meadows, 2002; Weiss, 2003; Yager, 2005).

In this publicly well-recognized report, the commission suggested that schools
adopt more “rigorous and measurable standards” and high expectations for students’ academic performance that would push students to the best of their ability to succeed (Collins, 1998; Delandshere, 2002; U.S. Department of Education, 1983; Yager, 2005). The report also makes specific recommendations for bringing quality teachers into public schools as well. For example, it highlights the need for raising the quality of teacher preparation programs and the standards for admission to those programs (The National Commission on Excellence in Education, 1983). It proposes better working conditions for teachers including competitive salaries, an array of professional development opportunities and empowerment in instructional decisions that would affect classroom practice (The National Commission on Excellence in Education, 1983).

This report has been severely criticized by some scholars partly because it uses test results as reference for its findings. Critics argue that test results alone fail to provide sufficient information for making judgments about the state of education in nation’s public schools (Berliner & Biddle, 1995; 2005; Brickhouse, 2006; Darling-Hammond, 2003; Delandshere, 2002; Harris & Herrington, 2006; Sacks, 1999). Drawing on a convincing amount of data on student performance, Guthrie and Springer (2004) challenge this report’s reasoning. They argue that test results that form the basis of A Nation at Risk (The National Commission on Excellence in Education, 1983) cannot be trusted for making judgments about the state of students’ performance as they often give us false readings of what students are able to do. Berliner and Biddle (1995) in their well-known book, The Manufactured Crisis, draw data from variety of sources to convince the educational community that the educational system in America is doing far better than the administration had claimed in A Nation at Risk (U.S Department of Education, 1983).

Several scholars maintain that although A Nation at Risk (U.S Department of Education, 1983) puts forward recommendations for content standards, pedagogy and teacher quality to improve students’ learning in science, politicians have used this report to justify their commitment to bringing about improvements in schools through standardized tests (Berliner, 2006; Brickhouse, 2006; Guthrie & Springer, 2004). These scholars state that although there have been some minor changes in teacher salaries and issues related to the classroom since the publication of this document, content standards and accountability have received most of the attention at the expense of pedagogy and
Many argue that without explicit attention to pedagogy and assessment, rigorous content standards and frequent testing is unlikely to change the state of students’ learning in science (Berliner, 2006; Darling-Hammond, 2003; Harris & Herrington, 2006; Sacks, 1999; Stiggins, 2004). Furthermore, significant evidence exists to suggest that the use of standardized testing as an accountability measure to bring about improvements in the educational system has resulted in dysfunctional outcomes (Darling-Hammond, 2003; Southerland & Hutner, in press; Stiggins, 2004). For instance, NAEP’s (2006) recent report on students’ science achievement trend for the last decade confirms that rigorous content standards and accountability measures have not led to significant increases in students’ learning in science.

In spite of recurring evidence suggesting that standardized tests do not increase students’ learning in science (Berliner, 2006; Delandshere, 2002; NAEP, 2006), the federal and local governments continue to invest efforts and money into formulating standardized tests to address the problems with students’ low academic performance. The recent policy initiative, the law of NCLB (U.S. Department of Education, 2002), is a strong indication of policy makers’ commitment to bring about improvements in education through standardized tests.

The No Child Left Behind Act


The NCLB Act (U.S. Department of Education, 2002) requires all schools to test all children in grades 3-8 every year in mathematics and reading beginning in 2005 and in science by year 2007. These test scores will determine if schools are making Adequate
Yearly Progress (AYP) towards the goal of proficiency for all children by the 2013-2014 deadline. Adequate yearly progress is the minimum level of improvement school districts and schools must achieve every year to meet this goal (U.S. Department of Education, 2002). As a result of the NCLB Act (U.S. Department of Education, 2002), states rank schools based on students’ test results in reading, mathematics and science and make funding decisions based on students’ tests scores. Through this system, the state rewards schools whose children score at or above achievement expectations and schools whose children fail to meet the achievement expectations will lose funding (Berliner, 2006; Brickhouse, 2006; Southerland & Hutner, in press; Stigins, 2004). However, many science educators criticize the NCLB Act for intending to bring about improvements to students’ learning by threatening schools (Brickhouse, 2006; Stiggins, 2004).

Critics of the No Child Left Behind Act

Although there is a heightened interest in bringing about improvements to the educational system through standardized testing, many find this approach problematic. Brickhouse (2006) maintains, “Tests are only a proxy for those competencies that we really value. Some tests are certainly better proxies than others. But none of them are the actual competences that are the aim of scientific literacy” (p. 3). Brickhouse’s (2006) argument presents serious challenges to the push for the quantification of students’ learning in science through standardized tests. The limitations of standardized tests are not limited to the construct validity or service to the wrong competencies that Brickhouse’s discussion highlights. The limitations of testing extend to include failure to serve the purposes for which they are designed (Davis et al., in press; Darling-Hammond, 2003). For instance, addressing the achievement gap between different sects of society is one of the core arguments of the NCLB Act (U.S. Department of Education, 2002). The NCLB Act aims to eliminate inequities by providing the same educational opportunities to the students regardless of their racial or ethnic backgrounds, socioeconomic status, disability and English language proficiency (U.S. Department of Education, 2002). Statistics show, however, standardized tests hurt students coming from deprived racial and socioeconomic backgrounds in many ways (Berliner, 2006). For instance, evidence suggests that the drop out rate for minority students is increasing because of standardized tests (Darling-Hammond, 2003; Moore, 2004; Popham, 2006; Stiggins, 2002).
Several science educators maintain NCLB reduces the role of assessment merely to accountability (Abell & Volkmann, 2006; Brickhouse, 2006; Southerland & Hutner, in press). Critics also argue that limiting the role of assessment in science education reform to the accountability does not hold potential to ascertain students’ acquisition of knowledge and skills deemed important by the NSES (Southerland & Hutner, in press) as it reduces the science learning to students’ acquisition of canonical knowledge (Brickhouse, 2006). According to Davis (personal communication, September 2, 2006) standardized tests fail to measure the depth of knowledge and the range of skills that the NSES (NRC, 1996) promotes in part because while we have developed much expertise over the last 50 years in measuring students’ acquisition of factual knowledge we have not come as far in developing techniques for measuring students’ thinking and their scientific inquiry processes. If the standardized tests fail to serve the purposes for which they are designed and fail to ensure students’ acquisition of knowledge and skills deemed important by the science education reform, alternative assessment perspectives must be brought fore to reinforce science literacy goals outlined in reform documents (AAAS, 1993; NRC, 1996, 2001). However, in spite of the reported limitations of standardized testing and the proven effectiveness of formative assessment for students’ learning, many states have responded to the federal call for greater accountability through standardized testing. Florida is no exception in this nationwide trend. I discuss how the political efforts to improve public schools’ performance in the state of Florida conceptualize assessment in the following passages.

Florida Comprehensive Assessment Test (FCAT)

As it was set by The Comprehensive A+ Plan (Florida Department of Education, 1996) and inspired by the law of NCLB (U.S. Department of Education, 2002), the state of Florida developed and disseminated a central measurement system, Florida Comprehensive Assessment Test (FCAT) to monitor the overall performance of public schools, student learning and teacher performance. The state of Florida designed FCAT to ensure a high quality educational system that will engage all students in rigorous standards of learning (Florida Department of Education, 2004; Gandal & McGiffert, 2003). The assumption guiding this policy is that standardized tests are valid constructs that can reliably measure students’ academic performance (Berliner, 2006; Brickhouse,
2006; Delandshere, 2002; Herman, 1992), thus, easily produce meaningful information that policymakers and teachers can later act on to make informed decisions about the future of students’ education (Florida Department of Education, 2004). For instance, high school graduation is contingent upon students’ passing grades on the FCAT mathematics and reading tests. According to this policy, schools are graded on an A-to-F scale based on how well their students perform on the FCAT. All the students of any school that receives an F twice in any four-year period become eligible for school vouchers, which may be used at other public schools or private schools (Florida Department of Education, 2004). Many argue that FCAT uses a language of fear that is consistent with the language of A Nation at Risk (National Commission on Excellence in Education, 1983) to bring about improvements to the educational system (Popham, 2006). Scholars maintain that the emphasis on FCAT forces teachers to alter their lesson plans and pressures them to spend more time preparing students for the test rather than fostering students’ acquisition of scientific inquiry skills (Blanchard, 2006; Yalaki, 2004). Davis maintains that this fear based notion of improvement contests the goals of science education reform (personal communication, September, 10, 2006). Furthermore, she maintains that the fear of receiving a failing grade on the FCAT pressures schools to encourage pedagogical and assessment practices that fail to ensure students’ acquisition of science literacy goals outlined in reform documents (Davis, personal communication, September, 10, 2006).

Although the state of Florida aims to produce maximum progress in students’ learning by measuring students’ learning through standardized testing, standardized testing fuels an intense debate among educational researchers, teachers, policy makers and parents at the national and local level. I introduce the nature of this debate in the following passages.

Equity and Access to Education For All

The national commitment to equity and justice in education is evident across policy and reform initiatives, which aim to bring about improvements to the educational system (AAAS, 1993; NAS, 2006; NRC, 1996; U.S. Department of Education, 2002). States, for instance, are currently busy designing standards and accountability systems that presumably ensure learning for all students (Brickhouse, 2006; Collins, 1998; Madaus, 1991). However, critics highlight the inequitable, discriminating nature of
accountability systems in the form of standardized testing (Berliner, 2006; Darling-Hammond, 2003; Marshall, 2006; Sacks, 1999; Stiggins, 2004). For instance, several scholars maintain that standardized tests undermine reform efforts aiming to create supportive learning environments for disadvantaged student populations. Furthermore, while reform documents (AAAS, 1993; NRC, 1996) address the concept of educational equity by focusing on the distribution of resources between racially, ethnically, gender wise, and socio-economically different groups within society and encourage teachers for the use of nurturing pedagogical and assessment practices, accountability policies address the concept of equity by raising standards of achievement for all and expecting all students to live up to those standards based on the assumption that all students have the same prior knowledge, access to the same resources and educational support (Darling-Hammond, 2003; Stiggins, 2004). However, this is far from reality (Berliner, 2006; Popham, 2006).

These scholars maintain that in spite of the federal and local intentions to eliminate inequities that standardized tests may potentially produce, these tests continue to perpetuate and endorse the already existing inequities within society partly because these tests are formulated based on philosophically and politically flawed assumptions (Berliner, 2006; Darling-Hammond, 2003; Popham, 1995; Sacks, 1999; Solano-Flores et al, 2001). For instance, instructional practices that focus on test preparation and, frequent use of testing as a means for assessment of student learning function to disqualify some students from access to participation and, therefore, to knowledge acquisition (Sacks, 1999; Shepard, 2000). Such instances produce severe results for minority and low-income students (Berliner, 2006; Carlone, 2003; Haberman, 1991; Sacks, 1999; Stiggins, 2004). Tests involving high stakes decisions then become perfect tools for perpetuating and reproducing the status quo that ensures discrimination. For instance, in spite of massive monetary investment in test preparation, racial and ethnic minorities continue to score lower than their classmates in reading, science and mathematics across grade levels. In a specific case in Pinellas County, Florida a group of independent consultants hired by a group of parents concluded that FCAT scores are as much as 28 points lower for black students than whites, and that black students are not promoted to the next grade at the same rate as white students (St. Petersburg, 2003). More than two-thirds of Florida's
Black students cannot read at their grade level, according to *FCAT* results, while only about one-third of white students performed so poorly. This achievement gap is intensified in mathematics test scores. Black students are 40% points behind white students in Pinellas and 35 points behind statewide (St. Petersburg Times, 2004).

According to 2004 data (Florida School Advisory Council, 2004) students in Gadsden County with 83% African American student population performed below state average both in math (31% compared with 57% state average) and reading (27% compared with 51% state average). These data show that minority students are not performing as well as their white peers on the *FCAT*. As a result of low-performance on standardized tests, schools in minority and low-income neighbourhoods receive less money to support students’ learning than the schools that already perform well on the tests. Because of the punishment and reward metaphor that the FCAT administration uses to bring about improvement to the educational system, many of Florida’s schools in poor neighborhoods face significant challenges for providing equitable educational opportunities for all students (Moore, 2004; National Center for Fair and Open Testing, 2005).

*The Influence of Standardized Testing on Teaching*

A large body of research addresses the effects that the standardized tests have on science teaching (Darling-Hammond, 2003; Popham, 1995; Stiggins, 2004; Madaus, 1991; Westerlund & West, 2001). Settlage and Meadows (2002) examine the effects of standards-based reform on science teaching at urban school settings. Their study reveals that standards-driven reform influences teachers in four negative ways: “undermining teachers’ professionalism, eroding teacher-student relationship, diluting the science curriculum and instruction based on predicted individual test performance” (p. 1). Settlage and Meadows (2002) argue that “educational reformists view standards as a mechanism for spurring greater accountability; in effect the opposite is taking place” (p. 117). These measures weaken not only teacher’s ability to become reflective professionals who are “capable of adjusting to the conditions of their work environment” but also their ability to help students engage, in their learning in innovative ways (Settlage & Meadows, 2002, p. 117). The testing accompanying these standards they argue, “is the destruction of the very professionalism that teacher educators hope to

Several researchers report that not only do accountability pressures motivate teachers to focus their instructional planning on test content (Whitford & Jones, 2000), they also encourage teachers to devote more instructional time to preparing students to do well on tests (Grant, 2000; Hamilton & Stecher, 2002; Mehrens & Kaminski, 1989, Shepard, 1991; Smith & Rottenberg, 1991). It follows that perceived improvements in students' achievement maybe a direct result of time devoted to test preparation (Amrein & Berliner, 2003).

Corbett and Wilson (1991) examine the influence of state-mandated testing on teacher practice in Maryland and Pennsylvania. They document that while state-mandated testing inspires some changes that take place in classrooms it is not clear whether these changes would be considered as reform because they do not bring about improvements in students’ learning. Zancanella (1992) and Cimbricz (2002) argue that the assumption that state-mandated testing leads to better teaching and triggers significant growth in students’ learning expresses more hope than reality. Overall, they suggest that despite the firm belief that standardized testing contributes by large extent to educational improvement for all, “evidence to support this claim yet to be established” (Cimbricz, 2002, p. 2).

Much of research supporting the potential power of assessment for learning, therefore, criticize traditional model of testing and concludes its adverse effects on teaching and learning (Berliner, 2006; Brickhouse, 2006; Sacks, 1999; Settlage & Meadows, 2002). These arguments, thus, a) suggest that standardized testing dilutes the purposes of science education reform (Brickhouse, 2006; Collins, 1998; Southerland & Hutner, in press) and b) indicates the shortcomings of standardized tests as means for achieving meaningful gains in students’ learning (Arter, 2003; Shepard, 2000).

In the following passages, I discuss why standardized testing fails to bring about improvements to the learning of students.

Concluding Remarks

What research has revealed up to this point is the systemic failure of efforts to
reform science education through external measures such as standardized testing. Furthermore, some scholars maintain that even reform efforts led by science educators are not being implemented comprehensively (Brickhouse, 2006; Southerland & Hutner, in press), partly because limited attention has been paid to teachers’ internally held conceptions about reform ideas (Carlone, 2003; Southerland & Hutner, in press). It follows that in order to understand the challenges associated with reforming the culture of assessment in science classrooms, we need to understand the ways in which teachers conceptualize assessment.

Assessment is important in educational reform because it shows how students progress towards acquiring the knowledge and skills deemed important in science education reform documents (Brickhouse, 2006; Gallagher, 2006). However, in order for assessment to indicate how students progress towards achieving science literacy goals outlined in reform documents, assessments must be designed to measure students’ learning across those goals (AAAS, 2005; NAS, 2006; NRC, 2005).

Hargreaves et al (2002) report that most teachers do not have sufficient knowledge of the most current learning theories and the associated assessment practices that these theories demand. Furthermore, the increasing focus on system efficiency through standardized testing undermines teachers’ professional capacity to use assessment formats to help students learn science continuously (Abell & Volkmann, 2006; NAS, 2006).

As Shepard (2000) suggests, “to accomplish the kind of transformation envisioned, along with the social definition, format and content of assessment teachers’ both beliefs and practices must confront traditional purposes and definitions of assessment” (p. 10). The challenge of assessment reform from this perspective is not devising valid, reliable and fair measures of assessment that will capture the complexities of students’ learning (Arter, 2003; Hargreaves et al., 2002; Torrance, 1995) or the power of externally imposed policies. Rather, it expands to include the difficulty associated with motivating teachers to understand the purpose of assessment reform, gain necessary knowledge and demonstrate motivation to enact curriculum accordingly (Hargreaves et al., 2002). In order to understand why teachers resist or hesitate to implement assessment reform in science classrooms we need to gain access to their internally held conceptions.
of assessment.

**Understanding Teachers’ Beliefs on Assessment**

Several scholars define beliefs as psychological constructs that can guide a person’s judgments and evaluation of knowledge and events, their decision making process and their actions (Nespor, 1987; Pajares, 1992; Richardson, 1996). According to Pajares (1992), “beliefs are created through a process of enculturation and social construction and with time and use, people’s beliefs become robust and difficult to change” (p. 316). However, alternative explanations exist and suggest that people’s beliefs can change in presence of more fruitful and plausible alternatives to their existing beliefs (Southerland, Johnston & Sewell, 2006). Although, whether a person’s beliefs change or not or under what conditions a person’s beliefs change is still subject to debate, there is no doubt that a teacher’s beliefs play a significant role in what a teacher does in the classroom in terms of teaching and assessment of students’ learning (Bruning, Schraw, Norby & Ronning, 2004; Pajares, 1992; Richardson, 1996; Tobin et al., 1996; Yalaki, 2004).

Like other people, teachers also hold beliefs about various things. For instance, teachers hold beliefs about the content they teach, beliefs about their students, beliefs about teaching and learning and beliefs about assessment (Brickhouse, 1990; Brunning et al., 2004; Pajares, 1992; Richardson, 1996; Thompson, 1984, 1992). Many educators, who have been involved in understanding teachers’ beliefs, maintain that teachers plan instruction in ways that are consistent with their beliefs about knowledge, teaching and learning (Bybee, 1993, Kagan, 1992, Lederman, 1992; Nespor, 1987, Pajares, 1992; Richardson, 1996; Southerland, Sinatra & Matthews, 2001).

For instance, literature on science teachers’ beliefs about nature of science reveals that teachers who view science as a concrete body of knowledge or an object for students to obtain are likely to focus on didactic instruction and students’ acquisition of prescribed knowledge in their teaching (Abd-El-Khalik & Akerson, 2004; Bencze, Bowen & Alsop, 2004; Tsai, 2002; Yalaki, 2004). The same line of research reveals that teachers, who view science as a way of knowing are likely to teach science for fostering a scientific habit of mind and focus on students’ acquisition of inquiry and critical judgment skills rather than their acquisition of prescribed knowledge (Abd-El-Khalik & Akerson, 2006;
Bencze, Bowen & Alsop, 2006; Kang & Wallace, 2005; Tsai, 2002).

The science education literature reveals that many teachers hold traditional beliefs about science, pedagogy and assessment (Gallagher, 2006; Genc, 2006; Southerland & Hutner, in press, Yalaki, 2004). Tobin et al (1996) maintain that teachers continue to hold traditional beliefs about teaching and learning partly because the teacher education programs do not challenge preservice teachers to view their beliefs about teaching and learning as problematic constructs that need to be rebuilt as they engage in teaching. Davis (1996) finds that teachers who view their beliefs about teaching and learning as problematic constructs are likely to reconstruct their beliefs about teaching and learning based on their experiences and teach science in a responsive manner to address students’ learning needs.

Several scholars maintain teachers’ beliefs about science, pedagogy and assessment play a critical role in the success of science education reform objectives (Bybee, 1993; Cuban, 1990; Davis, 1996; Gallagher, 2006; Gess-Newsome et al, 2003; Southerland, Sewell & Blanchard, 2006; Yerrick & Parke, 1997).

Science educators have shown interest in understanding teachers’ pedagogical beliefs; beliefs about teaching and learning (Haney et al., 2002; Bryan, 2003; Haney & McArthur, 2002; Levitt, 2002; Tsai, 2002), teachers’ epistemic views of science; beliefs about the nature of science (Abd-El-Khalik & Akerson, 2006; Cobern, 2000; Lederman, 1992; Tsai, 2002), in the last three decades. However, they have shown scarce interest in studying teachers’ beliefs about assessment. This lack of interest in studying teachers’ assessment beliefs indicates the need for understanding teachers’ conceptions of assessment. Understanding teachers’ conceptions of assessment is particularly important in an era in which teachers receive different messages about the role of assessment in students’ learning. For instance, while recent accountability measures call teachers to reduce the role of assessment in students’ learning merely to the accountability purposes, science education reform documents (NAS, 2006; NRC, 2001; 2005) ask teachers to use assessment for identifying students’ weaknesses and strengths and craft instruction based on students’ needs and strengths. Furthermore, while the NCLB Act and FCAT encourage teachers to use traditional assessment methods such as multiple choice tests, science education reform documents (NRC, 2001) encourage teachers to use alternative forms of
assessments such as portfolios, journals and essays.

These two calls, assessment for accountability and assessment for continuous improvement, are likely to present challenges for teachers in terms of making assessment decisions. Therefore, it is important to understand how science teachers internalize the two different perspectives on assessment as they design their lesson plans, scaffold instruction and assess students’ learning in science. However, looking at teachers’ understanding of assessment in isolation from their epistemic and pedagogical conceptions would limit the implications of this study for further research, for curriculum developers and for practicing community. To overcome this limitation, I examine teachers’ conceptions of assessment in light of their epistemic and pedagogical conceptions. Such approach to understanding teachers’ beliefs and practices of assessment is likely to produce useful knowledge and understanding that science educators can use to assist science teachers to change the culture of assessment in science classrooms.
CHAPTER THREE

METHODOLOGY

In this chapter, I provide a description of the research paradigm, the methodology guiding my inquiry and the methods I used to collect my data. I begin with a brief discussion of the nature of interpretive research, the methodological approach guiding my inquiry. Then, I introduce my research context, the process guiding the selection of my participants and the data collection methods. The last section of this chapter focuses on the data analysis methods and responds to the criteria of validity, robustness, and trustworthiness.

Denzin and Lincoln (2000) argue that the nature of the research questions should be the most significant factor on the choice of a research methodology. More precisely, they argue that what one wants to learn should determine how one should go about learning it. This purpose of this study is to understand the ways in which science teachers’ epistemic views of science and their pedagogical conceptions influence their conceptions and practices of assessment under the cultural and political influences of schooling. As a researcher involved in qualitative inquiry, I did not just want to present descriptions of what happened in science classrooms in terms of assessment, rather I was interested in exploring and characterizing the various reasons that generated those practices. Therefore, I found interpretive methodologies instrumental in helping me answer my research questions.

Theoretical Foundations

Denzin and Lincoln (2000) maintain that one cannot conduct qualitative research without making explicit or implicit references to an underlying theoretical framework. Interpretive methodologies and symbolic interactionism informed my data collection and data analysis process.

Interpretive Research

Several researchers (Biklen & Bogdan, 1992; Denzin & Lincoln, 2000; Janesick, 2000; Wilson, 1998) define interpretive research as an attempt to understand the world of
subjects’ experiences through their perspectives. An interpretive study focuses on the human meaning making aspect of teaching, seeing teacher practice as a product of interpretations, interventions and individual decisions (Crotty, 1998). Thus, interpretive researchers attempt to understand social phenomena through accessing the meanings that participants make of their experiences. Denzin and Lincoln (2000) maintain that interpretive researchers reject the objective accounts of knowledge, seeking instead a subjective, negotiated understanding of phenomena under investigation.

Marshall and Rossman (1995) recommends the use of interpretive research for the following contexts.

1. Research that attempts to develop in-depth understanding of the complexities and processes relevant to a specific context.
2. Research that attempts to explore why and how policy and local knowledge and practices are related to one another

These concerns were relevant to the purpose of my study; therefore, I utilized an interpretive mode of inquiry. In this study, I was interested in characterizing the challenges associated with implementation of assessment reform in science classrooms. Because teachers’ assessment decisions interact with different factors at professional, cultural and political levels, an interpretive approach allowed me to better understand the complexity of enacting assessment reform in science classrooms.

In the following passages, I explain how the interpretive research serves the purposes of my study.

**Interpretive Inquiry and the Focus of My Study**

Prominent science education reform documents (AAAS, 1993; NRC, 1996) recommend teaching of inquiry skills through specific pedagogy, and of assessing what students learn through alternative forms of assessments. On the other hand, the political structures such as standardized tests (e.g., *NCLB* and *FCAT*) impose certain roles on teachers, which contradict the goals of science education reform (Brickhouse, 2006). This produces a series of contradictions (pertinent to assessment) for teachers to resolve. How teachers go about resolving these contradictions depend on teachers’ personal capacity (e.g., pedagogical content knowledge). Therefore, what takes place in classrooms is a
result of how teachers’ interpret and attach meaning to these recommendations (Cimbricz, 2002). From this perspective, teachers are not merely the responders of the political norms and passive receivers of scholarly recommendations but instead they are intelligent interpreters of these norms and recommendations. These interpretations and the meanings derived from these interpretations guide teachers’ assessment practices. Both personal and structural forces influence how teachers’ interpret, acquire necessary knowledge to use (if they decide to) and implement calls for frequent use of alternative forms of assessment (NRC, 2001). Because this call for frequent use of alternative forms of assessment is part of science education reform agenda (NRC, 1996; 2001), understanding challenges to the successful integration of these forms of assessment in science classrooms is essential.

Because of the complexity involved in exploring, characterizing and understanding the challenges associated with the successful integration of alternative forms of assessment in science classrooms, the methodology used should be powerful enough to allow contradictions, coherences and power relationships to come to the fore. This method should allow the researcher to observe and interpret the meanings and actions of teachers as closely as it is possible to reflect “the insider’s perspective” (Crotty, 1998, p. 76). Furthermore, such a method should be able to generate information on how teacher agency interacts with the structures (cultural and political) of schooling as they perform their duties. Interpretive methodologies allow for these meanings, coherences and contradictions to be revealed. Interpretive methodologies are powerful in terms of generating knowledge and insights necessary for science educators to understand the complexity involved in enacting reform-based ideas of assessment in science classrooms.

Interpretive methodologies allow us to systematically examine the interactions between the various attributes of teachers’ conceptual ecology, and cultural and political attributes of schooling with a particular focus on the achievement of the enactment of reform-based ideas of assessment. Teachers’ conceptual ecologies can be defined as constructs that involve a range of epistemic and pedagogical attributes (Southerland et al., 2006; Tsai, 2002). The epistemic and pedagogical attributes of these ecologies can be considered internal constructs. Cultural and political attributes of the teaching profession
are external forces that influence the enactment of reform ideas either by exerting pressures on or by providing resources and support for the teachers’ pedagogical and assessment decisions. It is my argument that the relationship between internal and external forces that influence the change process need to be systematically and fully explored. Interpretive methodologies enable us to explore the structural attributes and professional contradictions that limit the enactment of assessment reform in science classrooms. Although researchers interact with the context and participants in different forms and varying degrees, interpretive research generally takes place in natural settings (Biklen & Bogdan, 1992; Denzin & Lincoln, 2000; Lincoln & Guba, 1985; Marshall & Rossman, 1995).

This study took place in a public high school as three science teachers conducted their professional responsibilities without an interruption from the researcher. Although I stayed in the context of my study working as a science teacher, my prolonged engagement allowed me to explore how and why things happened in the way they did in their natural conditions. Although my presence and my questioning might have had an effect on my participants’ instructional and assessment practices, and on how they conceptualized their practices, I did not have political power to have a direct influence on how teachers designed their lessons, how they went about teaching and assessing what they taught. However, I encouraged them to be as honest and open with me as it is possible.

Symbolic Interactionism

I use symbolic interactionism (Blumer, 1969; Goffman, 1958), a specific form of interpretive research, as a methodological approach for investigating the challenges associated with the enactment of assessment reform in science classrooms. The purpose of this research is to shed light on how science teachers interpret political and cultural norms surrounding their assessment practices by using critical attributes of their professional knowledge. This methodological stance proves to be consistent with the focus of my inquiry for the following reasons.

First, the interactionist camp views humans as engaging and creative participants who construct their social world, rather than passive, conforming objects of socialization.
(Blumer, 1969; Crotty, 1998; Goffman, 1958). Secondly, symbolic interactionism directs
the focus of social inquiry on the meanings and actions of acting persons (Blumer, 1969;
Coser, 1971; Crotty, 1998; Denzin, 1978; Goffman, 1958). Finally, it emphasizes
subjective meanings of human behavior as the participants in the study function within
complex social structures. More importantly this perspective stresses that these subjective
meanings are socially constructed, and that these subjective meanings interrelate with
objective actions and cultural and political structures (Crotty, 1998). Thus, symbolic
interactionism proves to be a strong foundation for looking at the interactions between
political and cultural norms, and teachers who interpret and develop responses to those
norms on a daily basis.

Research Questions

I adopted the following table from Schram (2002) to indicate my general
questions, the purpose and the rationale for choosing specific methods of data collection.

Table 2. Data Collection Matrix

<table>
<thead>
<tr>
<th>Main Question:</th>
<th>What are the perceived challenges associated with changing the culture of assessment practices in science classrooms at a public high school?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subquestions:</td>
<td>How do science teachers’ professional knowledge and understanding of assessment influence enactment of assessment reform?</td>
</tr>
<tr>
<td>Question 1:</td>
<td>Purpose of inquiry is to find out how teachers understand assessment and how they assume responsibility for assessment reform</td>
</tr>
<tr>
<td>Method:</td>
<td>Interviews, document analysis and observations</td>
</tr>
<tr>
<td>Rationale:</td>
<td>To achieve a better understanding of what science teachers know about assessment, and the promises and challenges that what teachers know presents for the enactment of assessment reform, it is important to take a closer look at the knowledge base of science teachers (Gess-Newsome &amp; Lederman, 1999). This enables teacher educators to understand where science teachers are, what their professional needs are, and how they can be assisted to attain their professional responsibilities for reinforcing the goals of science education reform.</td>
</tr>
</tbody>
</table>
Table 2- continued

**Question 2:** How does school culture influence science teachers’ assessment practices?

*Purpose of inquiry* is to understand how school culture influences what meaning is made of assessment, what practices are honored and rejected.

*Methods:* Interviews, field notes, reflection journal and artifacts

*Rationale:* School cultures can afford certain assessment practices to prevail and others diminish. Such information is not readily available through observations. Interview data and field notes may allow for better understandings to emerge. Interviews are useful in discerning interior meanings.

**Question 3:** How do policies and political structures and influences affect science teachers’ assessment practices?

*Purpose of inquiry* is to find out the extent of influence recent policies dealing with assessment have had on assessment practices in science classrooms.

*Methods:* Interviews, Examination of documents, Announcements, Observations

*Rationale:* While interviews can allow us to learn how teachers develop responses to FCAT, observations can afford the details of how such understandings are transferred to classrooms as assessment practices. Such information is likely to produce information that will allow us understand how teachers are situated within the larger context of schooling.

Adopted from (Schram, 2002).

I provide information about the context of this study, introduce my participants, and describe the specific data collection methods that I employed while conducting this research in the next sections.

**Design of the Study**

After introducing the methodology that guided my study, I discuss the design of my study in the following sections.

**Research Settings**

This study took place in a public high school, which I call Southern High School. The Southern High School is in the suburb of a metropolitan city in the State of Florida.
According to 2003-2004 school year records, the year I started my data collection, the average years of experience for teachers at this school is 14.1. Of teachers who work at this school currently 44.5% hold an advance degree. Reports (Florida Department of Education, 2003) show that average science classes hold 24.1 students at this school compared with 26.5 in the state. Dropout rate is 0.6% compared with 3.1% state average dropout rate. About 31 percent of students receive free or reduced lunch. According to 2004 records there were 2715 students enrolled at this school and that number exceeded 3200 at the beginning of 2005-2006 school year. This school hosts a very transient student population.

The state of Florida recently adopted an accountability system in the form of standardized testing to monitor the progress of the school systems for improving student learning. Due to the high stakes associated with schools’ and students’ performance on this standardized test (FCAT), there is a heightened interest amongst parents, teachers, students, administrators and politicians about the implications of this accountability system for teaching, learning and assessment of student learning.

Selection of Participants

The goals of quantitative and qualitative research are quite different; therefore, researchers in each research genre follow different procedures for selecting their participants (Biklen & Bogdan, 1992; Creswell, 2003; Denzin & Lincoln, 2000; Guba & Lincoln, 1987; Patton, 2002). In qualitative research, samples are not randomly selected from a representative population; instead the researcher carefully and purposefully selects participants who can serve as rich cases of information (Byrne, 2001; Creswell, 2003; Coyne, 1997; Patton, 2002).

In order to satisfy this condition, I used purposeful sampling (Patton, 2002) in selecting my participants. Patton (2002) states that purposeful sampling involves selecting research participants based on their capacity to reveal the amount of data and insight that researchers hope to explore. Purposefully selecting my participants helped me access rich cases of information pertinent to the focus of my inquiry.

I chose to work with public school science teachers. First, science education reform documents (AAAS, 1993; NRC, 1996) target public schools. Thus, public school
Science teachers are the chief consumers of the recommendations (pedagogical and assessment) put forward by these documents. Understanding how teachers respond to these recommendations becomes critical in our efforts to improve student learning in science. Secondly, public schools appear to go under intense pressure due to standardizing testing cross the nation (Brickhouse, 2006; Collins, 1998; Stiggins, 2004), therefore, it makes sense to work with teachers who experience professional, cultural and political tensions on their pedagogical and assessment practices on a daily basis. Exploring these tensions and finding ways to resolve them become essential for liberating change in what we teach, how we teach and how we assess what we teach in science classrooms.

I worked with three science teachers with varying years of teaching experience. Two of the participants in this study are male teachers between the age of 20 and 30. One of the male participants, Kevin, had taught science for seven years. Ryan, the other male science teacher, had taught science for three years. Melody was the only female participant. Melody had taught science for three years. The school administration chose Melody as the teacher of the year for two years in a row.

All teachers who participated in this study, demonstrated proficiency in the content knowledge that they taught. Colleagues in science department and school’s administration recognized these teachers as competent teachers of science. By competent I mean teachers who had sufficient knowledge of the subject that they taught and who were able to exert authority to keep students on task. In addition to content knowledge, the administration measured teacher competency through 1) teachers’ ability to maintain control in their classrooms and teach from “Bell to Bell”. 2) how well teachers responded to the administrator’s priorities (e.g., focus on reading comprehension skills, limiting hall passes to one per week among other indicators).

However, my observations of these teachers’ perceived competencies did not solely guide my participant selection process. I selected to work with these teachers because these teachers gave me the impression that they were very serious about their profession. Too, these were the teachers with whom I could relate and interact frequently. Finally, these were the teachers whom I thought would be willing to make the time commitment involved with their participation in my study.
My Involvement in the Context

Heshusius (1992) highlights the notion of participation as a meaning-making instrument for social researchers. She argues that participation implies “relatedness, equality, care and full somatic presence” and must be an integral part of the research endeavor (Heshusius, 1992, p. 10). In her definition of participation, Heshusius (1992) notes “participation is about active identification with the phenomena one wants to understand” (p. 18). Although my active participation in the context of this study brought my biases in my findings it gave power to my arguments.

I worked as a chemistry teacher in the same school in the context of this study, therefore, my involvement in this study brought my biases not only into how I collected my data but also in how I interpreted what I observed, what I heard and what I discussed in my findings. Although my primary data collection methods involved participant observations, semi-structured interviews and document analysis, I was constantly engaged in data collection informally as I interacted with my students, listened to the “all call” announcements, set at lunch table with my colleagues, attended workshops and faculty meetings. I interacted with other science teachers who taught chemistry; I exchanged ideas, lesson plans and worksheets as I was trying to find my place in this complex social system called school.

I had just started to work as a science teacher in the context of my study when I started to collect my data. As a first year science teacher coming from a different country trying to teach science in an American public school was a quite experience for me. My cultural background and my limited experience in teaching filtered how I made sense of my experiences and how I observed and interpreted the experiences and meanings of those teachers that I studied. More precisely, my struggle to address the administrative duties, with motivating students to stay on task, and to complete homework assignments influenced what I prioritized in my data collection and how I present my findings. My own struggle to maintain control and motivate students for learning, my hardship in attending to the administrative demands of the school, my fight with catching up with the deadlines in terms of my grading and preparing for my next classes led me to give heightened attention to the those attributes of the teaching profession that oppressed the teachers. I focused on those attributes that reduced individual teacher agency to
performing the political demands of the school instead of empowering the teacher to lead reform efforts. For instance, my awareness of the administration’s focus on classroom management in keeping students on task led me to concentrate on the power relationships between the teacher and the students.

In addition, I entered in the context of my study with extensive philosophical and pedagogical coursework in my educational background. These experiences contributed to the biases embedded in my study as they gave me skills to observe and make meaning of the events and interactions through a critical eye. For instance, I was able to attach meaning to the mundane practices of schooling such as the bell schedule and how it presented challenges to the “doing of inquiry” in science classrooms. Furthermore, my learning about pedagogy and philosophy allowed me to draw conclusions from my own experiences as a teacher and transfer that knowledge into how I made meaning of my participants’ instructional and assessment decisions and practices. My experiences in my graduate program provided me with skills to criticize and try to understand why certain practices that my participants believed helped their students to acquire skills of inquiry, in actuality did not help their students acquire those skills. Through this critical engagement, my ability to interpret my participants’ beliefs and practices while practicing under similar conditions gave strength to how I went about interpreting my findings.

However, my own experiences with the teaching profession as a first year novice science teacher limited my ability to visit and observe my participants’ classrooms as often as I would have liked. As I was trying to fit into the culture of Southern High School and as I was working with my participants, my association with the group influenced how I went about critiquing my participants’ beliefs and practices of assessment. I had a tendency to be more sympatheitc and less evaluative in my judgements. Moreover, partly because of my own experiences within the school, my thinking focused almost exclusively on the question of how students should learn and not on the equally important question of what students should be learning, as I tried to make sense of my participants’ conceptions and practices of assessment. I believe that these factors limited my findings.
Ethical Considerations

Gaining access to the participants is a political act that involves several stages (Denzin & Lincoln, 2000). First of all, I filed a human subjects consent form to initiate my study. As part of this process, I developed an informed consent form to inform my participants about the overall purpose of my research and about the methodologies that I would use for data collection. Although my participants were not likely to experience any particular harm or benefits through their involvement in this study, I mentioned about the risks and benefits that might have unexpectedly been involved in this study. In the consent form I outlined the procedures that I would use for ensuring the confidentiality of private information that they shared with me.

Data Collection Methods

This study involved three fundamental data collection methods; interviews, classroom observations and document analysis. I explain the contribution of each data collection method in my quest to find answers to my research questions in the following sections.

Classroom Observations

Understanding how the political and cultural influences shaped my participants’ conceptions and practices of assessment is one of the central components of my inquiry. Classrooms serve as special contexts for understanding how the cultural and political influences surface. For instance, classroom context enables the researcher to observe the nature of the power relationships between the teacher and students.

Cimbricz (2002) points to the lack of observational evidence that provides information on actual classroom practice due to the influence of state-mandated tests. She argues that while research reports some changes in instruction through looking at teachers’ beliefs, little is known about the nature of change that takes place at classroom level. Although in-depth interviews provide access to how teachers internalize state-mandated testing, coupled with observations, interviews “not only provide data on teachers’ understanding of what and how to teach, but also on how these understandings are operationalized and carried out”… thus [added], “contextualizing teachers’ beliefs
and practices hold considerable promise for future research” and policy implementation (Cimbricz, 2002, p. 4).

In order for me to gain access my participants’ instructional and assessment practices, I visited each teacher three times, observed how they conducted themselves as science teachers and how they interacted with their students. I recorded what happened in the class in my field notes. Those observations gave me access to more than one form of the reality about my participant’s teaching style and assessment practices. Furthermore, these classroom visits enabled me to ask more critical questions regarding their beliefs and practices relevant to assessment.

*Interviews*

Although classroom observations allow for collecting information from a widerange of sources, it is not as powerful as interviewing in terms of revealing information about subject’s thoughts, intentions and feelings about the phenomena being observed (Denzin & Lincoln, 2000).

Bogdan and Biklen (2003) identify three forms of interviews: structured, semi-structured and open. In a structured interview, the researcher uses a set of questions formed to probe subjects’ beliefs and practices. During a semistructured interview the researcher starts dialogue with a set of questions pertinent to the purpose of study and allows the flow of dialogues to take place with the participants. Semi-structured interviews give the interviewee an increased control over the flow of conversation. During an open interview, the researcher creates conditions and encourages the subject to talk freely and then probe the subject with deeper questions (Bogdan & Biklen, 2003).

I probed my participants’ epistemic views of science, their conceptions of science teaching and assessment and of their classroom practices in a series of audiotaped interviews. The interviews took place in a semi-structured form; I began with a set of formal questions pertinent to the purpose of my study, and then used participant responses as the foundation for forming new questions (see Appendix A) as my research evolved (Merriam, 1998). According to Bogdan and Biklen (1992), less structured interviews not only provide participants with more control over the content of conversation but also empowers participants to exert certain degree of control over the
meaning making process.

The semi-structured interview format ensured flexibility in that an answer to one question informed me about how I formulated the subsequent questions as the interview unfolded (Bogdan & Biklen, 1992; Denzin & Lincoln, 2000; Hiller & Diluzio, 2004; Kvale, 1996).

Interview Set

I collected my data through three sets of interviews. The purpose of the first interview set was to develop a “general understanding of a range of perspectives” (Bogdan & Biklen, 1992, p. 97). Due to the progressive nature of interview process, participants are usually able to reflect more deeply in the follow up interviews, often coming to a new understanding of their own experiences and beliefs as they articulate more precisely on their initial statements (Hiller & Diluzio, 2004; Kvale, 1996). As the interviewer develops deep insights and acquires new skills he/she is likely to invite the interviewee to be more precise and accurate about the phenomena under investigation. As the researcher attempts to clarify his/her understanding and explore the details the participant is forced to deal more rigorously with his/her responses and understanding (Kvale, 1996).

During the first interview set, I asked my participants general questions pertinent to science, science teaching, science learning and assessment of student learning (see Appendix A). The final two sets of interviews allowed me to ask deeper questions and push for further linkages, associations and contradictions that allowed for new information and understandings to emerge.

Document Analysis

I used two sets of documents; instructional and administrative documents in my document analysis (Bogdan & Biklen, 2003; Weber, 1990). The first form of document analysis involved critical examination of my participants’ instructional materials, such as lesson plans, course textbooks, lesson plans, homework assignments, lecture notes, tests, quizzes and laboratory worksheets. These documents proved to be powerful in terms of helping me recognize patterns and understand how the content of the textbooks, teachers’ lesson plans and assessments enriched or limited students’ learning in science. The analysis of these documents also enabled me to determine how well teachers’
assessments aligned with the purposes of the prominent national science education reform documents (AAAS, 1993; NRC, 1996). Furthermore, they helped me to determine whether teachers’ assessment practices reflected the fundamental characteristics of the forms of assessment advocated by the national science education reform documents on assessment (NRC, 1996; 2001; NSTA, 2006).

The second set of document analysis involved a critical examination of teachers’ handbook, administrative announcements, correspondences, national reform documents, newspaper articles, state standards and policy initiatives that aim to bring about improvements in educational system. Document analysis in this domain involved interpreting these documents, seeking for purposes, patterns, nuances of meaning and contradictions pertinent to the purpose of my research (Bogdan & Biklen, 2003).

Data Analysis

Several scholars (Biklen & Bogdan, 2000; Denzin & Lincoln, 2000; Guba & Lincoln, 1987) acknowledge that qualitative researchers analyze data throughout the research process. Therefore, data analysis simultaneously takes place as the researcher collects data. This ongoing process of analysis and feedback informs the researcher to reformulate the research question as s/he gains deeper understandings of the phenomena under investigation (Guba & Lincoln, 1989). Janesick (2000) argues that an inductive approach to analysis guides the qualitative researcher, which means that categories, themes and assertions come directly from data. This method of induction allows for the modification of concepts and relationships between these concepts to occur throughout the research process with the goal of most accurately representing the experience and understanding of participants (Janesick, 2000).

In my data analysis, I used grounded theory (Glaser & Strauss, 1967; Lacey & Luff, 2001). According to Miles and Huberman (1994) data analysis in qualitative research involves three stages; data reduction, data display and conclusion drawing and verification. Analysis of qualitative data usually starts with familiarizing oneself with the data through reviewing multiple types of data. Next, comes the organization of data into manageable forms for easy retrieval. After the researcher has organized knowledge into manageable forms comes the most important phase of data analysis. Then, the researcher
must ask himself/herself, *what do I want my data to tell me that might be significant in answering my research question?* The answer to this question will guide the researcher to choose an appropriate method of analysis.

The process of data analysis in grounded theory is “cumulative and can involve frequent revisiting of data in the light of the new analytical ideas that emerge as data collection and analysis progresses” (Lacey & Luff, 2001, p. 21).

I began analyzing data in the first stage of my research upon my arrival as a chemistry teacher in the context of my study. After I collected my data through the first interview set (see Appendix A), I started analyzing my data. I developed core coding categories, searched for consistent purposes, understandings, patterns and activities around the themes pertinent to the purpose of my study. I used some of these core categories such as teachers’ epistemic views of science, teachers’ conceptions of the learner and the learning process, teachers’ conceptions of teaching science and inquiry, and their conceptions of the purposes and practices of assessment.

Two analytic processes led the core categories to grow into the conceptual categories: constant comparison and continued questioning (Charmaz, 1990). After I completed my thematic analysis, I categorized the important themes relevant to the focus my study.

Strauss and Corbin (1990) maintain that all theoretical developments are considered provisional until proven by various data sources and by validation from others. If additional data need to be collected, the researcher should consider pursuing the goal of rigorous evidence.

After I developed my core categories, I confronted these conceptual categories with more data in order to define them carefully, delineate their properties, explicate the causes and demonstrate the conditions under which they operated (Charmaz, 1990). During this process of data analysis, I identified relationships between core concepts and categories, by constantly comparing them to form the basis of my conclusions and for verification purposes. After I generated my common themes and core categorie, I continued with developing my empirical assertions (Janesick, 2000). Then, I supported my assertions with direct quotations from the interview transcripts. Such method of data analysis makes it easier for the reader to trace the foundations of my assertions in the
findings of my study. This rigorous process of constant comparison and questioning of my conceptual categories led to the development of a core theory providing answers to my research questions (Strauss & Corbin, 1998).

Quality Criteria: Establishing the Rigor of the Research

The trustworthiness of qualitative research is a critical issue that demands special consideration throughout the research process (Glesne, 1999; Guba & Lincoln, 1987). Since interpretive research honors subjectivity and productively employs the researcher’s biases qualitative researchers need to establish a robust strategy for ensuring the credibility of their results (Glesne, 1999).

As I engaged in my inquiry by asking my research questions, choosing my theoretical framework, designing my methodology and choosing my subjects, I was cognizant that my perceptions, beliefs and values to certain extent guided my questions, observations and interpretations. Although this inquiry was vulnerable to subjective interpretations by design, I am aware that maintaining impartiality and credibility in the design, conduct, interpretation, and reporting of research findings is essential for the transferability of its results (Denzin & Lincoln, 2000; Guba & Lincoln, 1987; Patton, 2002). A key issue for this piece of inquiry, therefore, was to develop a shared understanding of the appropriate procedures for assessing its trustworthiness, thereby allowing the procedures and findings to be open to critical analysis from others (Denzin & Lincoln, 2000; Morse & Field, 1995; Patton, 2002). The following specific methods explain how I addressed these concerns.

Trustworthiness

Trustworthiness of a study corresponds to rigor of a study in conventional terms. There are two fundamental components of trustworthiness criteria, namely Credibility and Transferability (Guba & Lincoln, 1989). According to Guba and Lincoln (1989) Credibility criteria correspond to the term internal validity utilized in conventional research. The main promise of credibility criteria is to establish the match between the constructed realities of participants and those realities as represented by the evaluator” (p. 237).
There are several techniques that are used to assure credibility of a qualitative inquiry. Prolonged engagement, persistent observation, hermeneutic and dialectic cycle, and member checks are possible techniques of assuring quality in interpretive research (Guba & Lincoln, 1989). I used these methods to ascertain the trustworthiness of my findings.

**Hermeneutic Process**

I utilized the hermeneutic dialectic circle (Guba & Lincoln, 1989) to ensure the credibility of my findings. According to Guba and Lincoln (1989) hermeneutic dialectical circle involves continuous meaning making process with stakeholders actively providing feedback not only during the data collection process but also during data analysis and writing process. I taught in the context of my study as I collected and analyzed my data, therefore, I constantly shared my findings, insights and understandings with my participants and sought their contribution to my understandings. This process of hermeneutic circle encouraged diverse ideas to come to fore and allowed me to build comprehensive synthesis of the data that I collected through my observations, document analysis and interviews (Guba & Lincoln, 1989).

**Prolonged Engagement**

Guba and Lincoln (1989) define prolonged engagement as “substantial involvement at the site of inquiry, in order to overcome the effects of misinformation … to build trust necessary to uncover constructions” and engage oneself in the understanding of the context (p. 237). I was actively engaged in the context of my study for 18 months.

In order to meet the criteria of prolonged engagement, I stayed in the context of my study from the planning stage of my study through the end of data analysis period by constantly communicating with the stakeholders for 18 months. These interactions with my participants allowed me to construct and reconstruct my understanding of these teachers’ realities while going through and interpreting pertinent experiences in the same context. My involvement in this study both as a researcher and as a teacher allowed me to satisfy the criteria of prolonged engagement.
**Member Checking**

Guba and Lincoln (1989) argue that appropriate data collections methods constitute a prominent role in what meaning a social researcher attempts to produce. They argue that for authentic interpretations, one must fairly honor different constructions and values underlying participants’ meaning making and treat those constructions as invaluable sources of information (Guba & Lincoln, 1989). One may be able to capture such constructions through interview data, however; participants must verify researcher’s understanding of such constructions. Researcher can achieve that through member checking process (Guba & Lincoln, 1989; Janesick, 2000).

I used member checking to strengthen the authenticity of my findings and argument (Lincoln & Guba, 1985). I was often in contact with my participants both during data collection process and data analysis stage because we taught in the same school.

After I analyzed teachers’ interviews and received feedback from my major professor, I gave the participants a hard copy of their cases and asked them to comment on the findings that I thought my interview conversations and classroom observations revealed. One of my participants did not agree with my interpretations of his understanding of science. After I received feedback from my participant I made changes to my argument and presented cases to them for further authenticity. The participants thought that my final interpretations reflected their beliefs and practices across conceptual themes that I developed.

**Crystallization**

I employed multiple data collection methods in this study to ensure crystallization (Janesick, 2000; Richardson, 2000). To ensure contribution of crystallization to the trustworthiness of my study I cross-analyzed my raw data from formal interviews, informal conversations, reflection journals and classroom observations. Patton (2002) maintains that multiple methods of data collection allows researchers to observe whether the results from different methods lead to similar findings about phenomena under investigation. If findings from different sources lead to the same direction, then researchers feel more comfortable and confident about the credibility of their findings.
Janesick (2000) uses crystallization metaphor to describe robustness of qualitative research. She argues that crystallization allows for thoughts and experiences of participants and researcher to become alive and speak to the reader. I achieved the criteria of crystallization by using multiple sources for data collection and supporting my assertions by presenting relevant scripts from the interview transcriptions in my data analysis.

Moreover, my prolonged engagement in the study as a participant, a science teacher who went through similar experiences as my participants, allowed such quality to emerge more easily than if I was a researcher unrelated to the context of my study.

Peer Debriefing

Crotty (1998) maintains that the nature of interpretive inquiry allows biased interpretations. Although interpretive research honors researchers’ biases as long as supported by evidence, this bias is likely to limit the comprehensiveness of findings. By comprehensiveness I mean that there is enough evidence supporting my assertions that I assume to describe the reality of teachers’ experiences (Bogdan & Biklen, 1992). To reduce such influence and ensure comprehensiveness, I collaborated with my major professor and peers who were writing their dissertations at the same time as I was in an online reading group. This collaboration and sharing process ensured collective support and insight throughout my writing process.

I frequently communicated with my major professor through telephone conversations and e-mail, and shared my findings and interpretations with her. The feedback from our conversations substantiated my understanding of collected data by bringing in a critical perspective. This dialogue raised important questions and helped me to identify some important themes that I had not considered initially. I evaluated my major professor’s and my peers’ feedback to revisit my initial interpretations and, reflected on the comments from my committee members, thus, developed robust interpretations.

Transferability

Unlike the criterion of generalizability, transferability is the applicability of the
finding of the study to similar settings (Guba & Lincoln, 1989, p. 241). Tobin and Tippins (1993) argue that the fundamental purpose of interpretive inquiry is:

….not to convince readers of the generalizability of what has been learned but to provide sufficient details of the contexts in which the theory is embedded and to enable readers to decide on the extent to which what has been learned can help them meet their goals. (p. 9).

Tobin and Tippins (1993) remove the generalizability of the outcomes from the focus of inquiry and replace that with the transferability of the outcomes, thus, giving the readers the power to reflect on the details revealed in the findings and determines their usefulness for their unique cases. Guba and Lincoln (1989) further elaborate and state that “transferability is always relative and depends entirely on the degree to which salient conditions overlap” (p. 241). Although the arguments that my data analysis generated characterizes a unique school context, readers may be able to apply these findings in other contexts with similar purposes, activities, and characteristics.

**Thick Descriptions**

The purpose of thick description is to “provide as complete a data base as humanly possible in order to facilitate transferability of judgments on the parts of others who may wish to apply the study to their own situations” (Guba & Lincoln, 1989, p. 242). I met this criterion by constantly referring to quotations from interview transcripts, making references to my observations and sharing my own reflections about my experiences and understandings in the context of this study.

**Ontological Authenticity**

Guba & Lincoln (1989) propose ontological authenticity as one of key aspect of quality criteria, i.e., providing teachers with the opportunities to become more sophisticated in their response to the researcher’s interests. I achieved this criterion by utilizing a progressive interview process that involved three stages. Stretching timeline for the data collection increased my participants’ awareness of the focus of my inquiry (Guba & Lincoln, 1989). The interview questions during the first interview guided the teachers as they reflected on the key aspects of challenges and promises associated with
their assessment of student learning. For instance, Kevin revised his definition of science and volunteered to share with me his revised definition of science after he reflected on my question to him on the definition of science.

**Catalytic Authenticity**

Catalytic authenticity is another important aspect of Guba and Lincoln’s (1989) quality criteria, which means whether the evaluation process has contributed to stimulation of teachers’ awareness about their limitation in terms of assisting students with their learning through means of alternative assessment. I engaged in ongoing conversations with my participants at the lunch table and at the departmental meetings to heighten their interest in the purpose of my study. These ongoing conversations with my participants in the context of my teaching also allowed me to make insightful interpretations about the purposes of my study. For instance, these conversations enabled teachers to revisit what they already knew about assessment and the purposes that their assessments served.

**Educational Authenticity**

Educational authenticity refers to the extent to which the subjects come to understand, and learn from each other’s views. I met this criterion by engaging my participants in conversations that addressed my informational needs at the lunch table. I was able to track the progressive improvements in the sophistication of the arguments that they presented. These conversations convinced me that they had learned from their participation in my study. For instance, Ryan told me that his involvement in this study helped him to make adjustments to his assessment plans for the subsequent year. My conversations with Kevin reveal that Kevin’s participation in this study helped him to recognize the limitations of the projects that he assigns to his students each quarter. Kevin’s participation in this study encouraged him to think about ways in which he might be able to improve his projects to enrich students’ learning in science.

**Tactical Authenticity**

Tactical authenticity refers to the degree to which the research participants are empowered to act upon their understanding as a result of their learning from the research
process (Guba & Lincoln, 1989). Although tactical authenticity is central to the transferability of qualitative studies, the length of this study did not give me chance to observe if the teachers actually integrated assessment practices that serve the purposes of science education reform documents. This study ended at the end of the school year, and I did not have means to access that information in the summer.

Summary

The goal of my study was to characterize the challenges associated with the implementation of assessment reform in science classrooms. This required access to the ways in which my participants conceptualized assessment and assessed their students’ learning of science. My methodology and the methods that I used for data collection allowed me to access information on how my participants conceptualized assessment along with information on the influences of their epistemic views of science and their pedagogical conceptions on how they viewed assessment. These methods also enabled me to gather information about the cultural and political influences on my participants’ conceptualization of assessment and their assessment practices. As a result, the methodology I worked within the data collection methods used and the quality criteria I used helped me to substantiate my evidence that form the basis for my arguments I present in the following chapters.

I divided the findings section of my dissertation into five chapters based on the emerging themes. The following table summarizes the content of each chapter in my findings.

Table 3. The thematic content of the findings

<table>
<thead>
<tr>
<th>Chapters</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 4</td>
<td>Teachers’ Epistemic Views of Science</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>Teachers’ Pedagogical Conceptions</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>Teachers’ Conceptions of Assessment</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>Teachers’ Assessment Practices</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>Teachers’ Conceptions of the FCAT</td>
</tr>
</tbody>
</table>

I started the discussion of my findings by introducing the conceptions that make
up a teacher’s conceptual ecology (Southerland et al., 2006); teachers’ epistemic views of science, teachers’ pedagogical conceptions and teachers’ conceptions of assessment. Following my discussion of teachers’ epistemic, pedagogical and assessment conceptions, I introduced my findings pertinent to my participants’ assessment practices. Such organization of the findings enabled me to look at the nested relationship between various constructs that make up a teacher’s conceptual ecology (Southerland et al., 2006) and teachers’ assessment practices. However, to acknowledge the role of the school culture on teachers’ pedagogical and assessment conceptions and their assessment practices, I provided a brief discussion of the school culture at the beginning of my discussion of teachers’ pedagogical conceptions. I introduce my findings pertinent to my participants’ conceptions of the FCAT in the chapter that follows my other findings to better acknowledge the limitations of external structures for bringing about improvements to students’ learning.
CHAPTER FOUR

TEACHERS’ EPISTEMIC VIEWS OF SCIENCE

In this chapter, I present my findings pertinent to my participants’ epistemic view of science and their conceptions of school science. I finish this chapter by characterizing my participants’ epistemic views of science through Tsai’s (2002) classification of teachers’ epistemic sophistication.

Teachers’ epistemic views of science has become subject of several recent studies in science education (Abd-El-Khalik & Akerson, 2004; Abd-El-Khalik & Lederman, 2000; Applebaum & Kaplan, 1998; Bencze, Bowen & Alsop, 2006; Brickhouse, 1990; Kang & Wallace, 2005; Southerland et al., 2006; Schwartz & Lederman & Crawford, 2004; Tsai, 2002). Each of these studies describe that science teachers’ epistemic views of science are not consistent with contemporary conceptions of science (Abd-El-Khalik & Akerson, 2004; Lederman, 1992; Southerland et al., 2006). Many of these studies report that elementary and secondary school science teachers hold relatively naïve, unsophisticated views of science that fail to understand the complex structure of scientific knowledge (Abd-El-Khalik & Akerson, 2004). They further argue that teachers’ naïve epistemological conceptions lead them to limit the content of their teaching to students’ acquisition of canonical knowledge with limited or no focus on helping students develop scientific habits of mind (Applebaum & Kaplan, 1998; Bencze, Bowen & Alsop, 2006; Kang & Wallace, 2005; Tsai, 2002).

It follows that if teachers’ epistemic views of science are significant indicator of how they think about teaching science, it is essential that those beliefs be explored and made subject to critical investigation for improved practice. Furthermore, although science education scholars report significant correlations between teachers’ epistemic view of science and their pedagogical conceptions (Bencze, Bowen & Alsop, 2006; Kang & Wallace, 2005; Southerland et al., 2006; Tsai, 2002), the relationship between science teachers’ epistemic views of science and their conceptions and practices of assessment has not been pursued in science education literature. One of the purposes of this study is to address this gap in science education literature and develop an understanding into the nested relationship between teachers’ epistemic views of science and their conceptions.
and practices of assessment. It is my argument that developing such understanding can
guide us to help science teachers construct epistemic views of science that lead them to
teach science and assess student learning in ways that serve to the purposes of science
education reform (AAAS, 1993; NRC, 1996). However, before making that judgment we
need to present evidence on teachers’ epistemic view of science.

Epistemology is a branch of philosophy, which investigates the nature and origins
of human knowledge (Crotty, 1998). Many argue that science is a highly formulized and
complex branch of knowledge (Loving & Cobern, 2000; Crotty, 1998; Davis, 2004).
Therefore, analyzing the general nature of scientific knowledge is one of the main
interests of epistemological studies. This involves defining key characteristics of
scientific knowledge, understanding rules for the operation of a scientific habit of mind
and its means of justification (Abd-El-Khalik & Akerson, 2004; Loving & Cobern, 2000;
NRC, 2006; Southerland et al., 2006; Tsai, 2002).

Teachers’ Epistemic Views of Science

The findings from this multiple case study suggest that my participants held
distinct epistemic views of science in terms of their sophistication. I present my findings
on how these three teachers’ epistemic views of science differed from one another in the
following section.

Kevin’s Epistemic View of Science

The evidence below will show that Kevin’s understanding of science aligned with
characteristics of a positivistic view of science. Positivists maintain that in the social as
well as in the natural sciences, sense experiences and their logical treatment are the only
source of all worthwhile information (Loving & Cobern, 2000). They reject intuitional
attempts or cognitive speculation to gain knowledge and develop understanding (Crotty,
1998). Kevin’s epistemic views of science matched with the knowledge claims of
positivist camp as he viewed the presentation of tangible evidence as the only legitimate
claims to the knowledge.

Kevin looked for solid evidence in his definition of science. Kevin first limited his
definition of science to “anything that has to do with math and any kernel of knowledge
that has a mathematical application.” However, further evidence suggests that Kevin
struggled with conceptualizing his understanding of science.

Well at first I was thinking that [science is] anything that has to do with this planet and the different occurrences in there—but it is just more than just this planet and—I think science happens out in the space, too. I think it is just trying to find a valid explanation for why things happen. I think that is what science is—it covers chemistry, physics, biology and earth science. (Kevin interview, March, 12, 2005).

Although Kevin’s responses showed that his understanding of science lacked epistemological sophistication, I did not want to make a quick judgment about the epistemic nature of Kevin’s understanding of science. In order to clarify his response, I tried to ask my question in a different format. I brought religion in and allowed him to formulate his understanding of science in that context.

Well—Science is trying to prove valid reasons for why things occurred—whereas religion—that is generally a belief and not a firm fact of why things did occur that way—it is more faith—it tries to explain—this is why it happened—thegys[religions] try to explain things through some power or other supernatural reason—try to explain why it occurred this way—I personally feel more of a scientist than a religious person in that sense—I look for the firm scientific explanation for why things happen—rather than a religious explanation that attempts to convince people why things occurred—through some other unseen reasons—ethically, I think it is valid that people should know the scientific reason why things occurred. Now whether they want to believe that it is due to only just that, whether everything has to have a firm reason—I think that is up to each individual person to believe in, whether they think there should be extra explanation to why things happen—I think we should as scientists basically denounce that there is not anything else—I don’t think we should not go that far basically to totally cancel out religion that there is not such thing as this—is all scientific, that it happens for a reason and—I do think we should keep striving for more reasons for why different things happen so we can understand things better, but people find comfort and they should—in believing why things happen—and providing different explanations for why things happen. (Kevin interview, March, 12, 2005)

Although Kevin first limited his definition of science to a concrete body of knowledge, his answer to my question about the interaction between religion and science showed that his understanding of science involved making sense of how systems work through reasoning. Kevin believed that science’s mission is, “to try to have people understand that everything works for a reason.” He emphasized the role of “valid reasons” in his elaborated definition of science. However, valid reasons demanded presentation of tangible evidence in his elaborated definition of science. Such
understanding dictates that scientific facts and reason give power to the theories that people develop in their attempts to describe and understand the natural phenomena. Kevin’s commitment to limiting “valid reason” to the presentation of empirical evidence reflect the fundamental claims of positivist epistemology (Crotty, 1998; Davis, 2004).

Ryan’s Epistemic View of Science

Ryan attended a Christian college; therefore, religion played a significant role in how he viewed and understood the world. Furthermore, his religious beliefs not only influenced how he viewed science but also shaped how he taught science. Ryan said:

I went to a Christian school — and I am trying to incorporate what my Christian beliefs are into my teaching — obviously, I work at a public school you don’t mention Christianity or religion in general at all unless I am trying to making a broad statement that is— not any pro any particular religion — but it does influence the way I see the world — the way I understand the world. And I try to incorporate that into my teaching. That is what— I have gotten this through my education classes and my chemistry classes. And that was put in school where I was. For example looking at Chemistry not as just bunch of stuff to be taught, but as something made from a creator, God, definitely changes your opinion of what you are looking at. (Ryan interview, February 1, 2005)

This quote suggests that religion played a significant role in how Ryan looked at science. Ryan submitted to the God’s authority for explaining the natural phenomena. When I asked Ryan to share his understanding of science with me he made the following statement: “I mean obviously you think of science from a scientific standpoint” (Ryan interview, February 1, 2005). He stated, “Science is inquiry into understanding the world surrounding our lives and the world that extends beyond our understanding” (Ryan interview, February 1, 2005). Ryan maintained that the mass media frequently puts science and religion on the opposite ends of the same pole. He did not believe that this is true, however. Although Ryan’s religious convictions guided how he viewed science, he was able to interpret science and religion within their own separate spheres. For instance, Ryan viewed science and religion as two separate systems in terms of their explanatory power for understanding the natural world. Religion took precedence over science when the focus of understanding or claims to knowledge became issues of faith.

Ummmm [pause few seconds] — what is science? [pause] — I mean obviously you think of science from a scientific standpoint — a lot of time it is inquiry. I think a lot of times inquiry done for an ethical purpose — but it does not have to be
ethical — but it should be ethical, properly. I think those are kind of two components — just not inquiry, but doing it for a specific reason — but I think a lot of times these two components seem divorced from each other which I don’t think is true. (Ryan interview, February, 01, 2005).

Ryan rejected a view of science that built its foundations solely on pure evidence, logical thinking and experimentation. Although Ryan defined science as an enterprise for inquiry into the unknown, he expressed that as we inquire into the unknown, science should have boundaries. For instance, he mentioned science should not judge and disturb the validity of other spheres of knowing such as ethics or religion.

An analysis of Ryan’s comments suggests that looking at scientific research from an ethical perspective presents us with an opportunity to examine our world with regard to our daily experiences, our thinking and imagination without interference from an empiricist perspective of science. Further evidence suggests that Ryan’s understanding of science differed from the socially accepted and frequently referred definition of science, which confined science to a set of facts produced through careful experimentations.

Well I mean-it [science] means trying to find out answers to the problems that you don’t know. I mean trying to figure out through scientific method-through scientific way of reasoning and logic-the answer to a problem that you don’t necessarily know the answer to. (Ryan interview, February, 01, 2005)

Ryan viewed science as a method of finding the truth that is waiting to be discovered. Ryan’s understanding of science focused on the analytic method of thinking whose ultimate goal was to inquire about the unknown world and make sense of existing knowledge through logical reasoning. For Ryan, science meant deductive reasoning, which forms the foundations of Descartes’s philosophical argument, rationalism (Crotty, 1998; Davis, 2004).

*Melody’s Epistemic View of Science*

Melody’s understanding of science involved the discovery and exploration of facts, understanding of the processes of how things work and being inquisitive about the unknown. When I questioned Melody on her understanding of science, she said, “I think it is the exploration and discovery of knowledge. That can be new knowledge or knowledge that we think we already understand. That’s what I think science is” (Melody interview, February 18, 2005). This suggests that Melody viewed science as something
that people do rather than something that people know.

Melody did not limit the boundaries of science to academic subjects. In her extended response, Melody defined scientific method as “dealing with the ideas that are hard to find an answer to.” Melody said:

I think it’s like being inquisitive. Like trying to find answers to questions, like — I don’t think it has to be only chemistry, biology or physics. I think it could be like, “Why do people pick Coke instead of the generic brand?” you know — “Why are we gravitated towards things we like or why am I a teacher instead of an engineer?” Why do we do the things we do? I think it’s an answering of any kind of question. I have a sign on my walls it says science answers questions and questions the answers. And to me, it’s just kind of like that whole asking of “Why?” Like why things are they way they are and what are they made of. And like, why are we here, what is this table made of? It’s like the inquiry of life. You know, what happens if we put chemicals together? (Melody interview, May, 06, 2005)

Melody highlighted the importance of logical reasoning in her understanding of science. She maintained that science involves the process of thinking that leads to understanding of complex problems that demand analytical thinking by the analyzer to reach an answer. Melody established a connection between real life phenomena, human curiosity and the scientific inquiry. For instance, she said “science is the inquiry of life.” She viewed science as an endeavor into the unknown, a search for meaning in life and attempt for solving complex problems using logic rather than following a prescribed set of procedures to validate existing knowledge. Evidence presented above suggests that Melody held a relatively sophisticated understanding of science but her understanding of science reflected the attributes of rationalism (Crotty, 1998; Davis, 2004).

Characterizing Science Teachers’ Epistemic Views of Science

Science educators use different categorizations in order to understand teachers’ epistemic sophistication. I use Tsai’s (2002) model to characterize my participants’ epistemic views of science for the following reasons. First, this study by no means is an attempt to explore and understand the factors that mediate the development of teachers’ epistemic views of science. Rather, my purpose is to describe teachers’ profiles in terms of their epistemic sophistication. Secondly, Tsai’s (2002) model helped me to examine the nested relationships between teachers’ epistemic views of science, their pedagogical conceptions and their conceptions of assessment.
Tsai (2002) uses three different categories in his framework; traditional, process and constructivist. The following table provides a description for each category as well as information on how my subjects fit in those categories. This table will also guide my analysis of the nested relationships between my participants’ epistemic views of science, their pedagogical conceptions and their conceptions of assessment in the subsequent chapters.

Table 4. Teachers’ Essential Epistemic Views of Science.

<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptor</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>Science provides correct answers, or science represents the truth.</td>
<td>Accurate description; neutral or objective observations; objective interpretations; truths.</td>
<td>Kevin</td>
</tr>
<tr>
<td>Process</td>
<td>Scientific knowledge is discovered through the scientific method or by following codified procedures.</td>
<td>The scientific method; codified procedures; process of discovery; following scientific rules.</td>
<td>Ryan Ryan</td>
</tr>
<tr>
<td>Constructivist</td>
<td>Science is a way of knowing, and it is invented through scientists’ agreed conventions and paradigms.</td>
<td>Invented reality; imaginative acts; theory-laden observations; constructed through social negotiations.</td>
<td>Melody Melody</td>
</tr>
</tbody>
</table>

Adopted from (Tsai, 2002, p. 775).

Translating Kevin’s understanding of science through the lens of Tsai’s (2002) model suggests that Kevin fits into the “traditional” category because he viewed science as a collection of facts discovered through careful experimentation with limited reference to science “as a habit of mind” or science “as a way of knowing.” It follows that Kevin’s epistemic view of science did not resonate with the goals of science education reform, which advocate for a constructivist (Tsai, 2002) epistemic view of science.

Ryan fits into the “process” category because Ryan did not view science as a collection of facts and principles, rather he looked at science as a way of knowing with its own limitations. Although Ryan made references to science as a way of knowing he failed to articulate science as an evolving process. Rather, he viewed science as a mechanical procedure that people can use to reach an answer.

Melody looked at science as a way of understanding not only the natural phenomena but also the social phenomena. Because Melody used logical thinking as the
primary means for making sense of the natural and social phenomena, I placed Melody into the “process” category as well.

**Summary**

All of the participants considered the scientific theories and laws as explanations of natural phenomena built up through discovery-intensive procedures of making observations, proposing hypothesis and logically formulating arguments in the light of empirical evidence. Although they agreed on these features of scientific inquiry, subtle differences surfaced in how they perceived the impact of science on human life and human thinking. While Kevin viewed science as a collection of verified knowledge, my other participants acknowledged the role of the scientific process in the production of that knowledge. Furthermore, although Ryan honored multiple sources of reality both Melody and Kevin considered science as the dominating power in terms of its role in producing knowledge and developing understanding. I must acknowledge, however, that Ryan’s position towards science was filtered by his religious convictions. I use religious conviction in this context to refer to beliefs that are accepted without empirical evidence or beliefs that cannot be justified through logical reasoning. Ryan’s total submission to the God’s authority puts limitations on his epistemological sophistication. Such submission to a higher authority may inform how Ryan views teaching and learning and the ways in which he assesses students’ learning in science.

I present my participants’ conceptions of school science in the following section.

**Teachers Conceptions of School Science**

Science education reform documents place an emphasis on students’ understanding of the nature of science (AAAS, 1993; NRC, 1996), the relevancy of science to the real life context and fostering a scientific habit of mind (NRC, 1996). Reform documents maintain that understanding how science works and the processes that shape the nature of scientific knowledge empowers students to make informed judgments about personal, health and environmental issues (AAAS, 1993; NRC, 1996; 2001; NRC, 2006). It has been argued that education that focuses on the teaching of the nature and methods of science is “an active attempt to empower citizens for maintaining a democratic society” (Southerland et al., 2006). It follows that exploring teachers’ views
on the role of science in school curriculum is crucial as those views hold potential to inform us about the ways in which teachers teach science and assess students’ learning in science. I present my participants’ views on the significance of science in school curriculum in the following passages.

Kevin’s Views on School Science

Kevin maintained that every member of society should be exposed to scientific knowledge because it allowed the members of society to keep up with scientific advancements.

Well not everyone who takes Chemistry or Biology is going to become a research scientist but science classes kind of discriminate for those -who -who has the potential to become scientists and, hopefully, some of them will want to pursue a career in science by having a successful experience in those classes. But for the rest, they will still need to know that knowledge in order for them to be able to - like let’s say read a newspaper article which talks about a particular topic on science and be able to understand that. (Kevin interview, March, 12, 2005)

Kevin was cognizant that not all students in his classroom wanted to become a scientist or were interested in pursuing a career in science. However, he wanted to make sure that even those students who did not show interest in pursuing careers in science acquired the fundamental scientific knowledge. He maintained that even those disinterested students needed scientific knowledge to read a newspaper article that mentioned science. Kevin’s focus on the acquisition of knowledge in his understanding of school science reinforced his epistemic views of science in which he viewed science as a collection of facts and principles with limited reference to science as a way of knowing. It follows that Kevin held a naïve understanding of scientific literacy.

Ryan’s Views on School Science

Ryan maintained that the fundamental purpose of including science in school curriculum involved teaching for scientific literacy. Then, I asked him to define scientific literacy as:

Ummm — What I think it [scientific literacy] is that they [students] can take a basic scientific problem that they hear say in the media and they can evaluate it in the scientific perspective. You hear the most outrageous claims in the world, and people buy into them because they are not applying the basic principles of science
to that. Um — that would probably be the first thing about scientific literacy, the first thing that pops into my mind. I think as well, there has to be a basic understanding of science. I mean basic when you’re talking about people, they have to know the basic density of something will be the same no matter where you get it; basic things such as that. There is also scientific literacy as in the ethical questions of science. I guess that goes into evaluating scientific claims, based on complex of interest that people might have. People, naturally, I mean: you talk about the drug companies, students, parents, teachers, whatever. And they automatically know the implications of drug companies paying people to compromise. But really teach that into all contexts. (Ryan interview, May, 16, 2005)

Ryan focused on three components in his definition of scientific literacy. First, he believed that science should be part of school curriculum so that students are able to understand science and evaluate the credibility of scientific information they encounter in daily life through media. Secondly, he believed that people learn through science how the natural world functions (e.g. density). Finally, Ryan believed science should be part of school curriculum so that people are able to make informed decisions on ethical concerns about the influences of scientific developments on the environment and human health. Evidence presented above indicates that Ryan’s views on school science reflected his epistemic sophistication.

**Melody’s Views on School Science**

Melody viewed science as a significant component of school curriculum. Melody’s views on school science reinforced the fundamental goals of science education reform documents (AAAS, 1993; NRC, 1996), which advocates teaching for the development of a scientific habit of mind. Melody said:

I think that science should be taught as part of a required curriculum, because it is important for EVERYONE to have a working understanding of the way the world works, but more importantly, and the most important reason, is that everyone needs to know the scientific method, and how to go about solving a problem before they should be granted a diploma and the status of having a basic education. I don't think we should focus as much on breadth, but on depth into problem solving, and relating the scientific method to social problems, and basic problems that students encounter in their schools and daily lives. (Melody interview, May, 03, 2005)

Melody’s answer to my question indicates that she wanted science to be part of school curriculum for two fundamental reasons. First, she thought that the fundamental
scientific knowledge enabled students to develop a working understanding of the way the
natural world functions. Secondly, she believed that science fostered the development of
a scientific habit of mind that every citizen needed to make decisions in daily life.
Melody maintained that the scientific habit of mind empowered students to make
decisions about social and personal problems. This suggests that Melody’s views about
school science reflected her epistemic sophistication in which she viewed science as a
way of knowing and a search for meaning.

Characterizing Science Teachers’ Conceptions of School Science

The following table indicates the sophistication level of my participants’ views on
school science. Furthermore, this table indicates how my participants’ views on school
science reinforce the purposes of science education reform documents such as NSES
(NRC, 1996).

Table 5. Teachers’ Essential Views of School Science.

<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptor</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>Science needs to be part of curriculum because it contributes to our economical productivity.</td>
<td>Science helps us sort the best and motivated students and enables us to prepare them for future careers in science. The rest only need to acquire the scientific knowledge necessary to understand the media reports on science</td>
<td>Kevin</td>
</tr>
<tr>
<td>Process</td>
<td>Science needs to be part of school curriculum because it helps students learn about the scientific procedures necessary to solve complex problems</td>
<td>Science enables students to use the scientific method to conduct scientific experiments. Science empowers students to weigh the credibility of competing claims using the principles of science</td>
<td>Ryan</td>
</tr>
<tr>
<td>Constructivist</td>
<td>Science needs to be part of school curriculum because it helps students develop a scientific habit of mind</td>
<td>Science empowers students to become critical thinkers who are able to make informed decisions both about the natural and social phenomena</td>
<td>Melody</td>
</tr>
</tbody>
</table>

Adopted from (Tsai, 2002, p.775).

- **Summary**

Findings suggest that all of my participants viewed science as an important
component of school curriculum for people to acquire the scientific knowledge necessary to understand the ways in which the natural world works and understand the media reports on science. Melody and Ryan maintained that science should be part of school curriculum because it empowered students to think critically and weigh the credibility of competing claims using the principles of a scientific habit of mind. Kevin, on the other hand, focused on the role that the school science played in sorting the best and motivated students and preparing them for future careers in science. Furthermore, Kevin believed that acquiring the scientific knowledge for understanding the media reports that talked about science is essential for those who lacked skills and interest for pursuing careers in science. However, Kevin wanted his students to acquire the scientific knowledge in order to be able to read newspaper articles rather than helping students to make informed judgments about scientific issues surrounding their lives. It follows that while both Melody’s and Ryan’s views on school science reinforced the goals of science education reform, Kevin’s views on school science remained somewhat traditional and did not reinforce the goals of science education reform.

Also, while Melody and Ryan acknowledged parallel reasons for including science in school curriculum, Melody’s views on school science appeared to be more sophisticated than Ryan’s partly because Melody did not limit the boundaries of scientific way of thinking to the science itself. For instance, she wanted students to use scientific way of thinking in their everyday life decision making.

After presenting my participants’ epistemic views of science and their conceptions of school science, I present my findings pertinent to my participants’ pedagogical conceptions in the following chapter.
CHAPTER FIVE

TEACHERS’ PEDAGOGICAL CONCEPTIONS

One of the core arguments of recent science education reform literature is the belief that all students are capable of learning science (AAAS, 1993; NAS, 2006; NRC, 1996). The reform documents maintain that in order to create conditions necessary for success of all students regardless of their gender, racial and socioeconomic backgrounds, curriculum must be designed to meet interests, experiences, abilities and prior knowledge of all students (NRC, 1996). In addition to the improvements in curriculum, the reform documents state that teachers’ pedagogical conceptions and practices must also support the notion that all students are capable of learning science (NAS, 2006; NRC, 1996). In this chapter, I present my participants’ pedagogical conceptions; their beliefs about the learner and the learning process, and their beliefs about science teaching.

Kevin’s Pedagogical Conceptions

Beliefs on the Learner and Science Learning

Kevin acknowledged that some students were more intelligent than others, however, he believed that if properly taught, everyone had the intellectual ability to learn science. Kevin related the learning problems of his low achieving students to their lack of motivation rather than to their lack of intelligence. Kevin said:

Well — The intelligent students are not always the motivated ones but most of the time they are — I think they tend to have seen or understood that in the future they are better off for being better educated so — they will want to just learn for the learning sake and so that they are better educated and profit from that in the long run — but that is not always the case. Some of the smarter students still feel that they are too smart for it or that — may be they think they don’t need to learn this topic for what they might want to do with their lives in the future so they may not be motivated for this and concentrate on something that they are more interested in. The lesser intelligent students, perhaps some of them may just follow along with what they have been told perhaps by their parents or teachers that “you will need to learn this for a specific reason”. And they will say “ok” and they will just go ahead and do it. I don’t think that is majority for the lower level students either but it might be. I am not sure — but I am not sure may be that is a more closer percentage — a lot of times the lower level students are the ones — probably who have failed before and maybe — have kind of given up here and or don’t think they are smart enough and so they end up lacking the motivation maybe don’t see
in the end that they will be better off for it. So it is harder to get them and convince them that you need to learn it. Just may be you won’t be using chemistry in the future but just for the fact of learning things you learn how to learn things better, so that if you picked up and learned how to learn chemistry you will be better off for whatever you want to learn in the future, so it is kind of educating in general. (Kevin interview March, 12, 2005)

Kevin’s answer to my interview question indicates that he thought students failed in his class not because they were not intellectually capable of fulfilling the requirements of his class but because they were not convinced that education in general and chemistry particularly served their interests. Kevin held an authoritarian view on the learner in which he expected students to learn science by responding to the teacher authority rather than by searching for meaning in their learning. Thus, the students who were labeled as “underachievers” were not necessarily the incapable but they failed to relate to the purposes of school curriculum (Pinar et al., 1996). It follows that school curriculum as a structural factor failed to capture students’ interest and produce fruitful conditions for them to learn science to their maximum potential.

Although understanding Kevin’s beliefs about his students’ ability to learn science was of my interest, I was more interested in accessing his understanding of the process of learning. The following quote indicates how Kevin viewed learning.

Ummmm –learning is when they can explain something that you taught them back to you and maybe even in a different fashion and not in a complete memorized and straight back the same exact way that you taught them so to get them to learn, you need to explain the basis of it and then show them different examples of it and relate it to maybe other things so that when you can ask random questions about it, they can still respond with the proper answers. Even though they might not have seen the same exact question before but they can relate to something that you have taught them or relate it to something that they might have seen somewhere else. (Kevin interview, March, 12, 2005)

Kevin held conflicting beliefs in terms of how he thought science learning took place. Although some attributes of his beliefs reinforced the type of learning process defined by reform documents, the notion of “learning science for understanding and real life applications” (Gallagher, 2006; NRC, 1996), other attributes of his belief system contested that definition of learning (NRC, 1996). Although Kevin wanted his students to match their learning to his expert knowledge or the knowledge that could be found in the
textbook, he did not expect them to do so “in a complete memorized fashion.” Rather, he expected his students to apply the knowledge that they acquired to unfamiliar situations and understand that knowledge in the real life context.

However, additional data from my interview conversations with Kevin indicates that he defined learning as students’ ability “to come up with proper answers.” This definition suggests that Kevin viewed learning as students’ ability to match their learning with expert knowledge rather than a process that guides students to search for meaning making. This view of learning is limited and contests the learning objectives advocated by science education reform documents (AAAS, 2005; NAS, 2006; NRC, 1996).

**Contextual Influences on Kevin’s Beliefs on Science Learning**

Kevin primarily taught Health Academy students. Health Academy was considered as a special unit of the Southern High School, which prepared students for technical jobs in the health field. This perception shaped how Kevin thought about students’ learning of science. For instance, Kevin maintained that his students needed prerequisite knowledge necessary for entering the technical jobs that he envisioned his students would enter. Kevin was convinced that the best way to ensure that students acquire the prerequisite knowledge necessary for becoming successful in college classes was to deliver information through lectures.

> I do a lot of lecturing. Of course, the educators will say lecturing is not the best way to teach science but that is what students see when they go into college…and I still believe in lecturing that you need to be able to learn that way for one because that is what you will see when you go to college and because in different careers people will describe things to you and talk to you that you need to kind of be able to learn from it. (Kevin interview, March 12, 2005)

Kevin maintained that he taught science through lectures partly because that was what he expected his students would go through at college. He maintained that his students needed to develop skills of learning through lectures as he thought most college classes expect students to learn through lectures. However, this reasoning becomes problematic taking into account that he mentioned earlier that most of his students would not make it to college.

Kevin’s concerns about classroom management also shaped his ideas about students’ learning in science. For instance, Kevin often maintained that an effective
classroom management strategy is prerequisite to students’ learning. Kevin said, “You just can’t teach when the kids are not listening.” Thus, teaching through lectures enabled Kevin to maintain control in the classroom and empowered him to keep students quiet. The notion of maintaining authority in the classroom led Kevin to limit students’ learning to rote memorization of the scientific facts and principles with limited focus on the applications of that knowledge in everyday life context. Thus, Kevin’s pedagogical conceptions were not consistent with social constructivist pedagogy. Furthermore, his teaching goals contested the goals of science education reform. However, Kevin’s pedagogical conceptions were consistent with his epistemological conceptions.

Beliefs on Science Teaching

Kevin’s purposes for teaching science primarily involved students’ conceptual understanding of the chemistry concepts that he taught. Kevin maintained that he wanted his students to develop an understanding about how the natural world worked, more precisely how certain properties of chemicals enabled and limited the occurrences of chemical reactions.

Well — I am not covering all sciences. I am just covering some basic chemistry. I am just trying to explain to them and give them reasons [as to] why basic chemicals and chemistry do the things they do. [They need to know] they react this way because they have these properties. I could probably do more and in the future I would like to apply more how they see it everyday so they can relate it more and see more examples around them of why chemicals are doing the things that they do, but I generally end up with teaching them the fundamental principles. I guess mainly time concern sticking to just teaching the principles of it and not as much where they see it happening everyday. (Kevin interview, March, 12, 2005)

Kevin considered teaching as a vehicle for transmitting the scientific knowledge prerequisite to college science and expert practice in fields of employment that his students envisioned to enter. Therefore, he paid special attention to the teaching of scientific facts, logical reasoning and communication skills, which he considered were the skills that his students would need in the future. Kevin used logical reasons in his teaching of science to help his students understand the fundamental scientific laws causing the occurrences of natural phenomena. Also, helping his students to make connections between the content of his class and its real life applications constituted one
of the primary objectives of Kevin in his teaching. However, his commitment to content
coverage and his limited knowledge of the implications of chemical knowledge limited
his ability to help his students make the connection between the content of his course and
the applications of such knowledge in real life context.

Conceptions and Practices of Inquiry

Kevin maintained that he did not receive formal training on teaching through an
inquiry framework, however he believed his laboratory activities reflected attributes of
scientific inquiry. My analysis of his laboratory worksheets and my observations of his
laboratory activities indicate that his laboratory activities did not reflect the attributes of
inquiry. Kevin prepared solutions for his students, set up the laboratory tables for them
and expected them to follow the instructions given in the worksheet to do the
experiments. Then, he expected students to collect data and answer the questions on the
back of the laboratory worksheets to come up with an answer that reinforced what they
had learned in the lectures. Kevin’s understanding of inquiry reflected the attributes of
the “0” level of inquiry (Heron, 1971), in which the teacher asks students to confirm a
scientific principle by following a prescribed set of directions. However, we know that
reform documents (AAAS, 1993; NAS, 2006; NRC, 1996) define inquiry as fostering
students’ ability to form authentic questions, form hypothesis, design authentic
procedures to conduct an experiment and students’ ability to come up with an authentic
answer. Although Kevin considered his laboratory activities as teaching through inquiry,
the acquisition of conducting scientific research skills was just a peripheral outcome.

Kevin also considered the teaching methods that he used as inquiry partly because
he guided his students’ thinking through questioning. Kevin said:

I do try to answer their questions so that they understand things, but I do more of
a directed inquiry, where this is my goal for the day, this is what I want them to
learn, and then I try to go over it. In explaining things to them I kind of ask
questions and want them to figure things out on their own – It’s not free inquiry,
where they go off on their own -going off on this and that, it’s more directed to
where I want them to go, but I try to answer all the questions and try to get them
around back to the concept which I want them to learn. (Kevin interview, May,
06, 2005)

Although Kevin asked questions to generate interest in the subject he was
teaching, he answered the questions for students and expected them to understand the logical argument he used to answer the questions for them. Kevin’s naïve understanding of inquiry and his lack of the practical knowledge of inquiry reduced students’ learning to the acquisition of canonical knowledge. Also, Kevin adhered to teaching his laboratory activities through traditional means partly because the traditional methods gave him the control over the students. Then, he was able to use that authority to maintain his prestige as an effective teacher, as effectiveness was defined based on teachers’ ability to maintain control in the classroom.

Kevin defined inquiry as “students exploring on their own.” This definition did not hold the complexities of understanding what meaning students needed to be making or of the scaffolding necessary for inquiry learning. Thus, his definition of inquiry led to a duality of either he told students or students discovered knowledge for themselves. As a result of the fear to lose classroom control to his students, Kevin limited his inquiry to directed teaching. Kevin said, “I am just trying to explain to them and give them reasons [as to] why basic chemicals do the things they do and they react this way because they have these properties.” This evidence suggests that Kevin refused to give students freedom to construct knowledge on their own using their faculties. Instead, he constructed knowledge for them and expected them to understand the logic and evidence that he used to justify his answer.

Although teaching scientific inquiry skills was one of Kevin’s primary goals in his laboratory lessons, Kevin’s naïve understanding of inquiry reduced the purposes of laboratories to verification of existing scientific facts and laws. It follows that Kevin’s conceptions and practices of inquiry did not reinforce the inquiry goals advocated by the prominent science education reform documents; Benchmarks for Scientific Literacy (AAAS, 1993) and NSES (NRC, 1996).

Ryan’s Pedagogical Conceptions

Beliefs on the Learner and Science Learning

Ryan also believed that all students were born with innate ability to learn. However, Ryan viewed learning as achievement rather than a search for meaning making (Gallagher, 2006; NRC, 1996; Taylor, Gilmer & Tobin, 2004). Therefore, Ryan established a strong correlation between hard work and learning. Ryan said:
Everyone has natural intelligence, but I think; you either use it, or you lose it. It’s very important that people cultivate intelligence. The societal impacts on people obviously profoundly affect how they think about the world, how they see the world, and their ability to learn as a socially constructive thing. You take the extreme examples, where people have no human contact, and the people come quite unintelligent—and well, you don’t want to use the word dumb. But, it’s kind of a harsh word to say. And I think, especially with the students that I have, that maybe aren’t that smart at all-naturally. And you can tell that they are not quite smart—and you can tell as well that they—when they go through the year, they became smarter and smarter, just because they apply themselves. If you work on thinking, and you work on working hard just in general, and while you work hard you think; I think you are able to increase your intelligence just a bit. (Ryan interview, April, 03, 2005)

Ryan maintained that a student’s own intellectual ability determined their ability to achieve the learning outcomes. Ryan used the brain as a muscle metaphor in his description of students’ learning. Ryan believed that learning activities in his classroom provided a medium for his students to enhance their intelligence. Ryan was convinced that his assessment tasks contributed to his students’ cognitive development as he thought the questions forced his students to think about and relate to the content of his lecture in a critical manner.

Ryan also acknowledged the role of social interaction in the cognitive development of the learner. Although he recognized that the learners’ interaction with their environment shaped how they understood, thought about, and responded to the world around them, it was not clear from my observations of his teaching that Ryan incorporated this awareness into his teaching. According to social constructivism, the teacher uses various pedagogical techniques to enhance the interactions in classroom where the learners construct knowledge interdependently with their peers and with the teacher (Cook-Sather, 2002; Shepard, 2000; Solomon, 1987; Tobin et al., 1996; Vygotsky, 1978). My observations of Ryan’s teaching suggest that Ryan did not allow his students to interact with each other or with him during his teaching.

Ryan also acknowledged the relationship between students’ motivation to do the work and their learning. Ryan firmly believed in his students’ intellectual ability to deal with the challenges that he presented to them in his assessment practices. Therefore, instead of relating students’ achievement in his class to their intellectual ability, Ryan related students’ success or failure in his class to their motivation. Ryan said:
In IB, I think it’s when they hear you say that “you need to do your homework” “Yeah that’s what every teacher tells me, and I’m not going to do it.” As they say, “I’m lazy, I’m not going to do it.” I feel that at least in IB, the ability for me to help students is quite limited. They’re either going to do it or not. (Ryan interview April 4, 2005)

Ryan believed that students needed to complete the homework assignments in order for them to learn science. This notion, relating learning outcomes to the amount of work done by students is a reflection of an authoritarian pedagogy that is quite consistent with behaviorist pedagogy.

Contextual Influences on Ryan’s Beliefs on Science Learning

Ryan is the teacher with two years of teaching experience. Ryan primarily taught the IB chemistry classes.

Ryan expressed the influence of structural constraints on his teaching. Ryan maintained that the 45-minute classroom period did not provide enough time for students to intellectually engage in learning. Ryan’s perception of time limitation led him to reduce his teaching to the review of important chemistry concepts in the form of lecturing. However, Ryan himself did not believe learning took place by listening to lectures. For instance, Ryan once said, “If you want to learn chemistry you have to invest extra time at home by reading the book, thinking about it and doing the homework. You just can’t expect to learn chemistry by listening to lecture and taking notes. I mean you must study for it” (Ryan interview, March 2, 2005). This statement shows that Ryan used a banking metaphor (Feire, 1970; 2002; McLaren, 1986) to describe learning. Thus, Ryan’s perception of time limitation led him to conduct his teaching in a way that made students dependent on him for learning.

The administrative policies and pressures about classrooms management also influenced Ryan’s perception of how students learned. School administration was convinced that maintaining control over student talking and controlling the interactions taking place between them ensured learning. Ryan strictly limited student-student interaction and students’ interaction with him during his lecture because of the administrative pressure. By maintaining control over his classroom, Ryan met the perceived accountability demands of his professional contract. Not only administrative
pressures but Ryan’s own pedagogical beliefs on how learning took place also led him to correlate student’ learning outcomes to maintaining discipline in the classroom. This notion of maintaining control in the classroom led Ryan to focus on the efficiency of his teaching rather than creating a learning environment conducive to fostering students’ acquisition of science literacy goals advocated by science education reform documents (AAAS, 1993; NRC, 1996).

**Beliefs on Science Teaching**

Ryan summarized the purpose of his teaching around four objectives; 1) helping students acquire the fundamental facts of chemistry; 2) helping his students to be prepared for taking advanced chemistry classes; 3) helping his students acquire and appreciate the fundamental attributes of scientific thinking; 4) teaching his students about the ethics of conducting scientific research. Ryan said:

> Obviously you want them to understand the basic facts of your subject, the class you are teaching and whether they use the science on an on-going basis or not. I mean — I hope one of my objectives is that I hope it will make college chemistry much easier for my students, which is CHM 2 and a lot of people think that is one of the toughest classes at college. And so — no matter where they take chemistry next, I want them to have a solid foundation for what they have begun. But my other objective is to learn scientific ways of thinking and reasoning. I want them to be able to think of the world in the scientific way either through laboratory or whatever. And especially to understand that science properly [done] should be ethical. I guess that is another thing that came a lot from my religious beliefs. That you cannot divorce science from the ethics. Well, science does not have answers to all questions mankind has. Otherwise, science, well, is not useless but it is — it can be manipulated for every evil end known to man. So I also want them understand that as well. (Ryan interview, February, 01, 2005)

Ryan expressed conflicting purposes for teaching science. Although teaching the fundamental scientific content and fostering a scientific habit of mind was Ryan’s stated goals, he accommodated other goals in his teaching as well. For instance, one of Ryan’s objectives in his teaching was to assist his students to become better consumers of scientific knowledge. Ryan wanted to challenge science’s established, unquestionable authority in his chemistry classes. On the other hand, Ryan primarily focused on teaching science for students’ acquisition of scientific facts and laws in his classes.

Although teaching chemistry through inquiry and fostering a scientific habit of
mind among his students was Ryan’s stated objectives, his perception of time limitation and his commitment to covering the intense IB content guided him to exclusively focus on the transmission of established chemistry concepts. Ryan said that “you have to go through this entire curriculum and unfortunately that does not give you a lot of time for doing other things.” Although the IB program’s intensive curriculum drove his teaching during the lectures, Ryan was able to foster students’ acquisition of inquiry skills in his laboratory activities. Ryan said:

Well, ideally, if you had an infinite amount of time, you’d make them go through scientific inquiry every time. But you don’t have infinite time, and so you’re always trying to teach the parts of scientific method. Whether it’s just simple as data collection or whether it’s, and that’s actually when they go into IB, there’s actually seven different categories that they grade for. And the IB labs, they actually send them away; students do them, I teach the grades and they actually have to be sent into Europe. And they have graders there check the original teacher’s work, and they grade them. And one of the categories is just simple data collection. There’re other categories of planning and um, there are categories of evaluation of results, and I try to follow those categories when I make different labs. You’re trying to teach scientific inquiry but you’re also trying to teach a point as well; whether it’s Hess’s law, or Boyle’s law, and the process. So I mean, there’s a kind of a given thing of both takes that you’re trying to do. Scientific inquiry, and also teaching of a point. (Ryan interview, February, 01, 2005)

Although Ryan highly regarded the teaching of scientific inquiry, his version of scientific inquiry was to help students match their understanding to the established truths (e.g., Hess’s law). Ryan’s understanding of inquiry remained at “Level 0” (Herron, 1971) in which the teacher guides students to confirm an established scientific law by following a set of prescribed procedure. Ryan focused on teaching his students the proper use of instruments, data collection and interpretation skills in his laboratories. However, Ryan mentioned that his laboratory activities did not reflect the true characteristics of scientific inquiry but were just an introduction to the world of scientific inquiry. Ryan attempted to teach his students scientific inquiry skills partly because the IB program required teachers to teach science through an inquiry framework. However, Ryan’s commitment to content coverage combined with his limited knowledge of and experiences with inquiry generated a set of teaching practices that took away the scientific inquiry component from the heart of his science teaching practices.
Conceptions and Practices of Inquiry

Ryan viewed inquiry as a process and a habit of mind that students could use to find answers to the complex questions that we do not know answers to. Ryan said:

Well I mean — it means trying to find out answers to the problems that you don’t know. I mean — trying to figure out through scientific method scientific way of reasoning and logic, the answer to a problem that you don’t necessarily know the answer to. (Ryan interview, February 1, 2005)

Ryan’s definition of inquiry reduced inquiry learning to scientific way of reasoning, developing logical arguments for solving complex problems. However, science educators define inquiry as students’ ability to formulate questions, devise ways of solving complex scientific problems, collect evidence, thus, to produce authentic answers (Blanchard, 2006; NRC, 1996). It follows that Ryan’s understanding of inquiry did not reinforce the definition of scientific inquiry outlined in science education reform documents (AAAS, 1993; NRC, 1996).

Although Ryan thought highly of learning through inquiry, Ryan gave his students limited opportunities to learn science through inquiry. Ryan said:

I want them to have a solid foundation for what they have begun with but my other objective is to learn scientific ways of thinking and reasoning. I want them to be able to think of the world in the scientific way either through laboratory or whatever. (Ryan interview, February, 01, 2005)

Although Ryan wanted to foster a scientific habit of mind among his students, his teaching primarily focused on transferring expert knowledge to his students.

Ryan’s perception of the priorities of the IB program limited and shaped how he went about fostering a scientific habit of mind in his classrooms. For instance, Ryan thought that the IB program prioritized the content coverage over skills of inquiry. However, the IB program booklet encouraged the IB teachers to teach science through an inquiry framework. Ryan said:

I mean you have to go through this entire curriculum and unfortunately that does not give you a lot of time for other things. Although a lot of labs you do — you try to make — at least you get the students into thinking in a scientific manner. And you get them trying to look for an authentic answer so they are trying to find an answer for themselves. Or they are trying to verify something. Like even well — few weeks ago I did a lab where they had to completely come up with ideas as to why the precipitates were formed in groups such as that. Then there is one time it is not so much finding the answer but could you make two grams of a product
from stoichiometric calculations and so it does make them at least predict what is going to happen using scientific knowledge that they have learned to get to a real answer. (Ryan interview, February, 01, 2005)

This quote suggests that Ryan encouraged his students to take ownership over their learning and seek for an authentic answer. However, encouraging students to come up with an authentic answer was not a common practice in Ryan’s laboratory activities. Ryan’s laboratory activities mostly took place in traditional manner where students followed a set of instructions to conduct the pre-designed experiments. Since Ryan said that one of his purposes in the laboratory was to get his students to come up with an authentic answer and because authenticity is central to doing of inquiry (AAAS, 1993; NRC, 1996), I asked Ryan to tell me what he meant by, “trying to get his students come up with an authentic answer.” Ryan said:

I think in the context of labs, when I’m going for authentic, I want them to do something that has some meaning for them. Just not something that they do to get the answer and it’s kind of out there, for the number. I think one of the labs is can you make two grams of a compound. Well that’s more of a mission that they’re trying to do in a way. They’re trying- interest is the best way to look at it. You get them interested because they’re not just doing a lab assignment for the sake of doing a lab assignment. They’re actually doing something- maybe, to have the spirit of scientific inquiry. They’re really trying to go for an answer as opposed to saying ‘this is the answer. (Ryan interview, November, 14, 2006)

Ryan described authenticity as students’ ability to think about the problem presented to them in an authentic manner. Authenticity for Ryan meant students’ ability to find answers to the questions presented to them on their own using their unique methods. However, Ryan expected students to prove a point in their learning of inquiry rather than allowing students to search for meaning by using the methods of science. Ryan’s stated beliefs in the quote cited above suggest that Ryan held a rationalistic pedagogy in which he measured students’ learning by their ability to formulate logical arguments, methods for finding answers to complex scientific problems.

Ryan wanted his students to develop a love for learning, allowing their curiosities and their motivations to drive their learning. However, he did not think that it was his responsibility to exploit these attributes of student learning and scaffold his instruction; thus, make learning interesting for his students. Although Ryan expressed beliefs that
partially reinforced inquiry goals outlined in *NSES* (NRC, 1996), his perception of time constraints discouraged him from teaching chemistry through inquiry.

I think the biggest limitation in this school to me doing more inquiry labs and projects and stuff like that is the seven hour day. Because we have 45 minutes in a period. Trying to get them to do an inquiry-oriented activity in a 45 minute period I mean -if you break it up, all the thought process involved in doing an inquiry lab in a productive manner is lost. I mean you are trying to find an inquiry oriented activity that you can do in 45 minute is pretty difficult. I mean a good inquiry oriented lab will take about 7 academic hours and seven periods a day kind of makes them worry more about grades than thinking about the thought process involved in doing things. I mean learning chemistry through inquiry. You just cannot do it in 45 minutes. (Ryan interview, February, 14, 2005)

Although Ryan wanted to promote key skills of scientific inquiry in his laboratory activities, his perception of time required to cover curriculum mandates guided him to teach laboratories in a traditional mode. During my observations of his laboratory activities, Ryan was always in a rush helping his students follow the directions so they could complete laboratory activity before the bell rang. Ryan challenged his students to show that they were able to use the scientific method in an artificial manner to conduct pre-designed experiments; use the laboratory instruments, collect data and be able to interpret the data they collect by answering the questions that he posed at the end of laboratory worksheet. Although students went through the steps involved in doing scientific inquiry, they did not derive questions and formulated ways of solving them. In that sense, Ryan reduced inquiry to students’ ability to follow directions, think critically for matching or comparing their answers to the established scientific truths. Herron (1971) describes such form of inquiry as “Level 1” inquiry in which students investigate a question posed by the teacher through a prescribed set of procedures. Such laboratory practices do not reinforce students’ acquisition of inquiry skills outlined in NSES (NRC, 1996). Although Ryan’s conceptions of inquiry remained traditional, he held potential to develop sophisticated pedagogical conceptions, thus, teach science through an inquiry framework.

**Melody’s Pedagogical Conceptions**

*Beliefs on the Learner and Science Learning*

Melody believed that all students are intellectually capable of learning science.
However, she acknowledged presence of other factors besides intelligence that determined how students went about their learning. Melody said:

I tend to think that every kid can do everything, but it’s a matter of “How much it’s really worth.” So I think like, I mean even if you took an ESE kid, they could be a chemical engineer if they really wanted to. But, is it really worth their effort? I mean – is it worth all the tremendous months of tutoring, and studying or whatever. But I really think it’s a part of finding their own aptitude. Some kids belong in IB, some kids don’t. But I do think that all students have the ability to learn. I don’t believe that there are some kids that can’t be taught. But other than an injury or a trauma to the brain – other than that, I don’t think there’s any kid that really can’t learn. (Melody interview, May, 3, 2005)

Melody maintained that students’ interest in pursuing academic goals that the schools promote and their interest in science determined how much science a student can learn in school. Although Melody believed that some of her students were motivated to learn science, learning was defined as achievement; the willingness to put effort and ability to see value in the end product of learning.

Melody viewed science as an academic subject that she thought everyone needed to learn. However, she maintained that without first exposure to science, a lot of students are discouraged from learning science.

I think everyone should be exposed to it [science] at least. Just like I think that is the beauty of public schools, like kids should be exposed to music and arts and math and reading and you know, “How do you know you don’t want to do something unless you’ve tried it, and how do you know you like something unless you’ve tried it?” yeah, I do. I don’t think you can close the door on something unless you’ve had the opportunity to try it out. (Melody interview, May, 03, 2005)

She maintained that students’ perception of their cultural and logical capital, and their perceived capacity for success in science, influenced their participation in science learning. She argued that in order to overcome this limitation, all students should be exposed to science. She stated that making science learning mandatory for all students would allow some students to realize their capacity for taking advanced science classes or pursuing careers in science.

Melody maintained that in order for students to learn science to the best of their abilities, students should first value their learning experience and be able to connect what they learn in science classes with their everyday experiences. Melody said:
You definitely have to have something to attach to — either it is just their experiences in everyday life or something else — breathing in and out — and connecting that to the fact that the air is matter and you have to have different examples — And they have to hit different ways of learning like hearing, seeing, listening and discussing. (Melody interview, February 18, 2006)

Melody did not view learning in a stimulus-response fashion, rather, she thought of learning as a process that involved students’ engagement through different means of sense making (e.g., listening, seeing). Melody said:

I think it takes time and they kind of have to kind of digest it a little bit — you know similar to eating. They have to take it in chew it up for a while and you know toy around and sit for a while. I think there are a lot to be said for different types of assessments to really give them a chance and have them really work with the information and like working with play and moving things around and trying different scenarios. I think learning is similar to play in that sense. (Melody interview, February 18, 2006)

Melody viewed learning as something that students do on their own using different means rather than something that the teacher does to the students. It follows that Melody’s beliefs on how learning takes place are consistent with social constructivist pedagogy (Tobin et al., 1996). Although Melody presented sophisticated pedagogical conceptions about the process of learning, her teaching practices did not reflect this understanding. Melody scaffolded her instruction in ways that were not consistent with social constructivist pedagogy. Because covering curriculum mandates took precedence over students’ learning, Melody’s teaching remained traditional. For instance, curriculum mandates encouraged teachers to sequence the course content, Melody gave each concept limited coverage time, leaving students with scarce time to deal with the complexity of scientific knowledge presented to them. For instance, although Melody integrated hands-on learning activities in her teaching, students had scarce time to reflect, think and write about their experiences after they engaged in hands-on activities. In addition to curriculum mandates, structural factors (e.g., students’ socioeconomic background, their level of motivation and their reading comprehension skills) caused frustration for Melody and led her to equate students’ learning with task completion.
Contextual Influences on Melody’s Beliefs on Science Learning

Melody is the teacher with three years of experience, who taught physical science and earth space science classes. Melody primarily taught standard level classes. Melody mentioned that her students did not care about their own education. She highlighted the fact that these students generally came from generational poverty and their life consisted of struggle for survival on a daily basis. Melody knew that these students often lacked resources, had a hard time making an adjustment between their survival goals and the goals that the school expected them to accomplish. Although Melody recognized the learning potential for this particular group of students, she felt powerless in terms of her ability to motivate them for achieving academic goals for her class.

When I questioned Melody on what she thought her students needed in order to perform well on academic objectives, she gave me the following answer:

They need more outside of the classroom kind of stuff, like, they need the whole yard. If they are not getting it, they like, need it at home, they need the whole — like if they’re not getting read to, they’re not getting encouragement, and they’re not having a parent look over their homework with them, they almost need an after school program where somebody is checking that they like “What do you have first period, what do you have second period?” They just don’t have the basic skills about getting here, and like, having somebody care about whether they have homework or not, and how to help them with it afterwards. So I think just in general, they need like organization. I mean they need like the whole nine yards. I mean like, organization and how to care about the school. (Melody interview, May, 03, 2005)

By establishing a healthy rapport with her students, Melody was able to recognize that many of her students had challenging lives and unhealthy relationships at home, which in turn resulted in inadequate supervision of their academic progress. She was aware of the fact that those students brought their problems with them to the classroom, and so failed to focus on measuring up to the academic expectations of the school. To address this issue, Melody modeled study habits to help her students and presented herself as an adult, who cared not only for her students’ academic lives, but also for what happened at home.

Establishing a good rapport with her students enabled Melody to develop an understanding of their learning needs, thus, become responsive to the social and cultural constraints on some of her students’ learning (Berliner, 2006; Darling-Hammond, 2003;
Noddings, 1984; 1992). This responsiveness led Melody to focus the purpose of her teaching to social mobility (Labaree, 2000) rather than teaching science for intellectual inquiry, search for meaning and understanding (Gallagher, 2006; NRC, 1996). For instance, Melody’s major concern was to help students graduate from high school, thus, improve their lives. This reduced the purpose of science teaching to teach for certification and assess students’ learning in a way that helped them meet the requirement of high school graduation. This case highlights that structural factors such as students’ lives at home, thus, the cultural capital that the students bring with them to the school presents major challenges for teachers who want to teach science for understanding.

**Beliefs on Science Teaching**

The nature of Melody’s beliefs on science teaching varied with the group of students she taught. In her standard classes, Melody focused on teaching her students organization skills, reading comprehension skills, and helping them to establish study habits. It follows that in Melody’s standard classes teaching students the life survival skills took precedence over teaching science for conceptual understanding and students’ acquisition of inquiry skills. Also, Melody’s judgment of her students’ potential to pursue higher education determined the nature of her teaching goals. For instance, she stated that when her standard students finish high school they are most likely going to pursue a career in fire fighting, roofing or air conditioning. Only few of those students, she believed, had the motivation to attend a four year college. Therefore, Melody reduced the purposes of her teaching to students’ acquisition of the skills of organization, communication and assuming responsibility for responding to the authority, which she thought were the skills that her students would need upon their graduation from high school.

Melody’s primary goals for her honors classes, on the other hand, involved helping her students develop a solid understanding of physical science concepts and preparing them for taking advanced science classes. Melody also wanted her students to develop a positive attitude towards their personal success in advanced science classes. Melody said:

Some of the kids are terrified with touching anything that says chemistry or physics. But somehow if they have that class [physical science] with me, I mean,
by the time they are finished with me and by the time the school year is gone, they will have a higher level of maturity to handle a physics or a chemistry class. (Melody interview, February, 18, 2005)

Although Melody held differing goals for the two student populations she taught, some commonalities existed across the two groups. One goal that overlapped was Melody’s goal to help her students develop curiosity, a sense of wonder, understanding and a sense of appreciation for the complexity of the living and non-living world surrounding their lives. Melody said:

One of my biggest objectives is to get them understand the complexity of the nature. Like, get them understand why things happen the way they do. So I try to help them understand that. I think — like the biggest thing that is really important for kids is really is to be able to explore on their own and see it first hand and actually do it and create it themselves. Even though it is already been done and discovered and been written about many times and we are probably not going to be able to explore or discover anything new in the classroom. Even Newton’s Law the things that have been done so many times, that is very popular in the news, it is still new to them. So I want them to be able to think about the causes of the things they observe in their lives. Like, why we use seat belts in our cars? You know things like that. I also want them to be able to use the scientific language and explain things using the scientific vocabulary. Like I want them to know what it means when they see the word thermodynamics or acceleration. So, I want to help them get that, too. (Melody interview, February 18, 2005)

This suggests that Melody’s primary purpose in her teaching was to help her students understand the scientific theories. To facilitate students’ understanding of scientific laws, Melody encouraged her students to make connections between the content that she discussed in the class and the applications of that content in real life context. For instance, while teaching about Newton’s laws, Melody talked about how the law applied to the seat belt technology. Melody also promoted scientific thinking skills by facilitating the discussion of relevant and socially significant topics, such as the stem cell research, traffic accidents and environmental issues. In such cases, Melody did not expect her students to conform to her authoritative knowledge, rather she promoted free thinking by leading the discussion and encouraging students to express their opinions on how they felt about these issues.

Although teaching for scientific literacy was Melody’s primary objective in her teaching, additional factors influenced how Melody set her teaching priorities. For
instance, Melody’s recognition of her students’ low reading skills encouraged her to integrate more reading assignments in her teaching. Melody’s assessment practices had an explicit reference to the standardized test (FCAT). For instance, Melody integrated multiple choice questions on her tests, not primarily because she considered multiple choice questions as powerful means to enhance student learning, but only because she wanted her students to develop test taking skills. Melody wanted to equip her students with knowledge and skills necessary to perform well on the FCAT as she viewed FCAT a ticket for her students’ graduation from high school.

Melody formed her identity as a teacher not only by the present conditions of her teaching responsibilities but by the future possibilities that she envisioned for her students. Melody recognized the classroom not only as a place for creating opportunities for her students’ academic achievement but also a place for their social advancement. In the context of her standard classes, Melody looked at teaching as an educational activity creating opportunities for her students to expand their horizons and assist them to think beyond what their local communities could offer both socially and economically. It follows that although Melody’s primary teaching goals involved teaching for scientific literacy and helping her students to develop a scientific habit of mind, contextual influences and her sense of community forced her to teach science for social mobility (Labaree, 2000), thus, limit her teaching to the minimum requirements for graduation.

Conceptions and Practices of Inquiry

Melody mentioned that she did not have enough practical knowledge about teaching science through an inquiry framework. However, she believed that her teaching reflected the fundamental elements of inquiry.

I did not get any formal training on it -I never studied on my own, but I think, I do think that I teach that way. I have been told by many students that I never answer questions. I just ask more questions. I try not to tell answers to them like when they ask me “what is the definition of ? Well, you have the book you have your notes -go ahead and look at it and then we will discuss it.” (Melody interview, February, 18, 2005)

Inquiry for Melody meant encouraging students to find answers on their own to the questions posed by the teacher. However, the questions did not have to be of kind that would force students to think critically, research and elaborate on concepts deemed
important in science (NAS, 2006; NRC, 1996). Although Melody encouraged her students to ask questions and often responded to the questions that her students raised by asking more guiding questions, she made sure that the answers that her students provided conformed to the knowledge that she presented to them through her lectures or that could be found in their textbook. Melody considered learning as inquiry as long as students engaged in thinking for finding correct answers. Melody’s understanding of inquiry fits well with “Level 0” type of inquiry in which students use a set of prescribed procedures to conform an existing fact (Herron, 1971).

Melody maintained that she did not limit her instruction to transmission of information from the teacher to the students. However, she expected her students to match their learning with the textbook knowledge. Melody said:

I try to present as much information as I can in terms of basic kernel of information. I give them examples, chances for them to do, chances for them to ask, chances for them to discuss. I try to hit all the different intelligences so they can do, hear, see, ask, discuss and write all that. At the same time I try to give a little bit mystery to it so when they are doing the lab they are actually discovering something on their own where it is not so forced. (Melody interview, February, 18, 2005)

Melody viewed learning as students’ ability to discover the established truths presented in the textbooks or delivered to the students through the teacher by using the power of logical reasoning. Melody used diverse methods to help her students accomplish the goal of discovering the truths.

Melody recognized the importance of logical thinking in her instruction; therefore, she invested significant effort in logically sequencing her instruction while planning her lessons. When I asked Melody to tell me how she went about designing her lessons, she said “I try to think like, “Why? Why? Why? Like “Why is this true?” and try to answer that all the way through it. So I’m trying to give them the whole “Why is this true?” Because Melody believed learning science took place through logical reasoning, she encouraged her students to reach an answer through following a logical method.

Although Melody maintained that she taught through inquiry, inquiry was reduced to applying logic to reach an answer that matched her authoritarian knowledge. This suggests that Melody’s conceptions of inquiry did not encourage her students to seek authentic answers. However, the authenticity of answers to questions is a central
attribute of the scientific inquiry (NRC, 1996). It is my interpretation that although Melody thought highly of teaching science through an inquiry framework, she lacked practical knowledge of teaching science through an inquiry framework. Furthermore, Melody’s limited personal experiences with learning through inquiry also placed limitations in terms of her ability to teach science through an inquiry framework.

Research on student learning suggests that questioning plays a critical role in doing inquiry (Blanchard, 2006; Goldman & Wiley, 2002; NRC, 1996). Although questioning was a critical element of Melody’s instruction, Melody did not create opportunities for students to ask questions. When I asked Melody what she thought inquiry learning involved Melody said:

What I know inquiry learning to be — I think it is learning through questioning, which can be quite difficult for a teacher as it sounds like “just let the kids ask questions and go”, but I think it can be far more structured and work more on a basis of allowing students the right to ask questions at the beginning of a term or chapter and having more to say into the curriculum itself and what they want to learn and how they want to do it. Plus I think that having a question that students generate makes it easier to teach for me because it keeps a narrow focus. (Melody interview, March, 03, 2005)

This evidence shows that although Melody considered questioning as a form of inquiry, she used questions to shape her students’ thinking in order to generate answers that matched her expert knowledge. Melody did not believe that it would be a realistic expectation to allow her students to ask questions and find answers to those questions on their own. She expressed her frustration with open-ended inquiry by saying, “Any time I forget and give them a purpose, they get fed up, frustrated, and usually quit. I think students are used to Q/A and can think in that direction with more clarity” (Melody interview, March 03, 2005)

This evidence suggests that students depended on the teacher and expected her to define the purpose of the learning activity for them rather than feeling the freedom to search for meaning on their own terms. Furthermore, this evidence indicates that the dominant culture of learning at the school challenged Melody’s intentions to teach science through open inquiry. For instance, the structures, (e.g., bell schedule, the load of teaching, limited time for laboratory preparation) guiding teacher practice in schools are not conducive for teaching science through an inquiry framework.
Melody also viewed teaching through discussion as a form of inquiry. She maintained a position for the superiority of argument as a form of instruction over the traditional didactic instruction that focuses on the transmission of information. Melody said, “You know who has it and who does not when you make them talk --you see different answers to the same question. Also, that way, you can tell if the student is misunderstanding you” (Melody interview, March 03, 2006). Melody stated that teaching through argumentation is powerful in the sense that it makes students’ misconceptions visible, their reasoning explicit and it exposes them to diverse views on the same phenomena. Kuhn (1991) maintains that the use of arguments is central to students’ critical thinking ability. However, instead of creating opportunities for students to formulate an argument and defend that argument with evidence Melody used discussion as a means to shape students’ thinking.

In addition to questioning and discussion, Melody viewed hands-on activities as inquiry. However, my observations of Melody’s teaching suggest that students’ curiosities and questions did not appear to be the driving force in her hands-on activities. Melody described the following hands-on activity, which she thought reflected the attributes of inquiry.

I have a conservation of mass project- um, I have a bunch of little labs that deal with the conservation of mass- and it was a pretty easy concept- So I figured- ‘let me--you know I didn’t want to give lab after lab after lab. So what I did was I put like five little labs together. Really, little, like, ice melting, salt dissolving. And then they had to pick 3 out of 5. And there was a write up and stuff- I can give you a copy if you want it. They had to pick 3 out of 5 and one had to be a chemical change and one had to be a physical change out of the three. And they had to write their whole lab report, from start to finish, and they had to write their own procedures. So we went through and taught them of conservation of mass first- and they had to write their own procedures, and their eyes just fell out of their heads. You know “What do you mean we have to write out own procedures??” And finally by the time they--But they had to get them approved by me before touching anything. So they had to like write it, and then I wanted it on top of my desk, and it got to a point where I used a timer and I would give them two minutes and then they had to sit back down and try it again. If I couldn’t understand what they meant in two minutes, then they didn’t understand what they were writing. So they had to write the procedure, get it approved, then do the experiment, collect their data, then type it up. So it probably took us a week to work on that. And I liked it because they had to really think. If they really understood the first procedure, the other two were easy. But a lot of them just had a really such a hard time getting started that they would freak out for like a day
and a half. So they would waste a day and a half like, “I don’t know how to do this, I can’t figure out how to do this.” And then they finally keep going and going and they are like, “Well, this is kind of easy.” (Melody interview, May, 03, 2005)

Although the laboratory activity encouraged students to actively write the procedures, design investigations, ask new questions and seek solutions working together with peers, Melody constantly guided their thinking and provided directions as to what they needed to do. Melody’s understanding of inquiry appears to be influenced by efficiency model of teaching. For instance, she timed students’ thinking rather than allowing them enough time to think thoroughly.

My observation of the activity indicates that students of this school are not accustomed to learning through inquiry. These students have been in a school system that discourages them from asking questions; instead, mandating that they listen to the lectures and formulate an answer that matches the knowledge presented to them through lecture notes when challenged on the tests. Because students were not accustomed to learning through inquiry, they were confused and did not know their responsibilities as to what they needed to do. Although the culture mediated against it, Melody’s persistent attitude provided conditions for students to learn science through inquiry. Although this was a rare occasion in which Melody was able to reinforce the goals of science education reform, it is evidence that the teacher agency is able to dominate cultural structure and change the course of science instruction. However, Melody’s naïve understanding of inquiry, her limited practical knowledge of teaching science through an inquiry framework and her limited access to inquiry-based curriculum materials limited the opportunities for her students to learn science through an inquiry framework.

**Summary of Teachers’ Pedagogical Conceptions**

I provide a summary of my participants’ pedagogical conceptions; their beliefs on learning, teaching and inquiry in the following passages.

*Summary Teachers’ Beliefs on the Learner and Science Learning*

Science education literature reveals that examining teachers’ beliefs about the learners and science learning is crucial because it holds potential to inform us about the instructional strategies and assessment methods that teachers use to meet the learning
needs of students (Ballone & Czerniak, 2001; Black et al., 2003; Levittt, 2002; Tsai, 2002). In order to understand how teachers’ pedagogical conceptions reinforce or contest the goals of science education reform (AAAS, 1993; NRC, 1996), I focused on eliciting my participants’ beliefs on the learner and the learning process.

These teachers believed in their students’ intellectual ability to learn science, however, they claimed that students did not have necessary motivation to learn science. Although teachers blamed students for not showing enough interest in meeting academic objectives, the school curriculum and the pedagogical practices used to facilitate students’ learning produced students’ lack of interest. Teachers used an authoritarian approach (Davis, 2004), a punishment and reward approach to teaching science. This reduced students’ learning to task completion. The teachers rewarded those who worked hard enough to complete the tasks given to them and matched their learning to the teacher knowledge with high grades and punished those who failed to do so with low grades.

Evidence presented above suggests that we need to restructure curriculum goals and instructional practices to catch students’ interest, thus, increase students’ motivation so they can find meaning in their learning experiences. Such restructuring would make learning activities occasions in which students would develop conceptual understanding of important science content, understand the application of that knowledge in everyday life and search for meaning (AAAS, 1993; Gallagher, 2006; NAS, 2006; NRC, 1996).

In addition to a curriculum goals and pedagogical practices some structural factors outside of school also influenced how students went about their learning. For instance, all of my participants frequently correlated learning with hard work and defined hard work as a powerful skill that some students did not bring with them to the school because of the local environment in which they grew up and the type of family they had. The limitations that home life put on students’ learning presented fundamental challenges to the teachers who were trying to help them live up to the academic expectations of schooling. It follows that three factors influenced students’ approach to learning. These three factors include 1) curriculum goals, 2) teachers’ pedagogical conceptions, and 3) students’ socioeconomic background.
Summary of Teachers’ Conceptions of Science Teaching

The national science education reform documents (AAAS, 1993; NAS, 2006; NRC, 1996) and science education scholars (Brickhouse, 2006; DeBoer, 2002; Gallagher, 2006; Settlage & Meadows, 2002; Southerland & Hutner, in press) argue that the purposes of science education have changed with demands of a changing world. They maintain that the goal of science education has shifted from teaching for acquisition of prescribed knowledge to the teaching for conceptual understanding, teaching for applications of scientific concepts in real life context and teaching to foster students’ acquisition of fundamental skills of inquiry (AAAS, 2005; Gallagher, 2006; NRC, 1996; NAS, 2006; Southerland & Hutner, in press; Yager, 2005). However, ensuring students’ acquisition of science literacy goals requires a learning environment that fosters critical thinking, allows for student autonomy and encourages creativity rather than reliance on the teacher authority (Blanchard, 2006; Cook-Sather, 2002; Gallagher, 2006; Marshall, 2006; NRC, 1996; Tobin et al., 1996).

Teachers play a significant role in creating and maintaining such learning environments (Bianchini et al., 2003; Blanchard, 2006; Gallagher, 2006; Marshall, 2006; Moore, 2004; Shulman, 1986; Southerland et al., 2006). Furthermore, research reports how teachers design and facilitate learning is significantly influenced by their deeply held beliefs about teaching and learning (Brickhouse, 1990; Davis, 1996; Pajares, 1992; Sowell, Southerland & Blanchard, 2006). If teachers’ internally held conceptions play a significant role in what students learn, to foster students’ mastery of the new science literacy goals, teachers must adopt new pedagogical conceptions (NRC, 1996; NAS, 2006; Southerland & Hutner, in press).

The findings from the cases of these three teachers suggest that teachers held conflicting purposes for the teaching of science. Four factors; teachers’ beliefs about what education should be about; curriculum goals that the traditional school science curriculum advocates, teachers naïve pedagogical content knowledge (Gess-Newsome & Lederman, 1999; Shulman, 1986) and teachers’ perceived responsibilities of their contracts generated conflicting purposes. These factors combined with teachers’ lack of understanding of the purposes of science education reform led teachers to teach science in an authoritarian fashion.

Although all teachers held conceptions of science teaching that partially reinforced the goals of science education reform (AAAS, 1993; NRC, 1996), in practice
they focused on students’ acquisition of canonical knowledge. Teachers’ way of teaching science reflected the values and expectations of the wider society, which views schools as the institutions of certification for the job market (Marshall, 2006). These social values, expectations combined with local political pressures due to standardizing testing led these teachers to reduce the purpose of science teaching to the task completion and competition for social mobility (Labare, 2000). These pressures placed limitations on teachers’ ability to teach science through an inquiry framework. Also, because the politicians approached the implementation of reform from a top down approach (Brickhouse, 2006; Cuban, 1990), teachers were not taught how to teach reform objectives in their classrooms. Therefore, teachers had limited knowledge of reform objectives, thus, showed limited interest in teaching science to reinforce science education reform goals in their classrooms. Furthermore, because teachers were not held accountable for teaching science to reinforce reform objectives, they did not pay significant attention to using reform calls as a framework for adjusting their instructional goals and practices to meet students’ learning needs.

Additional factors also influenced the type of science literacy goals these teachers reinforced in their teaching. These factors involved concerns about classroom management, perceived time constraints, students’ perceived practical needs for passing the test and graduating from high school. For instance, although Melody firmly believed in teaching her students higher order thinking skills, she primarily focused on teaching her students study habits and helping them develop test taking skills to meet students’ practical needs-passing the class and performing well on the FCAT. Similarly, Kevin placed a greater emphasis on the learning of the content knowledge, not because what he believed science learning should be about, but because he thought that is what his students needed in order to pass college chemistry. In the same token, it was not Ryan’s beliefs about teaching and learning but far more practical needs such as students’ need to establish a strong foundation in chemistry content as a preparation for college chemistry that led him to reduce his teaching to content acquisition.

Summary of Teachers’ Conceptions of Inquiry

Prominent science education reform documents maintain that science learning in the 21st century, demands relatedness, ability to connect with the real life, opportunity to
do science, motivation to listen to and engage in discussions involving scientific issues (NAS, 1996; NRC, 1996; 2006). Science education reformists maintain that teaching science through an inquiry framework will prepare students for the 21st century citizenship (AAAS, 1993; Blanchard, 2006; Gallagher, 2006; NAS, 2006; NRC, 1996). There are numerous definitions of inquiry in science education literature. While some define inquiry as scientific skills of observation, controlling variables for experimentation, ability to hypothesize and think critically (Krajcik, Blumenfeld, Marx, & Soloway, 2000; Linn, Clark & Slotta, 2003; Songer, Lee & McDonald, 2002), others define inquiry as an open-ended process, which calls for developing well-reasoned arguments for explaining complex scientific and socio-scientific phenomena (Roth, 1995). Although inquiry has been defined differently by different scholars, a common goal across all inquiry-driven discussions is to engage students in higher order thinking and help them construct knowledge by doing science (Blanchard, 2006; Gallagher, 2006; NAS, 2006; NRC, 1996).

Although my participants recognized the viability of inquiry-based science teaching, evidence presented in the earlier sections of this chapter suggests that their teaching mostly took place in a traditional fashion where the teacher gave the students a worksheet and asked them to follow the instructions to conduct the laboratory experiments. My observations of my participants’ laboratory classes indicate that students’ learning in the laboratory activities involved:

1. making observations and measurements using different tools,
2. doing preplanned experiments to test existing scientific hypotheses and theories,
3. interpreting data using mathematical formulas and answering the questions on the back of laboratory worksheets that asked them to come up with correct answers.

Teachers’ naïve understandings of inquiry, insufficient practical knowledge of inquiry, their perceived time limitations and classroom management concerns constrained their motivation and ability to facilitate inquiry-oriented learning in their classes. However, the type of inquiry that the teachers wanted to promote in their teaching was guided, “Level 2” inquiry in which students are asked to investigate a question presented by the teacher through student-designed procedures (Herron, 1971). Even then teachers did not guide students properly and challenged them adequately for utilizing cognitive
processes needed to solve scientific problems in a guided inquiry framework. Evidence presented in the previous passages suggests that teachers failed to reinforce the type of inquiry that the science education reform documents promote. Science education reform documents such as NSES promote “Open Inquiry” or “Level 3” inquiry (Herron, 1971) in which students design their own procedures for answering a question that the students themselves formulate. If the purpose of science education is to create a democratic society and to promote democratic participation of citizens in controversial issues (Labare, 2000; Marshall, 2006) teachers should promote open inquiry in science classrooms. Although these teachers’ naïve understanding of inquiry and their professional knowledge of inquiry limited their ability to teach science through an open inquiry framework, structural factors such as short class periods also influenced how the teachers approached inquiry. For instance, teachers did not believe that they could teach science through an open-ended inquiry or “Level 3” inquiry (Herron, 1971) framework in a 45-minute classroom period.

These findings are consistent with a recent report published by the National Research Council, which documents inadequate standards of laboratory teaching in American high schools (NRC, 2005). The outdated equipment, poor integration of laboratory activities into the rest of the curriculum, unclear goals, and teachers’ lack of practical knowledge of doing laboratory-based teaching are the reported factors lowering the quality of laboratory-based teaching in the U.S. high schools (NRC, 2005). The findings from this report suggest that students’ poor laboratory experiences not only influence their current learning but also influence their academic learning in college.

In a recent article published in the Journal of Research in Science Teaching, Tai, Sadler and Loehr (2005) look at the correlation between undergraduate freshman students’ performance in an introductory chemistry class and their learning experiences in high school using multiple regression analysis. They report a significant correlation between students’ performance and their experience in chemistry laboratory activities at high school. They find that students who went through laboratory experiences that heavily focused on laboratory procedures performed poorly in chemistry. On the other hand, students who experienced a laboratory learning experience that focused on developing inquiry skills and understanding performed well on learning objectives in the
introductory chemistry class. Such evidence clearly indicates that high school learning experiences not only affect students’ current understanding of science but also their future performance in science classes. Therefore, policy makers and science educators must pay special attention to understanding the ways in which science teachers view scientific inquiry and scaffold inquiry-based learning activities. Furthermore, this observation is a call for science educators and curriculum developers to cooperate on developing laboratory materials that would enable teachers to teach science through an inquiry framework. Although improvements to curriculum goals and learning materials are necessary, teachers are unlikely to teach science through an inquiry framework without a sophisticated pedagogical content knowledge base (Gess-Newsome & Lederman, 1999; Shulman, 1986).

These findings lead us to conclude that three fundamental factors limited teachers’ ability to teach science through inquiry. First, teachers did not have adequate preparation for teaching through an inquiry framework. Thus, their understanding of teaching science through inquiry was limited to structuring teaching to engage students in logical thinking. Secondly, the cultural norms and expectations about teaching school science forced these teachers to exclusively limit their teaching to content coverage. For instance, because teachers believed that teaching through inquiry threatened their commitment to content coverage, they refrained from teaching science through an inquiry framework. Finally, teachers held naïve understanding of inquiry and lacked pedagogical knowledge base to teach science through an inquiry framework. This in turn limited students’ learning to the acquisition of discrete bits of scientific facts and laws rather than allowing them to develop skills of inquiry and search for authenticity and meaning in their learning of science.

Science educators can build on this observation and search for ways to help science teachers teach laboratories in an inquiry framework. This can be a challenging task for science educators as pre-service teachers often meet their science requirements in science departments in which science is still taught in traditional fashion (Aydeniz & Gilmer, 2003). Teachers’ college science learning experiences may guide teachers to teach science laboratories in a traditional fashion as their learning in the laboratories are often limited to following directions to conduct experiments (Aydeniz & Gilmer, 2003).
This suggests that model inquiry-based laboratories must be designed and taught in science education programs as well as in science departments. This finding can be considered as a call for science education departments and science departments to collaborate on transforming the nature of laboratory activities both at the high school level and at the college level and structure them in a way so the laboratory-based learning activities reinforce the fundamental goals of science education reform documents (AAAS, 1993; NRC, 1996).

Characterizing Science Teachers’ Pedagogical Conceptions

As I discussed at the beginning of my findings, I use Tsai’s (2002) model to look at the nested relationship between various constructs that make up teachers’ conceptual ecology. Tsai (2002) established three categories for describing teachers’ pedagogical sophistication; traditional, process and constructivist. In the following tables I describe how my participants’ pedagogical conceptions compare to Tsai’s (2002) model.

Table 6. Teachers’ Essential Pedagogical Conceptions.

<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptors</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science teachers’ beliefs of teaching science</td>
<td>Traditional</td>
<td>Science is best taught by transferring knowledge from teacher to students.</td>
<td>Transferring of knowledge; giving firm answers; providing clear definition; giving accurate explanations; practicing tutorial problems; presenting the scientific truths or facts.</td>
</tr>
<tr>
<td></td>
<td>Process</td>
<td>Science is best taught by focusing on the processes of science or problem-solving procedures.</td>
<td>Teaching the scientific method; following problem-solving procedures; experiencing the processes of (self) discovery; working on the processes of verification.</td>
</tr>
<tr>
<td></td>
<td>Constructivist</td>
<td>Science is best taught by helping students construct knowledge.</td>
<td>Helping students make interpretations; providing authentic experiences; interacting with students; encouraging discussion and cooperative learning; paying attentions to students’ prior knowledge or misconceptions.</td>
</tr>
</tbody>
</table>
Table 6- continued

<table>
<thead>
<tr>
<th>Science teachers’ beliefs on learning of science</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traditional</strong></td>
</tr>
<tr>
<td>Learning science is acquiring or ‘reproducing’ knowledge from credible sources.</td>
</tr>
<tr>
<td>Transferring of knowledge; memorizing formula, definition, keywords and scientific facts; copying what teachers do; hard work on practicing tutorial problems; passive listening; finding the right answer; accurate calculation.</td>
</tr>
<tr>
<td>Kevin</td>
</tr>
</tbody>
</table>

| **Process**                                    |
| Learning science is focusing on the processes of science or problem-solving procedures. |
| Understanding the scientific method; following problem-solving procedures; learning through the processes of (self) discovery; working on the processes of verification. |
| Ryan                                             |
| Melody                                          |

| **Constructivist**                              |
| Learning science is constructing personal understanding. |
| Making interpretations; exploring or coping with authentic experiences; discussing with peers and teacher; relating to prior knowledge or (personal or daily) experiences. |
|                                              |

Adopted from (Tsai, 2002, p.774).

Evidence from this multiple case study indicates the limitations of Tsai’s (2002) model. Tsai’s (2002) has established discrete categorizations, which fail to capture the complexity of teachers’ internally held conceptions. For instance, while I placed Ryan in the process category in terms of his pedagogical sophistication, his pedagogical conceptions reflected some attributes from the traditional category. Similarly, although I placed Melody in the process category, Melody’s pedagogical conceptions reflected some attributes of constructivist pedagogy. However, we can expand Tsai’s (2002) model to have two levels; naïve and sophisticated, in each category to better accommodate the complexity of teachers’ internally held conceptions.

Tsai (2002) created his categorization based on the desired epistemological and pedagogical attributes needed to support science literacy goals that the prominent science education reform documents such as *NSES* (NRC, 1996) promote. Tsai (2002) is not alone in his attempt to deliniate the characteristics of a “quality teacher”. One can find thousands of articles attempting to describe the quality of the good teacher across
different fields and grade levels. However, these generalizations have not been able to predict teachers’ behavior in class as the issue is complex and involves multiple factors on which teachers have limited control. Therefore, in spite of countless research on the subject over a century, a great number of teachers still teach science in a traditional fashion and the impact of teachers’ categorization on students’ learning has been limited. However, we can build on Tsai’s (2002) model and establish a model that will enable us to better understand the connections between teachers’ epistemological, pedagogical and assessment conceptions. I discuss the details of such model in my discussion chapter.
CHAPTER SIX

SCIENCE TEACHERS’ CONCEPTIONS OF ASSESSMENT

I start this chapter by introducing my findings pertaining to my participants’ conceptions of assessment. More precisely, I present evidence on science teachers’ purposes for assessment, their conceptions of grading and feedback. In my argument, I discuss how my participants’ conceptions of assessment reflect or conflict with the fundamental attributes of social constructivist accounts of assessment. Then, I elaborate on how well my participants’ conceptions of assessment reinforce the goals of science education reform. I end this chapter by classifying my participants’ conceptions of assessment based on their consistency with social constructivist understanding of assessment.

Kevin’s Conceptions of Assessment

Kevin taught standard level chemistry classes. Kevin’s students were in the 10th and 11th grade Heath Academy and were standard level students.

Purposes

Kevin’s stated purposes of assessment focused on three objectives; a) students’ acquisition of canonical knowledge, b) students’ understanding of the real life applications of such knowledge, and c) students’ higher order thinking skills. Kevin said:

Assessment is for — to see how much they have learned — I mean — there is lots of ways you can assess — you can just ask a verbal question umm — you can get them to do a simple problem and you can get them to a higher thinking problem — either you can get them [to] speak back to you exactly what was given to them — or get them apply it to something — so they can learn to apply what they learned to different situations. They need to adjust what they have learned to a new situation — in order to see how well they have learned or what you have taught them. (Kevin interview, March, 12, 2005)

This statement indicates that Kevin’s thinking was in line with the reform suggestions (NRC, 2001). For instance, Kevin was cognizant that the learning construct (e.g., higher order thinking, applications in new context) must be defined before one could assess students’ learning. Furthermore, Kevin wanted his students to apply their
learning to different situations. However, because Kevin viewed teaching of science as
the transmission of expert knowledge to the students, his assessment practices were
limited and did not contribute to students’ acquisition of science literacy goals that
science education reform documents advocate (AAAS, 1993; NRC, 1996).

Kevin maintained that assessment was a powerful tool, helping teachers to
understand not only how much students have learned, but also providing information on
how much students have not learned. However, Kevin reduced assessment to grading.
While initially Kevin provided an answer that appears congruent with reforms, as the
interview unfolded the image of what he actually did was different:

…you don’t know how well they have learned until you go through and grade it
[their tests, homework assignments and labs] and find out “ohh they got an A they
must have learned really well or they got a C then they might have confused some
stuff.” You have to get some sort of feedback from them — so assessment is a
means that helps you to find out that. The grading itself is a number. I guess it is a
scale to see how well they have assessed and learned something. (Kevin
interview, March, 12, 2005)

Although Kevin used assessment as a means to identify his students’ strengths
and weaknesses in terms of their learning of science, the type of knowledge and skills he
assessed did not reinforce science literacy goals advocated by science education reform
documents. Kevin’s assessment measured the degree to which his students acquired the
canonical knowledge of scientific facts, procedures and formulas. An example of Kevin’s
test question would be, “A compound whose atoms are held together by covalent bonds
is___________________________” To ensure students’ acquisition of scientific
vocabulary and facts, Kevin frequently used puzzles and vocabulary quizzes as part of his
assessments. In addition to puzzles, Kevin made his students memorize the names,
symbols and the atomic number of elements.

It follows that Kevin’s assessment did not ensure students’ conceptual
understanding and their acquisition of skills of scientific inquiry, rather it focused on
students’ recall of scientific facts and statements (NRC, 1996).

Grading

Kevin established a strong correlation between his grades and his students’
learning of chemistry. Kevin firmly believed that his grades were valid measures of how
much his students knew about chemistry. When I asked Kevin to tell me how much information his grades told him about his students’ learning of chemistry, he said:

A fair amount — I think I might be a little bit easy on them and I need to learn how to be a little harder on them — at times I think I still see students struggling beyond and failing — I guess — I can’t be too easy or else other students will probably pick on it too and that of course causes a lot of problems with managing class — with the years of experience, I have kind of fixed and adjusted my problems a little bit — umm — and trying to push them and maybe a little more and maybe give them a nudge there — I guess the grades seem fair enough to me — and I accept them for what they are. Like — if they have As they must have learned quite a lot — and if they have Fs they must not have — I guess it [the letter grades that he gave his students] is a fair assessment of what they know. (Kevin interview, March, 12, 2005)

Kevin used grades as a feedback mechanism for checking on his students’ knowledge of the fundamental chemistry concepts that he taught in his classroom. Kevin also used grades to motivate his students to invest increased effort for meeting his learning expectations. It follows that Kevin used grades as a punishment and reward mechanism to facilitate students’ learning rather than as a means to engage them in critical thinking and a search for meaning. Kevin held a view of learning that is consistent with behaviorist pedagogy in his assessments. Kevin expected his students to match their learning with his expert knowledge or with the information that could be found in the textbook rather than helping students develop scientific habits of mind in his assessment.

**Feedback**

Kevin took his responsibility to provide feedback to students very seriously. However, Kevin gave his students feedback on their learning primarily through grades. Kevin stayed after school for an extended period of time to grade his students’ tests, homework assignments and quizzes. Kevin made sure that he took points off on each question that they his students did not get right. He thought that his letter grades were worth the effort that he put in his grading; thus, he was convinced that his grades adequately showed how much his students had learned or what they had not learned. Kevin said, “If they got an A they must have learned really well or they got a C then they might have confused some stuff.” (Kevin interview, March, 12, 2005)

Although Kevin primarily provided feedback to his students through grades, he
used alternative methods as well to assess the effectiveness of his instruction and the 
learning level of his students. Kevin said:

… in talking to them and asking them questions as I am trying to explain things to 
them I can assess how well they have learned it and to make sure that I have 
prepared them well enough for a graded assessment. I will ask questions to see 
where they are at, right before a test. (Kevin interview, March, 12, 2005)

This quote indicates that Kevin reduced the purpose of science teaching to test 
preparation and of assessment to grading. Furthermore, Kevin viewed his role as an 
authority that judged students’ learning rather than someone who assisted students to 
learn continiously.

Although Kevin used a variety of different assessment methods, he said that the 
oral assessment methods provided him with more valuable information about his 
students’ academic progress. Kevin said:

…the oral ones tell me more than the traditional methods of assessment when it 
comes to students’ overall progress. I guess oral assessments work better for me 
and for fair number of students — as I am trying to gear them towards where I 
want them to be. Most of the time, I will get their letter grades through the 
traditional methods of assessment — because that is what motivates them to learn 
— but I want to make sure through oral methods of assessment — I kind of guide 
them through where I want them to be. (Kevin interview, March, 12, 2005)

Kevin mostly used oral assessments to identify his students’ weaknesses and 
strengths and provide guidance in terms of what they needed to do for passing the test. 
Consistent with his authoritarian view, Kevin used grades to motivate his students to 
complete the learning tasks. For instance, Kevin used oral assessment methods, and 
quizzes to motivate his students to meet the academic expectations that he set for his 
students.

Although Kevin claimed the primary purpose of grading was to motivate his 
students to learn science, grading failed to influence Kevin’s students’ motivation to learn 
science. For instance, Kevin said:

I say to each class at the beginning of the year. You can get an A in this course as 
long as you turn in work on time, listen and follow directions, and stuff like that. 
I’d like to expect all of them to complete all of my work, turn it in on time, and 
get A’s in the class, but that’s my expectations, to try to get all of them to do that, 
they won’t all do that. Not everyone turns their assignments in on time, not 
everyone understands each assignment. I mean, I’d like to expect to everyone to
get an A. But I understand that it is not going to happen. No matter what I do, some students are just not motivated to push themselves as far as they might need to. (Kevin Interview, May 6, 2006)

Kevin failed to engage his students in learning partly because Kevin focused on assessing students’ recall knowledge in his assessments rather than using assessment as a feedback mechanism to help his students acquire critical thinking and inquiry skills, and to guide them to search for meaning in their learning.

Ryan’s Conceptions of Assessment

Ryan taught the IB Chemistry classes. Ryan primarily taught 10th grade students.

*Purpose*

Ryan used assessment for two fundamental purposes. First, Ryan used assessment as a means to measure his students’ acquisition of the learning goals that he held for his students. Secondly, Ryan used assessment to motivate his students to do the work necessary to learn science. Ryan said:

Assessment is for — seeing if they actually picked it up. I mean, if you don’t assess things students won’t learn it. I mean a) there has to be some motive [for] assessment or they won’t learn it or b) you can never understand if they learned it. So, how are you going to adjust your teaching in the future to see if they did learn it? I mean one of the best functions of assessment is just learning what you did not teach properly so you can go back and fix that. (Ryan interview, February, 1, 2005)

Ryan presented an authoritarian view of assessment, in which, he focused on “fixing” what needs to be taught by the teacher rather than focusing on creating opportunities for his students to reflect on their learning, identify what they don’t know and what they need to do to improve their learning. Ryan’s view of assessment was consistent with behaviorist conceptions of stimulus and response as he used assessment to motivate his students to complete the learning tasks rather than using assessment to engage his students in a search for meaning. It follows that Ryan’s understanding of assessment did not reinforce students’ acquisition of science literacy goals such as developing a scientific habit of mind and conceptual understanding of fundamental chemistry concepts (AAAS, 1993; NRC, 1996).
Grading

Ryan primarily taught the IB students. Therefore, Ryan’s case provides a special context for the interpretation of grading and the purposes that the grades serve. As the IB programs selects the most competitive students (NRC, 2002), the majority of Ryan’s students strived for high grades. Because high grades served as reward for hard working students and gave them access to promising futures in terms of college entrance, the IB students showed strong commitment for receiving high grades. However, Ryan maintained his students’ focus on grades influenced the way he viewed assessment. Ryan said:

A lot of times people call them [grades] necessary evil. I mean if you don’t grade it [student work] students will not learn it. At least especially in IB that IB students can have the attitude of — umm — there is a joke about pre-med students that I always say cross the road because they are required to do it. I mean IB students have the same grade thirsty attitude. If you don’t have to do it they are not going to do it — but they will do it if you make them do it — If you say there is a grade for this, wow! — then all of sudden it pops up and they will learn — and sometimes it is frustrating especially in projects where you are trying to teach about the scientific ethics and that is very hard to give a grade for — and they will always come and ask “Did I do a good job on this?” I am like, “Well did you learn something from it?” and it is hard for them to kind of grasp that concept but I mean grades — I mean you do not necessarily want to give a grade but it definitely motivates IB students greatly — to do the work. (Ryan interview, February, 1, 2005)

Although Ryan did not reduce assessment of students’ learning to grading, he viewed grades as a powerful tool that he could use for motivating his students to learn science. This view of grading is consistent with behaviorist pedagogy in which instruction is viewed as a stimulus and grades as rewards for reinforcing positive behavior. While Ryan strongly believed in his students’ intellectual potential to work with challenging problems, he did not believe that he could engage students in learning activities without punishing or rewarding them with grades. Although Ryan blamed the competitive nature of the IB program and students’ “grade thirsty” attitude for focusing on grades in his assessments, Ryan emphasized grades because of the strong correlation that he established between effort and learning (Marshall, 2006). For instance, Ryan said “the students who study hard end up learning.” Although Ryan maintained that students with grade-thirsty attitude learned science, whether students developed a deep
understanding of chemistry concepts, thought about the applications of chemistry content in real life (Gallagher, 2006) and maintained such understanding over time is highly questionable. Ryan thought students with grade thirsty attitude learned science partly because he equated learning with acquisition of canonical knowledge and high grades. Also, Ryan placed significant emphasis on grades partly because grades served as an effective means for him to prove to the administration that he was meeting the goals of the program in the most effective manner. For instance, because a teachers’ effectiveness by the administration was measured by students’ passing rate on the IB exam, Ryan viewed grades as the most effective means to increase his students’ passing rate on the IB exam.

*Feedback*

Ryan used grades as a primary form of feedback for his students to ensure their success in his classroom. However, because Ryan focused on students’ acquisition of canonical knowledge in his assessments, the feedback that he gave to his students did not ensure students’ acquisition of knowledge and skills that the science education reform documents promote (AAAS, 1993; NRC, 1996). At the same time, Ryan’s answer indicated how he used feedback on the laboratories to direct student learning to the outcomes he desired to generate.

Um, for labs, they just don’t learn that until they see red pen, they don’t actually do it. I had them go back into their lab folder, and look at their flame test, and that was a good awful lab, they hate me. It should have taken me a few hours to do it, and it took me all Thursday and most of Fall Break. Then, going to one which they did two weeks ago, which I’m halfway done grading, which is observations again, a response to much like the responses on the flame test. And now they’re doing a beautiful job, and most are getting A’s. You can see the same thing from mathematical labs, from getting data and learning data. I always ask sources, and I finally see them understand it. (Ryan interview, April, 4, 2005)

Ryan gave his students directions so that their laboratory reports matched his expectations. Because Ryan provided feedback to his students by leaving comments and taking points off, feedback helped his students to better match his expectations on the subsequent laboratory activities. The feedback that Ryan gave his students did not encourage students to search for meaning, acquire inquiry skills or develop conceptual understanding. Rather, Ryan’s feedback encouraged his students to match their learning
with Ryan’s expert knowledge so they would not lose the benefits of receiving higher grades.

Ryan had a consistent grading policy. He took points off for every little mistake that his students made on his tests and laboratory reports. Ryan’s consistency in his grading informed his students about Ryan’s expectations. It is obvious that marking and taking points off had a symbolic meaning, thus, served as a powerful tool for executing the teacher authority and for motivating students to study hard. Instead of motivating students to develop a love for learning and engage them in a search for meaning, Ryan’s feedback encouraged his students to focus on grades. Grades worked as a motivational factor for students to study hard, partly because Ryan had students who cared about their current and future academic achievement.

Ryan’s focus on grading tells us that he used assessment primarily to measure students’ motivation and effort rather than their learning. This value of hard work was mirrored in Ryan as he put significant effort in giving his students feedback on their laboratory reports. Ryan said:

I try to make it [feedback] complete enough so they can understand better — especially in the labs — and it does take a lot of time — to write lab critiques good enough so they will understand — when I see the same mistake repeating itself — I will — just write a number or a circle on their paper — if you have a three on your paper you know what that means — but I really try to give them time in class and specifically encourage them to read what I put on their papers -and hopefully they will learn from that mistake and you do see if you concentrate on saying this is not how you do something — you will see it — not in the future labs. They know this is why the points are taken off — they are not going to do it in the future and actually doing that well is going to save you time in long run — because then you find the labs at this point you are in the fourth quarter much more better — so I was going through some labs last night, ohh! They show all their work like they are supposed to in the lab, they record all their data properly and report the procedures that they used, they did all the stuff right so. I give As and Bs just because they know how to do all these stuff correctly. (Ryan interview, March 2, 2005)

This evidence suggests that because Ryan encouraged students and gave them enough time to reflect on his feedback, students actually benefited from the feedback that he provided on their laboratory reports. Ryan reported evidence of growth in students’ learning of scientific investigation skills towards the end of semester, as documented in
their laboratory reports. Because Ryan’s understanding of inquiry reflected the attributes of “Level 1” or “structured” inquiry (Herron, 1971), Ryan thought his students developed better scientific inquiry skills as the semester progressed. However, in reality, Ryan engaged students in answering a question posed by him and expected his students to follow a set of procedures to match their learning to his expert knowledge or to the knowledge that could be found in the textbook. It follows that Ryan’s feedback promoted dependency on the teacher authority rather than engaging students in a search for meaning.

Although Ryan provided detailed feedback in his assessment of students’ laboratory reports, students’ reaction to his feedback on the tests and quizzes and the level of learning he thought this type of feedback generated discouraged him from providing detailed feedback on tests and quizzes. Ryan said:

Like I said I kind of try to judge well, if I give that amount of attention to it they are going to do better in the future — I guess on tests. I don’t give quite as detailed feedback. Because my assumption is that I have already went [sic] over it two or three times, what is the chance that they are going to get it the fourth time I go over it? And they might remember it next day, but they are not going to remember it the next week. I had some teachers go over the entire test, and you can’t remember what they did a week later because if the guy is just saying something to you that is not going to mean anything to you that much. If—especially to learn something, you have to go over the test in class and then go home and study it, if you want to learn, but that is not going to happen even with test corrections. For instance, one time they did not do well on the test, and I asked them to retake the test. My assumption was that they were going to perform better but no — it did not happen. They still performed lower in that section — there is a pattern to it, there is a structure a pattern to learning and you just can’t go over it once and expect to learn something complex that way and so I guess the reason why I give them feedback is I guess more for their curiosity only—almost. I mean you want to know why your points are taken off. There is a little satisfaction there, I guess. (Ryan interview, March, 2, 2005)

Because Ryan correlated students’ performance on the tests and quizzes with the amount of effort that they invested in studying the classroom material, he did not provide detailed feedback to his students on tests and quizzes. Two assumptions convinced Ryan that proving feedback to his students on tests did not contribute to his students’ further learning. First, Ryan believed that his students were intellectually capable of answering the questions on the test. Secondly, he felt that his students were familiar with the
concepts covered on the tests. Ryan’s assumption was that if students had studied hard enough for the test they would have passed. Thus, Ryan viewed students who performed low on the test as those who lacked motivation to invest necessary effort to study for the test. It follows that Ryan viewed assessment, thus, grades, as indicative of hard work rather than a reflection of students’ conceptual understanding of the scientific concepts.

Ryan provided feedback to his students partly to meet his accountability responsibilities against community members: administrators, students and the parents. Ryan said,

I had a parent conference the other day. It was pretty interesting actually. I mean —I had to go through the actual test and explain to her question by question why I took points off. It was ridiculous!!! —I mean they make you believe that they know a lot of chemistry, but the funny thing is, — they don’t. So it [taking points off] helps that way as well. (Ryan interview, March, 2, 2005)

The parents of Ryan’s students cared about their children’s education; therefore, they questioned Ryan for their children’s grades. Ryan felt that he needed to justify his grades. Ryan justified his grades by giving his students some feedback on their tests and laboratory reports through marking.

Also, Ryan considered providing feedback as an ethical issue rather than looking at feedback as a process that created appropriate conditions for students to continuously learn. For instance, Ryan said, “I guess the reason why I give them feedback is I guess more for their curiosity only —almost. I mean you want to know why your points are taken off.” (Ryan interview, March, 2, 2005)

Although Ryan was willing to provide sufficient time to provide feedback to his students, the majority of students did not show responsibility to reflect on their strengths and weaknesses as they went through the feedback process. The following quote indicates the struggle Ryan experienced in his attempt to help his students self evaluate their learning. Ryan said:

Well if I don’t give them time in class—a lot of them won’t [think critically]. Ok, I will tell them that I am going to give them a few minutes to go over it, and if they don’t then, I will say, “No, no, no.” Pull back, I guess they just read it. Can I make them critically think about it? I think a lot of them don’t think about it critically. At least, I can make them read it. It is not so much that they care about it but - they care more about how many points I take off especially with IB students. They are so obsessed with grades and they [could] care less about what I
put on their paper. They want to make sure that their points are not taken off. I
even sometimes admit sometimes kind of feel guilty for taking half a point off,
but I tell them, “If I don’t take half a point off you are not going to do better next
time.” I know this is not worth taking half a point off, but if I don’t and they will
kind of give me that, “Mr. H. I guess you are right,” which seems a little cruel but
you kind of do it and that is the lowest amount of point I can take off. If you don’t
they will look at it and kind of whisper, “oh he did not take points off.” (Ryan
interview, March, 2, 2005)

Ryan maintained that he wanted his students to reflect on his feedback in a critical
manner. However, because Ryan’s perception led him to believe that students lacked
interest in thinking critically, he limited his efforts to make critical thinking a central
element of his assessments. This evidence highlights the need for scaffolding the
feedback process in an effective manner. This may be achieved by holding students
accountable for engaging in the self assessment of their own learning. However, whether
the time limitations and teachers’ commitment to content coverage will allow that to
happen or not needs further inquiry.

Ryan maintained that establishing and enforcing a tough grading policy served as
a powerful feedback mechanism for ensuring students’ learning. For instance, Ryan
maintained that taking points off for the mistakes that his students made, helped his
students avoid making similar mistakes in subsequent tests and quizzes. However, what a
students learns as a result of a tough grading policy depends on the type of knowledge
and skills that a teacher promotes in his/her assessment (NRC, 2001). My analysis of
Ryan’s tests, homework assignments and his quizzes shows that Ryan measured a)
students’ factual knowledge; b) their mathematical problem solving skills, and c) their
conceptual understanding of the fundamental chemistry concepts (see Appendix C).
However, I must note that Ryan expected his students to match their learning to the
textbook knowledge instead of asking his students open-ended questions that engaged
them in meaning making.

Melody’s Conceptions of Assessment

Melody primarily taught physical science and earth space science to standard
level students.
**Purposes**

Melody’s understanding of assessment reflected some of the fundamental attributes of the reform-based assessment in which assessment is primarily used to identify students’ weaknesses and strengths, and for the purposes of improving the quality and effectiveness of science teaching (NAS, 2006; NRC, 2001; Shepard, 2000). Melody said the following about the purposes for which she used assessment, “I try to see where they are, what they know, and what I need to keep doing, what I need to change and improve — I mean — you have to do that. (Melody interview, February, 18, 2005)

Although Melody expressed beliefs that are consistent with the purposes of assessment advocated by reform documents (NRC, 2001), the learning goals that she promoted in her actual assessment practices failed to reinforce the science literacy goals that science education reform documents (AAAS, 1993; NAS, 2006; NRC, 1996) promote. Melody said:

I want to see do they know the basic concepts, do they know the basic concepts with vocabulary, do they know just that kernel of knowledge or do they know enough to actually be able to speak language to me as well and explain it back to me in a scientific language? (Melody interview, February, 18, 2005)

Melody’s assessment practices were limited to measuring how well her students knew the scientific vocabulary and how well they knew the scientific content to carry a conversation with her using the scientific language. Furthermore, instead of fostering students’ acquisition of inquiry and critical thinking skills and helping them develop conceptual understanding, Melody expected her students to memorize scientific vocabulary and statements in her assessments. However, Melody’s view of assessment was more sophisticated than either Ryan’s or Kevin’s because she used it both to assess students’ understanding and to assess her own teaching. Her intentions were to have the students using the language of science in dialogue “explain it back,” however her practices were limited to matching expert knowledge.

A reform-based understanding of assessment does not limit students’ learning to reproduction of the vernacular of the science content rather assessment tasks are designed to help students develop conceptual understanding, make connections between content taught in class and its real life applications using canonical science as a medium (NRC,
Melody used grades primarily for accountability purposes rather than viewing grades as an indication of students’ learning. Melody rejected establishing a direct link between students’ grades and their learning of science. Melody said:

I think the real purpose of grading is to have something tangible to show the people who are not in the room. Because I think when they leave my classroom, I really don’t believe that an A, B, C, D really says anything about what or how much a student has learned. I think it tells me more about how willing they are to do what I ask them to do, and it says more about how much they appreciate an ABCD or an F. In a sense I don’t think it really assesses their knowledge. I think it assesses how willing they are to play the game, and some kids learn a lo-o-o-t and are not willing to play the game, and they get Fs and they may have learned more than anybody else has in the room and vice versa. And some kids get A’s but may have not learned anything — because they are very careful about filling the blanks and care much about grades. I think a letter grade is really for their transcripts and information to their parents — and somehow to assess the school and the teacher. I think it is more about that really than it is about the student learning. (Melody interview, February, 18, 2005)

Melody viewed grades as an indication of the value that the students attached to their schooling experience, a measure of students’ motivation to do the necessary work for meeting the classroom expectations rather than as a measure of students’ learning of science. Melody believed that grades showed how well students “played the game” rather than what they learned. Melody maintained that grades quantified students’ learning, thus, met the information needs of the administrators and the parents. Also, Melody viewed grades as a form of currency that the students used for pursuing advanced academic goals and for job applications. For instance, she said, “I think a letter grade is really for their transcripts.” This view of grades as accountability represents a fairly sophisticated understanding of grading. However, Melody felt that she needed to document students’ academic progress through grades for meeting the perceived expectations of her contract. As a result of this perceived responsibility, Melody focused on meeting the accountability demands of the administrators and parents instead of focusing on creating opportunities for her students to improve their learning.
Feedback

Melody used feedback primarily as a system for making sure that her students fulfilled their responsibilities of completing the learning tasks. Melody also believed that her feedback taught her students study skills necessary for meeting her classroom expectations. Melody said:

I think the main information I get from my assessment is that it tells me about how willing my students are in trying to do things that I asked them to do. I feel — like in the long run — there is value to how good they are at following directions, and following the rules and willing to accept another person’s help. I guess also how to learn what needs to be done first, second and third — and how to go about accomplishing goals — when they accomplish an A, they have done all the things that I asked them to do in the order that I wanted them to and on time. (Melody interview, February, 18, 2005)

Although the purpose of Melody’s feedback involved helping students perform better on academic expectations, she designed the feedback to increase students’ effort to do the necessary work for passing the test. As Meloy set the learning goals for students, her assessments became a measure of students’ conformity to her authority. Thus, Melody’s feedback failed to engage students in a search for meaning and help them develop critical thinking and inquiry skills. Nevertheless, Melody firmly believed in providing feedback to her students. Melody said:

I do spend a lot of time writing to them, like in my general classes their homework is just a check mark. That is just like they do it, and I go over with them — I try to get them understand that that is just practice — When it comes to a test or a lab, I try to really write to them and discuss certain things with them. If they have done something really good on a paper, I will ask if they would mind reading it out loud to the whole class. It is kind of stupid, but I do a lot of smiley face stickers on their tests and that — I am surprised to see how many of them will actually put the stickers on their foreheads and faces, and they are really proud of what they have earned on a test — I try to pull them to a side and kind of discuss things with them and conference with them like throughout when other kids are working on a worksheet or a project. (Melody interview, February, 18, 2005)

While Melody used assessment as means to suggest further improvements for her students’ learning, Melody primarily focused on ensuring students’ conformity to behavioral expectations. For instance, Melody defined learning as students’ ability to follow directions to match their learning with her expert knowledge rather than a search for understanding the complexity of the natural phenomena.
Melody also used grades as a feedback mechanism to motivate her students to learn as well. For instance, Melody’s praised the best performing students for their completion of tasks by giving them sticky faces. Melody said:

I try to really write to them and discuss certain things with them. If they have done something really good on a paper, I will ask if they would mind reading it to the whole class. It is kind of stupid but I do a lot of smiley face stickers on their tests and that. I am surprised to see how many of them will actually put the stickers on their foreheads and faces and they are really proud of what they have earned on a test. I try to pull them to a side and kind of discuss things with them and conference with them like throughout when other kids are working on a worksheet or a project. So, by talking to them and discussing things with them you may be able to convince them to do something and hopefully to get them learn enough science to pass the class (Melody interview, February 18, 2005)

Melody maintained that by giving students progress reports every other week and by talking to them, she intended to guide her students to reflect on their progress and help them take necessary actions to pass her class. However, she did not think that progress reports made a significant change to the performance of her lower-end students. Melody was most concerned about her students’ passing or failing grades rather than about their learning of science. Consistent with this belief, Melody structured her feedback to motivate her students to follow directions necessary to match their learning with her expert knowledge and pass her class.

Summary of Teachers’ Purposes of Assessment

These teachers’ purposes of assessment indicates that they primarily focused on students’ acquisition of information in their assessments rather than fostering students’ ability to apply the acquired knowledge to the real life situations. Furthermore, these teachers’ conceptions of assessment failed to help students to develop scientific inquiry skills and help them view learning as a search for meaning (NAS, 2006; NRC, 2001). More precisely, assessment of students’ learning was limited to testing students’ obedience to the teachers’ authority and reflection of the knowledge that the teacher authority held or of the knowledge that could be found in the textbook (Marshall, 2006). Although Melody expressed relatively sophisticated conceptions of assessment, the political influences on her assessment limited the nature of her assessment practices to students’ acquisition of recall knowledge and scientific vocabulary.
Summary of Teachers’ Conceptions of Grading

Grades are central to the way students’ learning is measured in American school system across grade levels. Because grades are reported to and interpreted by different stakeholders, they serve different purposes (NRC, 2001). For instance, parents look at grades as a measure of their children’s success at school or their ticket to college. Administrators consider grades as a measure of students’ competency to take advanced classes and fulfill the graduation requirements. State officials look at grades as indicators of public schools’ performance on state standards (NRC, 2005). Furthermore, grades serve as means to determine students’ future college admission and future employment (NRC, 2001). Grades are used to make decisions about the effectiveness of a teacher as well. For instance, school administrators use grades to make judgments about teachers’ performance in the subject area that they teach, thus, make future employment decisions (NRC, 2001).

These teachers recognized grades as the outcome of an accountability system, which requested tangible evidence of students’ learning. Grades served a means for the administrators, parents and students to check on the teachers; making sure that teachers were accountable for what was occurring in their classrooms. Although this accountability system establishes a direct link between students’ grades and their learning, the utility of grades is often limited to administrative decisions about students’ future learning and their certification for graduation (Marshall, 2006).

If we look across these teachers we notice that each teacher valued the currency of grades, but for differing reasons. Melody said grades reflected the students’ ability to “play the game,” Kevin thought grades measured students’ knowledge. Ryan while recognizing the limit of grades to show learning, valued the motivational aspects of grades.

Although both Kevin and Ryan established a link between grades and students’ learning, grades frequently measured how well students performed on task completion, how motivated students were and how properly students followed the directions given by the teachers. Melody did not think grades served as a valid measure of students’ learning. However, even she placed significant emphasize on grading partly because of her perceived responsibility to document students’ learning in the form of grades.
Summary of Teachers’ Conceptions of Feedback

The feedback component of assessment can be described as a means for checking on the progress of the students throughout the learning cycle (Bell & Cowie, 2001; Brookhart, 2006; Shepard, 2000). Researchers who advocate for meaningful assessment practices establish a strong link between students’ learning and the quality and amount of feedback they receive on their learning (Black & William, 1998; Davis, et al., in press; Shepard, 2000; Wiggins, 1998). These scholars argue that feedback gives students the opportunity to learn from their mistakes and take the necessary actions to improve their learning, and gain a deep and rich understanding by continuously engaging in learning (Black & William, 1998; Davis, et al., in press; Marzano, Pickering & McTighe, 1993; NRC, 2001). It follows that teachers need to use assessment not only to see how well their students measure up to the learning expectations but also to help them continuously learn and search for meaning.

Although all participants recognized the importance of immediate feedback on students’ learning they lacked pedagogical knowledge and necessary commitment to challenge and engage students in the assessment of their own learning. Furthermore, the norms and expectation of school culture were counter to and did not support teachers’ willingness to provide extra class time for students to assess their own learning. Although the school culture had an influence on how teachers went about assessing students’ learning, teachers’ pedagogical beliefs also played a significant role in how they went about assessing students’ learning. For instance, Ryan and Kevin believed that their lectures generated more learning than students’ self assessment of their own learning partly because of their authoritarian pedagogical conceptions.

The cultural understanding of assessment common among students also limited the power of formative assessment in science classrooms across my participants. For instance, students failed to realize that assessment is a form of feedback on their learning which calls for further action. Instead, for students, assessment meant a number that indicated whether they failed or passed the test. It follows that the feedback that the teacher provides on students’ learning becomes functional only when students use the feedback as a guide for taking subsequent actions to substantiate their learning. However, teachers are responsible for creating a culture of assessment in which students value
continuous improvement and learning rather than focusing on getting the right answers and high grades (Davis, et al., in press). Teachers can create such culture of assessment by focusing on teaching science literacy goals that will make learning meaningful and engaging for students. In addition to change in learning goals, science teachers may be able to engage students in the assessment of their own learning by placing greater accountability on the students for learning.

Characterizing Science Teachers’ Conceptions of Assessment

As I discussed earlier, I used Tsai’s (2002) model to characterize my participants’ epistemic, pedagogical and assessment conceptions. The following table indicates how my participants’ conceptions of assessment fit into Tsai’s (2002) model.

Table 7. Teachers’ Essential Conceptions of Assessment.

<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptor</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>The purpose of assessment is to monitor students’ progress in order to meet the accountability purposes.</td>
<td>Teachers assess students’ work to record their academic progress for certification purposes; passing the class and grade reporting. Assessment of students is limited to acquisition of prescribed knowledge.</td>
<td>Kevin</td>
</tr>
<tr>
<td>Process</td>
<td>Although the sophisticated purposes of assessment are well understood, the purpose of assessment is to assist students to acquire knowledge and skills deemed important by curriculum demands.</td>
<td>Assessment is viewed as a vehicle to monitor students’ performance for acquiring minimum knowledge and skills required by SSS with limited reference to real world application of scientific knowledge.</td>
<td>Ryan, Melody</td>
</tr>
<tr>
<td>Constructivist</td>
<td>Assessment focuses on the improvement of both the instruction and students’ learning growth.</td>
<td>The purpose of assessment is to monitor students’ progress for the purposes of helping them identify their weaknesses and strengths. Assessment involves giving students feedback for the subsequent actions for their learning growth. Assessment forces students to make connections between the science content and its real life applications. The teacher uses assessment results to craft instruction; both for adjustments to the type of knowledge and skills promoted and the instructional methods used to facilitate students’ learning</td>
<td></td>
</tr>
</tbody>
</table>

Adopted from (Tsai, 2002, p.774).
If we analyze these teachers’ epistemic views of science and their pedagogical conceptions, we notice a strong link surfaces between teachers’ epistemic and pedagogical conceptions. For instance, Kevin, who held relatively naive epistemic views of science, was also the teacher who held traditional pedagogical conceptions. Ryan and Melody, who held more sophisticated epistemic views of science, also were the teachers who held more sophisticated pedagogical conceptions. However, when we analyze science teachers’ conceptions of assessment in light of their epistemic and pedagogical conceptions we cannot say that science teachers’ conceptions of assessment reflect their epistemic and pedagogical sophistication. The inconsistency between teachers’ epistemic and pedagogical conceptions, and their conceptions of assessment was generated by the political influences. For instance, although Melody expressed assessment conceptions indicating her entrance to the constructivist camp, the increasing focus on measuring students’ learning through standardized testing encouraged Melody to assess students’ knowledge acquisition rather than assessing students’ inquiry and critical thinking skills. Furthermore, because the administration demanded that teachers quantify and report students learning in the form of grading, teachers focused on measuring students’ acquisition of canonical knowledge rather than creating opportunities for students to develop critical thinking skills and helping them to search for meaning in their assessments. That is because measuring students’ canonical knowledge was a more efficient way of reporting students’ learning than measuring students’ critical thinking skills and their ability to do scientific inquiry.
CHAPTER SEVEN

SCIENCE TEACHERS’ ASSESSMENT PRACTICES

As society shifts from an industrial age to an information age (Hart, 2001) competencies that schools promote will need to change as well. Prominent science education reform documents maintain that science teaching should focus on fostering students’ critical thinking skills, ability to analyze data, making inferences, and communicating and working well with colloquies (AAAS, 1993; NAS, 2006; NRC, 1996). Reform documents advocate students’ acquisition of knowledge and skills that cannot be measured and ascertained through traditional forms of assessments. Thus, helping students develop skills of inquiry, ability to understand the applications of scientific content in real life requires changes not only in the way we view assessment but also in the forms of assessments that we use to monitor, diagnose, guide and interpret students’ learning (Brickhouse, 2006; Gallagher, 2006; NAS, 2006; NRC, 2001; Southerland & Hutner, in press).

Many new forms of assessments have been discussed and designed under the name of alternative forms of assessment to promote and ascertain students’ acquisition of the knowledge and skills that the science education reform documents promote (Brookhart, 2006; NRC, 2001). Performance-based assessments, essays, presentations, collaborative learning projects, portfolios and journal assessments are all examples of alternative forms of assessments in science learning (Delandshere, 2002). These are the assessment tools that teachers are encouraged to use for ascertaining students’ acquisition of inquiry skills, engaging them in a search for meaning making and helping low achieving students to perform well in science (NRC, 2001). In order to see how my participants’ assessment practices served those purposes, I present evidence on the types of assessments that my participants used in their teaching.

Kevin’s Assessment Practices

Kevin used two fundamental assessment techniques; traditional and alternative forms of assessments. Kevin’s traditional assessment practices included, vocabulary quizzes, puzzles to help students memorize the periodic table, multiple choice tests that he gave to his students at the end of each chapter. The alternative assessment methods
that Kevin used in his assessment involved group learning and projects. Kevin’s assessment practices that involved group learning included laboratory activities and grouping students to complete the traditional worksheets in class. Kevin constantly circulated around the room and helped students during the laboratory activities and when he made his students work together on completing worksheets. Although Kevin circulated around room to facilitate students’ learning, the learning objectives that he promoted in his assessment practices failed to reinforce science literacy goals that science education reform documents such as NSES promote (NRC, 1996). In addition to traditional assessment means, Kevin used projects to facilitate students’ learning of science.

Projects

Kevin assigned one major project for each quarter (a nine-week period), which required his students to write a three-page report and make a visually appealing poster. Kevin also required his students to do a presentation of their project in front of the class using PowerPoint. Kevin aimed to accomplish the following by assigning projects to his students. First, he wanted to teach his students how to search for relevant information for an assignment by using the library resources and the Internet. Secondly, he wanted to help his students to learn how to write an essay using the scientific language. In addition to learning to write an essay using the scientific knowledge, Kevin wanted his students to develop better communication skills. Kevin said “a lot of these kids are going to become nurses so they are going to talk to patients. So my objective is to teach them how to better communicate. So the presentation part kind of helps with that.” Kevin focused on teaching his students communication skills partly because he thought his students would need to develop effective communication skills in the health field which was the focus of the academy in which the students were enrolled.

Kevin incorporated projects in order to honor his students’ diverse learning skills. He said, “Because some students are more creative than they are verbal or intellectual; they’re more creative type. It kind of feeds into the other aspect that it helps out creative students.” (Kevin interview, May 6, 2005)

Findings from Kevin’s case suggests that although he held an understanding of assessment that matched the definition of assessment advocated by the prominent
national science education documents (AAAS, 1993; NRC, 2001), the outcomes that his assessment practices generated did not reinforce the science literacy goals that science education reform documents promote (AAAS, 1993; NRC, 1996). Although reform documents advocate the use of projects, Kevin’s use of projects did not reinforce the goals of science education reform. For instance, my observations of students’ presentations show that the projects did not engage students in critical thinking. Furthermore, the presentations did not encourage students to ask questions and did not engage students in thought provoking discussions.

Kevin said:

The projects just get them to learn about a variety of things that they wouldn’t normally be taught in the course. If they hear each other’s presentations in a short amount of time, they can just scratch the surface about the scientists of the past, or the different careers they might go into or learn about different elements and their uses in everyday lives. So it gives them kind of an idea about that subject. (Kevin interview, May, 06, 2005)

Kevin used projects primarily to help his students make connections between the science content he covered in his classroom and the applications of such content in real life context. Kevin also expected his students to explore various career opportunities in science by making them study different scientists of the past in their projects.

Although Kevin wanted to use projects more frequently in his teaching he did not think he had enough time to do that partly because of his commitment to content coverage. Kevin mentioned the challenges of designing, administering and evaluating project-based assessments in the following quote. He said:

They take a lot of time up. I spend a lot of time designing them. I have them do two days research in class, which two days isn’t taking up too much class time. But then, they have to present them. In the presentation, that eats up almost a whole week, to get through everybody’s presentation through the whole class, and, if they forget or something. It ends up taking a whole week to get through all the presentations. If I do too many of them, that’s eating up too much class time. And then grading it also takes a lot of time. You want to make sure that the student has the knowledge about the topic. You want to make sure that the student uses proper language and of course you want a poster appealing to eye. You want to make sure that the student has spent at least some time doing the poster not just a sloppy drawing. So that takes time, too. Whereas, I’m not discussing or going over too many concepts in the course, and how to deal with the math problems involved I lose time to do all those conceptual and mathematical practice things,
and then- it’s kind of a toss-up to how much I want to do. If I had extra time, I’d do more projects, they could learn more about the variety things that way. (Kevin interview, May, 03, 2005)

Kevin considered time to be a fundamental impediment to his motivation and efforts to use alternative forms of assessment frequently. Kevin also maintained that grading the alternative assessment tasks is hard because “you have to assess not only students’ knowledge of chemistry but also their creativity, their language skills and their organization skills.” Time became issue for Kevin partly because of his commitment to teaching for content. Although Kevin believed in educational value of alternative forms of assessments, his commitment to content coverage did not leave him time to design alternative assessment tasks and use alternative assessments to facilitate students’ learning.

Although Kevin showed motivation to use projects for assessing students’ learning of science, he did not think his students learned a great deal from the presentations. He said, “I think they learn more from the lectures than they do from the presentations” (Kevin interview, May, 6, 2005). Kevin did not think his students learn from doing projects partly because Kevin defined learning as the transmission of prescribed knowledge to his students. He said, “they don’t end up learning a lot of the stuff from each other’s presentations. They learn a lot from their own, but if they listened more, that’d probably be a better way but unfortunately that does not happen” (Kevin interview, May, 6, 2005). This evidence indicates that Kevin strongly believed in teaching through telling, thus, he expected his students to listen to one another’s presentations and learn from each other.

Although his expectations for students’ learning of science remained high, projects did not significantly contribute to students’ learning. That is because he expected his students to learn by listening rather than structuring students’ presentations to afford opportunities for students to reflect on, negotiate and construct knowledge in a collaborative manner. It follows that the way Kevin facilitated presentations reflects the attributes of an authoritarian pedagogy.

Ryan’s Assessment Practices

Ryan primarily relied on traditional assessment methods to facilitate his students’
learning of science. Ryan frequently used quizzes, tests and homework assignment to assess his students’ learning of science in his classroom. However, Ryan’s test questions were more sophisticated than Kevin’s test questions. For instance, although Kevin primarily tested students’ acquisition of scientific facts and laws on his tests, Ryan’s test questions frequently required students to apply the scientific laws to unfamiliar problems and make inferences. The following two questions on the same topic illustrate the difference between Kevin and Ryan’s test questions. However, I must note not all of Ryan’s test questions were as critical as this question.

Table 8. Sample Test Questions.

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Sample question on the structure of the atom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kevin</td>
<td>____T___F Electron is the negatively charged particle</td>
</tr>
<tr>
<td>Ryan</td>
<td>When a scientist passes a beam of particles by a negatively (-) charged plate the particles travel away from the plate. The beam of particles is likely to be a. protons  b. electrons  c. neutrons</td>
</tr>
</tbody>
</table>

Although Ryan asked questions that required his students to think more deeply, he expected his students to match their learning with the textbook knowledge rather than encouraging his students to search for meaning and develop understanding of important scientific concepts in his assessments.

Ryan viewed group assignments and laboratory activities as alternative forms of assessments that he used in his instruction. Although Ryan frequently used laboratory-based activities in his teaching, he rarely incorporated group projects in his teaching.

*Group-based Learning Activities*

Ryan’s assessment practices that required students to learn in a group setting included laboratory activities and on rare occasions small group projects that focused on the teaching of socio-scientific issues (e.g., the relationship between energy and the environment). Ryan’s understanding of assessment in group settings was heavily influenced by his perception of accountability. Ryan expected each student to submit a separate laboratory report for grading purposes although he knew that all students did the same experiment. Ryan was convinced of the intellectual ability of his students to solve
problems independently. Therefore, he encouraged his students to develop understanding on their own by investing time and effort in thinking about the problems. This effort and learning metaphor is consistent with Ryan’s authoritarian pedagogical conceptions.

Ryan occasionally incorporated activities that allowed his students to discuss scientific concepts in his class with their group members. Ryan said:

Then, I have some days I do want them to do group work. Umm — I wish I had more time for that, for doing different things like environmental problems that I had two weeks ago where they just sat down and thought about nitrogen and stuff like that in groups and answered questions as a group and thought about how beneficial they thought nitrogen was and so those days were obviously assessment days and I think they learned a lot about nitrogen. (Ryan interview, February, 01, 2005)

Although Ryan incorporated few group assignments into his instruction he acknowledged the limitations of his ability to incorporate group-based learning activities in his instruction. Ryan’s commitment to content coverage led him to believe that he did not have enough time for scaffolding his instruction to create collaborative learning opportunities for his students. Thus, Ryan’s perceived time constraints limited the opportunities for his students to learn science in a collaborative manner.

Ryan’s distrust in his students’ responsibility to work collaboratively also discouraged him from using collaborative learning as means to help his students to develop conceptual understanding of significant chemistry concepts. Ryan came to this conclusion partly because of his commitment to content coverage.

Although Ryan believed in the role of collaborative activities in learning, he wanted his students to develop independent learning skills as well. Ryan said:

I think you have to do both. Because group work--it depends on the question with group work--if the questions are just hard enough for them that they cannot get it on their own. But with just talking with each other they are able to get it. I think it is all about the idea of proximal development. I think that is the point. You want to make it so the questions are interesting enough so that they will engage students and to get them interested in talking to each other. Just enough so that they don’t need my help when they are doing group work, my help at all but yeah they need each other. But if you do it too much what always ends up happening is one person in group kind of takes over for everyone and so I think having them think independently is good at times so they don’t become to dependent on the person dominant in their groups. (Ryan interview, February, 01, 2005)

Although Ryan recognized the power of group work, he firmly believed in
helping students develop independent thinking and problem solving skills. Ryan recognized that group work meant designing appropriate learning tasks, utilizing an effective management strategy for coaching students’ metacognitive activities pertinent to their interactions in group settings. He noted, however, students’ level of responsibility challenged his ability to turn group activities into meaningful learning experiences for them. This discouraged him from doing group-based projects in a frequent fashion.

Ryan maintained that one person in the group usually dominated the discussion and the rest of group members became dependent on the dominant member. This evidence shows that community members failed to share responsibilities in a professional manner. Ryan’s view of students reflected his authoritarian epistemological perspective. He did not trust students to have the capacity or commitment to learning. The lack of responsibility among community members discouraged Ryan from using appropriate assessment tools. Ryan assigned independent assignments to make sure that all students learned to think, work and solve problems without depending on the students who studied hard and understood the chemistry concepts covered in his class more rapidly. However, he noted that it was not realistic to expect everyone to work equally hard. In the same token, he mentioned it was an unrealistic expectation to make collaborative learning the only means of instruction to ensure students’ learning.

Although Ryan believed in the contribution of group-based project assignments to student learning, his students’ challenged Ryan’s commitment by not working to their maximum potential when Ryan tried to teach through assigning collaborative projects.

Well, I think they [projects] are worthwhile the time invested in them but it is funny because students don’t think it is. Because to them it is just almost like a pass/fail grade — so trying to get them understand that there is a purpose on those things is very hard, for them it is whether you got 90 or 80% on the last test. To them that is far more important than understanding the actual process involved in doing it. (Ryan interview, February, 01, 2005)

This quote cited above suggests how cultural understanding of assessment influenced the way teachers made assessment decisions. Ryan continued to share his thoughts on this issue with me after the interview was over, and he mentioned that students’ perceptions of projects discouraged him from integrating more project-based assessments into his teaching. He also mentioned that students did not study to their
maximum potential on projects because they had a misunderstanding that projects were just a pass or fail type of assessments. Rather than attempting to address students’ misunderstanding through appropriate pedagogical strategies, Ryan decided to limit the use of alternative learning and assessment opportunities.

Students were accustomed to an assessment system, which consisted of tests and quizzes. Students’ reactions to the project-based assignments indicated that the the true meaning of assessment was embedded within a number grade rather than an artifact that they could have produced through their involvement in doing projects. For instance, Ryan said, although the nitrogen project encouraged students to think critically and encouraged them to use the scientific method and allowed them the opportunity to make connections to the real life, students did not think they learned a great deal by working on the nitrogen project. That is partly because the teacher did not change the assessment system to value the learning. This suggests that students viewed assessment as an opportunity to increase their grades and enjoy the further benefits that high grades promised rather than as an opportunity to learn something about chemistry.

Melody’s Assessment Practices

Although the traditional assessment forms dominated most of Melody’s assessment practices (quizzes, multiple choice tests, true/false and fill in the blank type of assessments), she recognized the significant contribution of alternative forms of assessment to her students’ learning of science. I describe the types of assessments that Melody used in her classroom in the following passages.

Discussion

Melody used discussion as an alternative form of assessment to ascertain the knowledge and skills that she thought were important for her students to learn. When I asked Melody to tell me the forms of assessments that she used in her teaching, she said:

I think the one that I like the most—that is I guess the most informal one that I like and I use and the hardest to put a value to, a grade to and to put in a grade book is discussion. But discussion to me is—it assesses motivation and determination as well as the knowledge. I mean if they are willing to raise their hands, willing to participate and really involved in the discussion—and try. Today like we were doing a discussion in fifth period. I have these kids—he started the discussion and then. Corbin jumped in and Andrew jumped in. I mean one kid
started the discussion, but other kids participated. All these kids were jumping in and talking about it. I was kind of running through the overhead and writing things down on the board. “Alvin you want to say something, too?” “Alan, do you have something to say about it?” Even if Alan got two words in—you know that means I have all these seven kids participating, and they are all discussing and answering the questions and asking questions seven out of 25 is a pretty good number for the first 20 minutes of the class. I think you know and we went through the answers together. Then we started to piece things together and trying to see if the things we pieced together make sense and to make sure that it is cohesive. I felt like we are doing pretty good. Now, whether we are doing good or not depends on how much they really learn and that is kind of hard to assess right away. If they can answer the questions, if they have the positive experience for the first ten minutes of the class then they are going to write things down. They feel like that they are going be pretty good at writing things down. (Melody interview, May, 03, 2005)

Although Melody recognized the power of discussion as a means to help her students substantiate their learning, she unconsciously promoted a culture of competition between students who were able to get their words out and those who lacked power to participate. Melody believed that discussion helped all students to learn, however, she ignored the fact that she was not able to reach students who thought quietly and lacked power to get their words out quickly. Discussion as an assessment method has significant implications especially for learning of ESOL students in science classrooms. For instance, ESOL students lack power and skills to participate in the discussions due to the limitations that their language proficiency places on their communication skills (Hamilton & Zimmerman, 2002; Warschauer, 1999).

In spite of all good intentions that she had for her students, Melody unconsciously rewarded assertiveness and recognized the assertive students as “the motivated ones, the ones who really want to learn.” Melody targeted some students as motivated and achievers and targeted some as not motivated ones. Therefore, her teaching targeted only the group of students who she labeled as the ones with motivation to learn.

Although my other participants also used discussion as a form of assessing student learning, they did not make explicit references to the use of discussion as a form of assessment. However, Melody’s discussions were well designed and more structured than the way in which my other participants used discussion to assess students’ learning in science. For instance, Kevin and Ryan randomly used discussions for short period of
time to guide students to make connections between the content of their teaching and the relevance of the content they taught to the real world.

**Group-Based Learning Activities**

Scaffolding instruction to provide opportunities for students to learn science in collaborative group settings is one of the core arguments of science education reform advocates (NAS, 2006; NRC, 1996). The proponents maintain that group activities enhance the learning of lower-achieving students and help those already achieving think more critically (Brown & Palinscar, 1989; Cramer, 1994; Felder & Brent, 2001; Lord, 1998; Johnson, Johnson & Smith, 1998; Gokhale, 1995; Kagan, 1990; Slavin, 1990).

The literature on group learning exists under two terms; cooperative learning and collaborative learning. Although these two terms are frequently used interchangeably, there are significant differences between them in terms of the processes involved and the outcomes generated by each of these forms of group-based activity (Kagan, 1990). Although they both have students work in groups, assign specific learning tasks, and have the members of group share and compare their procedures and conclusions with one another, there is a significant difference between them (Rockwood, 1995). While collaborative learning ties into the social constructivist movement, in which students assume responsibility, discuss and develop understanding on their own rather than depending on the teachers’ authority to develop understanding of concept deemed important in science (Bruffee, 1993), cooperative learning bears fundamental attributes of behaviorist pedagogy in which students follow the procedures given by the teacher to match their understandings to the canonical knowledge presented to them through the teacher authority (Kagan, 1990; Rockwood, 1995).

Melody frequently incorporated group-based activities in her instruction, however, whether the group activities that she designed enabled the participation of all students is highly questionable. My observations of Melody’s group-based learning activities indicate that Melody engaged her students in “structured” or “Level I” inquiry (Herron, 1971). Melody put students in groups and expected them to learn science together by following directions provided by her. Melody walked around the room and frequently checked on her students’ learning by asking them questions and answering the
questions that her students raised. Some students did the work for the whole group, and the rest of them just depended on those students for their participation. They all turned in separate papers with similar answers on them, and they all received the same grades as long as they turned in their papers on time and did not “cheat.” My observations suggest that Melody focused on individual accountability and timed her students for learning during group-based learning activities. Furthermore, although Melody encouraged collective learning, she made sure that her students matched their learning to her expert knowledge.

Although Melody considered group learning as an effective way for enhancing students’ learning, she was cognizant of the challenges as part of her experience with grouping students for learning purposes. Melody said:

You know when there is a group activity some kids will go ahead and do the whole work, and there are always some who will just sit down the whole period and copy things off of other students towards the end of class, just to get grades. If things come back to me word for word. First of all they always have to turn in a separate sheet. Everybody has to turn in their own assignment. I personally don’t really believe in grading two people on the same product. Because one kid may not do what they are supposed to do, you know, and then whatever so they all have to do their own thing. If it comes to me word-for-word it is a zero, if two kids write the same thing word-for-word it is zero. I don’t care if it comes from book I don’t care what it is, and I call it cheating and make a big deal out of it and call home and do whole nine yards. I mean even the lowest end kid can look at someone else’s paper and read it and even if they have to reword somebody else’s answer and really don’t get it. They are at least having to think about what word they could substitute, and there is no way I could be over their shoulder every minute and really know exactly what they were doing. I will at least take that but they can’t copy word-for-word. (Melody interview, May, 03, 2005)

Although Melody believed that group projects helped her students learn better, she acknowledged that students did not show a sense of responsibility to construct their own knowledge. Rather, they preferred to depend on one another to complete the learning assignments. Melody’s learning activities failed to engage students in learning in a meaningful way, so she turned to an old practice of holding students’ accountable for submitting separate papers for grading purposes. Although accountability drove Melody’s assessment practices in group activities, Melody viewed accountability as a motivational factor for helping students meet the requirements of passing her class rather than as a means to motivate students to search for meaning.
Teachers’ Use of Peer and Self Assessment

Self and peer assessment practices are important components of social constructivist pedagogy, which advocates that learners construct knowledge on the basis of their interaction with their social environments (Tobin, Tippins & Gallard, 1996). Because of their centrality to the social constructivist pedagogy and because of their promise to enhance students’ learning in science (NAS, 2006; NRC, 2001), I elaborate on self and peer assessments in the context of my study in the following passages.

Self Assessment

The latest research in human cognition and learning suggests that students frequently need self-assessment opportunities to reflect on what they have learned and determine what they need to do for further improvement (Bell & Cowie, 2001; NRC, 2001; Shepard, 2000). An analysis of research on new understanding of assessment also reveals that in addition to helping students consolidate knowledge, self-assessments encourage students to reflect on the learning process and help them identify the strategies of learning that would allow them to make realistic judgments about their academic performance and develop ability to set learning goals for their own improvement (Bell & Cowie, 2001; Davis et al., in press; Shepard, 2000).

Although teachers acknowledged the contribution of self assessment to students’ learning, incorporating self assessment as a way of improving student learning by passing on skills of evaluation and critical judgment to students was a challenge for my participants. For instance, teachers showed a tendency to blame students for not using self assessment as a learning tool, however, teachers themselves lacked practical knowledge to put self assessment into work for enhancing students’ learning of science. Furthermore, teachers thought of self assessment as a time consuming process mainly because of students’ lack of interest in self evaluating themselves in a meaningful way.

The challenge of incorporating self assessment into science instruction is two fold. First, self assessment is directly linked to metacognition. Metacognition in this context refers to students’ knowledge about, awareness of and control over their own thinking and learning (Perkins, 1998). Bransford and Donovan (2005) maintain that development of metacognitive skills is one of the three strategies for effective science
teaching. Evidence emerging from this study shows that students were not metacognitively engaged in their learning process, which in turn discouraged teachers from integrating self assessment as a means for improving students’ learning of science. Secondly, teachers’ concerns about classroom management and efficient delivery of content discouraged them from using self assessment as a tool for enhancing students’ learning in science. In addition to its perceived threat to the teacher authority, teachers considered self assessment as threat to their commitment to content coverage. It follows that although teachers held conceptions that recognized the value of self assessment to students’ learning, teachers’ naive pedagogical knowledge and the intentions and values of other community members challenged teachers’ ability to transform this understanding into classroom practice.

**Peer Assessment**

Peer assessment is a dynamic process which involves learners in assessing others’ learning, providing feedback to other learners; critiquing and making judgments on the quality of their work and helping them enhance their performance on meeting learning objectives (Davis et al., in press; NRC, 2001; Topping, 1998).

Although teachers believed in the contribution of peer assessment to students’ learning, peer assessment was not a common practice in my participants’ assessment of students’ learning. Although students were allowed to compare their answers with one another after tests and check whether teachers were consistent in their grading or not, none of my participants formally integrated peer assessment in their instruction. Teachers provided diverse reasons for not integrating peer assessment as a form of assessment in their teaching. For instance, Ryan provided the following four reasons. First, he did not believe that students’ had the ability and knowledge to evaluate each other’s learning. Secondly, he considered peer assessment as a time consuming process and limited his ability to cover content. Third, Ryan did not think that students are motivated and skillful enough to peer evaluate one another. Finally, he thought that students would not make fair judgments and often favor their friends. Although Ryan held some practical knowledge of peer assessment and recognized the contribution of peer assessment in student learning, he was concerned about losing his authority to his students, which in turn discouraged him from integrating peer assessment as a means for his students to
enhance their understandings of the fundamental chemical concepts. Ryan said:

One thing, peer assessment takes a lot of time. You got to teach them how to grade it, and I got to make sure that they are fair but that is not going to happen. I think where a nice place peer assessment can work is in writing. I don’t do enough writing though to make it worthwhile at all. I have come to conclusion that students are completely incapable of judging themselves and the quality of instruction. For the same reason I don’t think they will do a good job of grading each other. (Ryan interview, March, 02, 2005)

Ryan came to this conclusion based on his previous experience as an intern. He maintained that he did not believe that students were responsible enough and capable of peer assessing one another. Although Ryan was teaching advanced students, he felt that while some students know the content well to assess their peers, others lack enough knowledge to assess their peer’s learning. Ryan’s view of students as limited thinkers discouraged him from using peer assessment as means to enhance students’ learning in science.

When I asked Kevin if he used peer assessment as a means to help his students learn, he said, “I never thought of using it but I know what peer assessment is.” Kevin expressed similar concerns with Ryan but my conversations with him suggest that Kevin a) lacked practical knowledge of peer assessment b) did not believe in the educational contribution of peer assessment to his students’ learning. Kevin came to this conclusion partly because of his distrust in his students’ maturity and responsibility to fairly peer assess on another. Also, Kevin viewed peer assessment as a time consuming process compared with the amount of learning he considered took place through lecturing.

Although Melody had a basic understanding of what peer assessment meant and she expressed sympathy towards the contribution of peer assessment to her students’ learning of science, her students’ level of responsibility discouraged her from integrating peer assessment as a model of assessment to enhance her students’ learning of science. Melody said, “To be honest with you, I don’t think they can peer assess one another. You are talking about kids – kids that cannot read…” Melody maintained that she avoided using peer assessment partly because she did not trust her students’ intellectual ability to peer evaluate one another. Melody did not encourage peer assessment partly because she reduced the definition of learning to a match between students’ learning and her expert knowledge.
Three factors limited the integration of self-assessment and peer assessment as means to improve student learning in science. First, teachers did not consider students as individuals who are capable of making fair and informed judgments about the academic progress of themselves and that of their peers. Secondly, concerns about classroom management discouraged teachers from integrating self and peer assessment as means to enhance students’ learning of science. Finally, teachers’ concerns about the content coverage left teachers no time to integrate self and peer assessment to supplement students’ learning in science. It follows that although teachers expressed some of the educational contributions that peer and self assessments could have made to students’ learning in science, the structural limitations of classrooms inhibited what teachers thought was the right thing to do in terms of assessment. These influences include; teaching for content coverage and assessing students’ knowledge gains for accountability purposes.

In addition to interviewing teachers for understanding the culture of assessment in Southern High School, I analyzed the assessment tools that my participants used to assess students’ learning.

*Teaching and Assessment Tools*

Assessment tools are instrumental to maintaining a specific culture of assessment in science classrooms (NRC, 2001). Therefore, special attention must be paid to the nature of assessment tools used to assess students’ learning. I analyzed my participants’ instructional and assessments instruments. My analysis of teachers’ instruments, the classroom materials that my participants used to support their teaching, resulted in six different categories of assessment practices. I summarize my analysis of these instruments in the following table. There are two columns in the table. The first column describes the forms of assessment instruments that my participants used to support students’ learning in science. The second column provides a description of the nature of assessments that each instructional instrument reinforces.

| Table 9. Learning and Assessment Tools |
| Assessment Forms | Nature of Assessment |

141
Textbooks

**Standard Level Textbooks:**

**Content:** The textbooks that the teachers who taught the standard level classes had the following characteristics. They situated the content in the real life context, by providing numerous examples. These textbooks were full of images. At the end of each chapter, there was a section devoted to careers in science with particular attention to the encouragement of minorities and women to pursue careers in science.

**Assessment:** Assessment is a significant component of modern textbooks. The textbook encouraged teachers to use alternative assessment methods, however, the assessment questions targeted a) students’ acquisition of scientific vocabulary, b) their conceptual understanding of scientific laws and theories, c) their ability to compare, contrast and draw inferences from laboratory data.

**Honor Level Textbooks:**

**Content:** This textbook discussed the content with particular references to the everyday applications of content. The textbook used images and graphics to help students visualize abstract concepts.

**Assessment:** Assessment questions tested students’ scientific vocabulary, their ability to restate the scientific laws and theories, and finally measured students’ mathematical problem solving skills.

Problem-solving worksheets

Engaging students in problem solving through worksheets was a common practice across all participants. Some of these worksheets probed students’ understanding of the fundamental theories and laws of science, asking students to restate the scientific knowledge presented to them in class and some of them involved mathematical problems asking students to translate what they learned about the fundamental theories and laws into mathematical formulas. Most homework assignments also fell into this category.

Vocabulary quizzes

These quizzes were frequently used by teachers to help students understand the scientific terminology. All of my participants used the vocabulary quizzes.
Table 9- continued

| Non-Laboratory Hands-on learning activities | These activities involved exercises such as building molecular models, building different models of atoms, building different compounds and naming them through drawing. All of my participants engaged students in this type of activities with varying frequencies. Melody and Kevin were the ones who used hands-on activities the most. |
| Laboratory activities | Assessing students’ learning through laboratory activities was a common practice across my participants. Students collected and recorded data, organized data in a logical manner and were asked to draw conclusions from their data. Students were not given the chance to design the experimental procedures by themselves, rather they were instructed to follow the procedures given to them. Sometimes teachers challenged student learning in the laboratory by asking them mathematical problems. The attributes of the laboratory worksheets remained similar across my participants. |
| Student presentation | Students researched a topic or a scientist of their interest, created a poster and presented the results of their research to the class. Student presentations were more popular among teachers who taught the lower end classes. Only Kevin and Melody took advantage of this type of learning activities. |
| Games, competitions | Teachers used games and competitions to help students memorize the names of elements, the names of polyatomic ions and the names of ionic and molecular compounds. Again, only the teachers teaching to the standard level students frequently incorporated these forms of activities in their classes. Only Melody used this type of learning activities. |

My analysis of the textbooks suggests that although the textbooks reflected some key elements of reform-based ideas of science instruction and assessment, other documents such as lesson plans, worksheets and other instruments of assessment; tests, quizzes and homework assignments and laboratory worksheets remained traditional. If the curriculum materials and the teaching of science continue to reflect the fundamental attributes of an industrial metaphor of schooling (Marshall, 2006; Tobin, 1990), educational reform is unlikely to take place. Two fundamental reasons led to the
traditional assessment of students’ learning in science classrooms to prevail. These reasons include; 1) teachers’ lack of practical knowledge of reform-based ideas of teaching and assessment, and 2) teachers’ understanding of what schools are designed for. It is my interpretation that these two limitations led teachers to focus on students’ acquisition of canonical knowledge in their assessment of students’ learning instead of focusing on students’ acquisition of inquiry skills.

*Characterizing Science Teachers’ Assessment Practices*

After presenting my findings pertinent to my participants’ assessment practices, I characterize my participants’ assessment practices using Tsai’s (2002) model.

**Table 10. Teachers’ Essential Practices of Assessment.**

<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptor</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>Assessment is limited to rote memorization of facts and a test of students’ acquisition of canonical knowledge. The role of assessment is limited to accountability purposes.</td>
<td>Providing answers that match expert knowledge, assessment of students’ learning involves checking their scientific vocabulary, the memorization of formulas and theories</td>
<td>Kevin</td>
</tr>
<tr>
<td>Process</td>
<td>Assessment is limited to students’ ability to solve mathematical problems following codified procedures and to their ability to reproduce existing knowledge.</td>
<td>Students answer the test questions by following problem-solving procedures; test questions engage students in the processes of verification of the existing scientific theories. Feedback is given to the students is limited to grades.</td>
<td>Melody Ryan</td>
</tr>
<tr>
<td>Constructivist</td>
<td>Assessment of students’ learning of science involves engaging students in the process of doing science and construction of knowledge.</td>
<td>Assessment tasks are designed to engage students in the construction of knowledge, assessment tasks encourage students to make connections between what they learn in class and their everyday experiences. Feedback is given to the students for the purposes of engaging students in further thinking to assist them in their quest to construct knowledge.</td>
<td>Ryan</td>
</tr>
</tbody>
</table>

Adopted from (Tsai, 2002, p.775).
Summary

My findings on teacher’s conceptions and practices of assessment suggests that although teachers’ stated assessment purposes reinforced the goals of science education reform, their assessment practices contested those purposes. Teachers considered assessment as means to ascertain students’ acquisition of the fundamental scientific knowledge and skills of critical thinking, not necessarily purposefully selected methods of identifying the failing and over performing students. However, their assessment practices did not serve these purposes. For instance, in order to meet their accountability requirements to the school administration, to parents and to students, teachers often used assessment as a means to rank students, test their level of responsibility and motivation to do the work required for taking advanced classes or graduation.

This observation suggests that teachers’ assessment practices failed to ensure students’ acquisition of knowledge and skills deemed important by *NSES* (NRC, 1996), which encourage students to construct knowledge rather than being passive recipient of canonical knowledge displayed in the textbooks. Furthermore, my subjects’ instructional and assessment practices suggest that the perceived practical needs of students to meet the requirements of high school graduation took precedence over the purposes that science education reform documents promote (AAAS, 1993; NRC, 1996). More interestingly, such instructional and assessment practices contest the popular motto “raising the standards of public school education,” promoted by the *NCLB Act*.

Although all participants held a view of assessment that is partially consistent with reform-oriented definition of assessment, the lesson objectives that they measured students’ learning against did not reinforce the science literacy goals outlined in *NSES* (NRC, 1996). For instance, although all participants maintained that they wanted their students to acquire the scientific inquiry skills, appreciate the processes involved in scientific discoveries by engaging them in the laboratory activities, their teaching primarily focused on students’ acquisition of scientific knowledge. Although the acquisition of inquiry skill was an important component in their beliefs about learning science, it was just a peripheral outcome of students’ learning experiences in the laboratory.

Although my participants were motivated and willing to design assessment
materials that perceivably reflected the premise that science is an active process
demanding high-order thinking skills, their knowledge of alternative forms of assessment
proved insufficient to engage students in higher thinking. Furthermore, the contextual
influences on teacher practice convinced teachers that alternative forms of assessments
did not hold power to ensure students’ learning of science. Although teachers blamed
students’ lack of responsibility and their motivation to do the work necessary to learn
science, teachers’ naïve understanding and insufficient knowledge of alternative forms of
assessment limited their integration of alternative forms of assessment in science
classrooms. Finally, because teachers reduced learning to the acquisition of canonical
knowledge, using alternative assessments did not become particularly important.
CHAPTER EIGHT

SCIENCE TEACHERS’ CONCEPTIONS OF THE FCAT

High enrollments and ethnic and racial diversity of students have placed new pressures on public schools across the nation (Marshall, 2006; NAS, 2006). In order to make adaptations to the continuing changes in public schools both the federal government and local governments are now taking significant measures (Stiggins, 2004). For instance, the recent NCLB Act (U.S. Department of Education, 2002) expands the federal role in education and sets requirements that affect every public school in America (Brickhouse, 2006; Southerland & Hutner, in press). The fundamental goal of NCLB is to close the achievement gap between different segments of U.S. population by raising the standards of learning, increasing the standards for hiring teachers in core subjects and allocating federal money to the after school programs (U.S. Department of Education, 2002). The NCLB Act requires that all students be tested on mathematics, reading and science annually, demands that schools document students’ annual progress, and threatens schools who fall short of achievement benchmarks (U.S. Department of Education, 2002). This policy significantly impacts how schools function and how teachers meet the curricular demands and address students’ learning needs (Brickhouse, 2005; Southerland & Hutner, in press).

As a consequence of the NCLB Act, states are now asked to take responsibility for improving the achievement of their poorly performing students on standardized tests (Brickhouse, 2005; Southerland & Hutner, in press). In response to the federal NCLB Act, the state of Florida formulated FCAT to measure students’ achievement on standardized tests in reading, mathematics and science (Florida Department of Education, 2006). Scholars argue that tests such as FCAT pressure schools and teachers to bring about improvements to students’ test scores using a punishment and reward system rather than using an assessment system that is responsive to students’ learning needs (Berliner, 2006; Brickhouse, 2006; Moore, 2004; Stiggins, 2004). In spite of the widespread criticism of the FCAT, it is relatively new to the science educators and teachers, therefore, a significant number of questions are yet to be answered before making judgments about the influence that the FCAT has on science curriculum, science teaching, students’
learning and assessment of students’ learning in science. In the following passages, I present evidence from three science teachers to shed light on the influences that the \textit{FCAT} may have on their science curriculum, science teaching and students’ learning. However, I find it useful to introduce the structural influences that the pressure of the \textit{FCAT} generated on the operation of Southern High School prior to introducing science teachers’ conceptions of the \textit{FCAT}.

\textit{FCAT’s Influences on School Curriculum}

The Southern High School administration made significant changes to school curriculum and operations to meet \textit{FCAT}’s achievement objectives. Kevin’s answer to my question on the influence of the \textit{FCAT} on school curriculum indicates the level of pressure felt by the school administration. Kevin said:

Well the fear of—with the Florida [practice] of letter grading schools—and it kind of gets into punishments—that we don’t want to end up with those low scores, and we—there is sometimes, there is scare—scare of getting one of those so called low scores—to get the punishments and not have the benefits of having the high scores and we try to focus more on the \textit{FCAT} and focus more on teaching to the test. So we end up with—so last year we had—I guess it was called “SHS [Southern High Reads] reads” to get their reading scores up. We altered the whole schedule so the students have a little bit of time where students were reading—and [in our] lesson plans we needed to cite specific examples of topics and some sort of specific codes. (Kevin interview, March 12, 2005)

The Southern High School administration made significant changes to both topics to be covered in the curriculum and to the schedule of classes to avoid the punishment of receiving a low grade on the \textit{FCAT}’s grading scale. School administrators encouraged teachers to make references to the \textit{FCAT} objectives in their lesson plans. Furthermore, the school administration spent substantial amount of money on test material and offered more programs in reading and mathematics to improve school’s scores on the \textit{FCAT}. They purchased books for teachers that focused specifically on the science \textit{FCAT} and shared those with science teachers expecting them to allocate a section of their lesson to the “\textit{FCAT Dailies}.” These “\textit{FCAT Dailies}” were practice test items that mirrored what the authors of the resource book thought might be the design of the science \textit{FCAT}. Teachers were to have a “Bell ringer” activity during the first five minutes of the class in which the administration expected teachers to engage students in an \textit{FCAT} related...
activity that did not have to have science content.

The school administration also encouraged teachers to become certified in reading to avoid sanctions posed by the NCLB. Thus, the school administration threatened teachers of losing their jobs if they did not get certification in reading. Although when this study took place, the FCAT science scores did not count against or for school (the FCAT science scores will become part of the FCAT policy starting 2007), the administration felt the need to make changes to science curriculum, and train teachers to meet students’ learning needs for passing the FCAT.

**FCAT’s Influences on Science Curriculum**

The NCLB Act motivated the school administration to make changes to science curriculum in order to meet students’ perceived learning needs, as determined by the FCAT. For instance, Melody said:

I have seen them [school administrators] drop the physical science in favor of earth-space science. Now, I disagree with their decision. For the same reason that they did that, I think the opposite should have been true. Because of the FCAT I think they should have kept the physical science, because a lower end kid if they went to biology, physical science, earth space science. I think they would be just fine as long as they were getting a good physical science curriculum and a good earth-space curriculum. They would learn enough science to pass the FCAT. (Melody interview, February, 18, 2005)

All of my participants expressed their dissatisfaction with the FCAT-driven curriculum changes. They argued for a sequential model of curriculum that started with basic science to the teaching of advanced science classes. For example, Melody said:

I think physical science is a class that all students need to take. Because I think physical science is the basis of all sciences. Like physics and chemistry —some of the kids are terrified with touching anything that says chemistry or physics. But yet somehow if they have that class with me -I mean I love that class because it covers all sciences. If they absolutely hate chemistry or if they absolutely hate physics there is a change — Like if they know Newton’s three laws they are doing pretty well. And if they go on to physics or chemistry. I mean by the time they are finished with me and by the time the school year is gone by, they will have a higher level of maturity to handle a physics or a chemistry class a lot better. Having that background and I think offering physical science especially to some of the lower-end classes is going to help them. It is not going to fix all the problems with the FCAT. You can go a little bit slower and kind of take away the frustration that the majority of those lower end class kids have with physics or chemistry. (Melody interview, February, 18, 2005)
Melody acknowledged the presence of a culture of fear among low-level students towards taking advance science classes. She explained that most of students’ fear came from students’ low reading skills and their lack of fundamental scientific knowledge. She maintained that this fear can partly be taken away by restructuring the sequences of the science courses offered. However, Melody maintained that the pressure to perform well on the FCAT took precedence over students’ learning needs as the administration made curriculum changes. This finding is consistent with Settlage and Meadows’ (2002) findings which report that the administrative pressures due to standardized testing ignores teachers’ professional wisdom, knowledge and skills to address students’ learning problems.

Teachers expressed their dissatisfaction with the changes that took place. However, the teachers’ dissatisfaction was not considered as the administration made curricular changes. Neither were the goals of national science educational reform efforts considered. The changes in curriculum were made solely on the basis of the perceived effect on student’s test scores on the FCAT. The pressures of the NCLB Act encourage school administrators to focus on increasing students’ test scores rather than improving the quality of science teaching that the students receive. It follows that NCLB and FCAT inhibit students’ acquisition of science literacy goals (Brickhouse, 2006).

Teachers’ Conceptions of FCAT

I present evidence pertinent to my participants’ conceptions of the FCAT in the following passages.

**Kevin’s Conceptions**

Kevin fundamentally supported the idea of putting an accountability system in place; however, he questioned the technical quality of the FCAT science questions.

I do believe that we do need something like this — it is an important necessity to have a standardized test to see that they are hitting like on some sort of minimum expectations. Science questions — I still think that they are roughly done and I don’t think that they provide enough detail about how much a student knows. (Kevin interview, March, 12, 2005)

Kevin maintained that the FCAT ensured students’ acquisition of minimum learning expectations in science which he considered to be the Sunshine State Standards.
The SSS outlines what Florida public school students should know and be able to do in four different grade clusters; PreK-2, 3-5, 6-8, 9-12. The standards describe the student achievement that the state holds schools accountable for students' learning in the subject areas of language arts, mathematics, science, social studies, music, visual arts, theatre, dance, health, physical education, and foreign languages (Florida Department of Education, 2004).

However, Kevin did not think the FCAT questions captured the range of knowledge that students learned in science classrooms. Taking into account that Kevin did not have access to the test questions his statement was quite problematic in terms of informing us about the role of the FCAT in bringing about improvements to students’ learning in science. Kevin’s limited knowledge of the test questions led him to pay heightened interest to SSS standards in his teaching, thus, cover more topics to prepare his students for achievement on the FCAT.

Ryan’s understanding of the FCAT’s role in bringing about improvements to the educational system was an outcome of the collective influences that the media reports and his conversations with his colleagues had on his thinking. Ryan said the following about the FCAT:

I actually heard that said it was actually a pretty good thing — from a survey coming from Harvard. I was talking to some other people this weekend, is it? Can it do a good job of motivating low end schools to do better? Umm sometimes there are some other little motivations within school and outside of school and now there is the point of some of the discussions that I have had with other people. I think the biggest problem with FCAT though is that the assumption is that the students will naturally get better at understanding deep questions as they get older. I think that whole idea that the state has — that it is just suddenly going to happen for students is completely false — if you don’t make teachers teach at the beginning it is just a rote memorization test so I am — I guess I am not dead against state testing but I think there are some other problems with the FCAT as well. (Ryan interview, March, 02, 2005)

Ryan maintained that by holding the poor performing schools accountable for students’ low achievement, the FCAT had the potential to encourage or threaten the administrators for making improvements to their performance. Although Ryan mentioned that the FCAT held power to motivate schools to change their operations, he did not
believe students’ learning in science would improve without taking additional measures such as hiring highly qualified teachers. Furthermore, Ryan maintained that the FCAT questions should target higher level thinking skills rather than measuring students’ memorization of scientific facts. Additionally, Ryan’s attitude towards the FCAT can also be explained by the students he taught. Ryan taught a selective group of students in IB, who have historically performed well on standardized tests (Berliner, 2006). Thus, the threat of standardized tests did not affect his students or him.

Melody’s Conceptions

Melody also acknowledged the viability of the FCAT as an accountability system. Melody said:

I don’t think I am as negative about the FCAT as some of the teachers that I know are. In a way I almost kind of ignore it. Because I think if you are really teaching what you are supposed to be teaching and teaching your material good enough, the kids should be just fine. If your students are still failing the FCAT I think there is a problem. And I don’t think that is a problem with me. I try not to take it personally if students don’t get a good score on it [FCAT]. (Melody interview, February, 18, 2005)

Melody felt confident about her teaching; therefore, she did not think that the FCAT specifically asked her to make changes to her instruction. She maintained that if teachers focus on teaching to the Sunshine State Standards (SSS), they did not need to make any adjustments to their instruction to bring about improvements in their students’ learning in science. It follows that the problem of student learning in science, for Melody, was a matter of low test scores, rather than the promotion of the type of skills and knowledge deemed important by the prominent science education reform documents (AAAS, 1993; NRC, 1996).

These teachers’ attitudes towards the test need to be considered also in light of the grade levels that were subjected to the tests. As Melody taught ninth graders and the science FCAT was administered at the 11th grade, her students were also not subjected to the test. Kevin primarily taught the 10th graders, thus, his students were subject to the test. Although Ryan’s students also were subject to the test, Ryan did not feel the pressures of the test as his students had a proven record of success on standardized tests. Ryan said:

I don’t pay attention to it [the FCAT] too much. That has been the focus especially with low-end students. Since I don’t teach low-end, I have not really
been there. I mean all my students, when I look at their FCAT scores, every student has passed the FCAT. For 120 students and in each individual subsection only one student in all of my classes failed one section. So to me the FCAT, for me to get the students to do better on the FCAT is not a concern. I mostly have tenth graders so they are pretty confident in the way they handle the questions on the FCAT. (Ryan interview, March 02, 2005)

It follows that the teachers who had more confidence in their students’ test taking skills felt the least pressure to teach to the test. Furthermore, this evidence suggests students’ test scores do not reflect teachers’ performance rather it reflects students’ knowledge and skills that they have accumulated over years of schooling or the skills and knowledge that they bring with them to the school. For instance, students who are drawn into the IB program generally come from families who have high socioeconomic status. Therefore, they perform better than their peers in regular classrooms on standardized tests (Berliner, 2006; Sacks, 1999).

**FCAT’s Role in Improving Students’ Learning**

These teachers held conflicting beliefs on the effectiveness of the FCAT on students’ learning in science. Although they all were supportive of the administration of an accountability system such as the FCAT, they failed to provide a firm explanation for why they thought such a test is necessary for improved students’ learning.

**Kevin’s Conceptions**

Although Kevin did not think that the FCAT held potential to bring about significant improvements in students’ learning of science, he wanted the FCAT to be administered because he thought that the FCAT held schools and teachers accountable for students’ learning. Kevin said:

But it [FCAT] helps to see whether the teachers in different areas are covering certain minimum categories in their subjects, which we should be doing and citing in our lesson plans anyways. So it is a way for the legislatures and government just to see that we as teachers are at least teaching these minimum things. So when a student leaves the high school that they have learned at least some minimum education skills like using algebra and math and using language and writing skills. (Kevin interview, March 12, 2005)

Kevin considered the FCAT as a means to ensure students’ acquisition of minimum mathematics and language literacy skills. Although Kevin believed the FCAT
ensured students’ acquisition of minimum literacy goals, he did not believe that the FCAT held potential to improve students’ learning in science. Kevin said:

I don’t think it [FCAT] does improve [students’ learning] I don’t think it has improved, but I don’t think it [FCAT] really does too much harm either — it may give us a false reading for the moment because like I said I don’t think it is assessing properly what students know. (Melody interview, March, 12, 2005)

Kevin’s limited knowledge of the FCAT and his limited knowledge of alternative conceptions of instructional improvement led him to view the FCAT as a means to ensure students’ acquisition of minimum learning expectations. Furthermore, the media reports released by Florida Department of Education (Florida Department of Education, 2006) that advocate the positive influence of the FCAT on students’ learning may have had an influence on the way Kevin viewed the FCAT’s role in bringing about improvements to students’ learning. Also the FCAT scores did not affect Kevin because the FCAT science scores did not count for high-stakes decisions.

Ryan’s Conceptions

Ryan supported the administration of the FCAT as an accountability measure to motivate the lower end students to perform better. However, he maintained that the FCAT reduced the purpose of education to high test scores. Ryan said:

After reading their little booklet they say it is supposed to be just one of the assessment, but I think it is THE assessment and I think that is a problem for education so you can get better FCAT scores and that may help in the low-end but is that really the goal of education so -I think the education is much much more than what is just there. (Kevin interview, March, 02, 2005)

Ryan expressed his dissatisfaction with the ways in which the FCAT influenced science instruction. Ryan maintained that the FCAT reduced students’ learning to the increased test scores and he called for the teaching of higher-order thinking skills. Ryan’s beliefs on the role of FCAT in students’ learning reflected his pedagogical sophistication. Consistent with his rationalistic pedagogical beliefs, Ryan wanted the FCAT to measure students’ critical thinking skills. Although Ryan’s sophisticated pedagogical conceptions influenced how he interpreted the influence of the FCAT on other science teachers’ instructional decisions, he himself failed to translate that sophisticated understanding into his own assessment practices. It follows that in addition to political influences teachers’ own curriculum shapes how a teacher scaffolds instruction and assesses students’
learning in science classes.

**Melody’s Conceptions**

Melody supported the administration of *FCAT* as an accountability measure partly because she thought the *FCAT* will bring about improvements to students’ learning. Melody said:

> I think *FCAT* raises standards of achievement, and that is something that we need for these kids. I see people being all negative and all reactionary and that bothers me. There is nothing wrong with having an assessment, and there is a lot of problems with the type of assessment that we have with the *FCAT*. I would like to see it fix the problems that we have with the *FCAT* on an intellectual basis instead of the reaction — soo — it is not a matter of who passes and who fails — it is about why those students are failing. We need to decide whether we need that test, and if we need it what is wrong about it and what to fix about it. I think we are focusing too much on who is giving the test, who is passing it and who is failing it. (Melody interview, February 18, 2005)

Melody felt that people opposed the administration of the *FCAT* for political reasons rather than for its educational impact on students’ learning. She argued we should focus on fixing the problems with the content, structure and the administration of the *FCAT* instead of focusing our argument on political concerns. Melody maintained that the *FCAT* had potential to impact students’ learning if the problems with the test are addressed.

However, Melody failed to realize that the *FCAT* itself is a political assessment practice that has significant consequences for socioeconomically disadvantaged student populations (Popham, 2006). This quote suggests that Melody supported the administration of the *FCAT* because of its role in placing greater accountability on students through high stakes, rather than informing the role of *FCAT* in students’ acquisition of science literacy goals advocated in science education reform (AAAS, 1993; NRC, 1996).

**Summary of Teachers’ Conceptions of the FCAT**

All of my participants agreed on the importance of an accountability system that would set higher standards, adjust curriculum according to those standards, and pressure teachers to teach to the new standards and measure students’ progress against those standards. They were convinced that such an accountability system in place would
improve the quality of education, thus, increase students’ learning in science. However, they argued in order for the *FCAT* to bring about improvements in students’ learning, the problems pertinent to *FCAT*’s content must be addressed.

Teachers’ dissatisfaction with the state of education in Florida’s schools led them to support the administration of the *FCAT*. In addition to their dissatisfaction with students’ performance, teachers’ limited knowledge of the *FCAT* influenced their perceptions of the *FCAT*. The State Department of Education denied teachers’ access to the test questions because it did not want to invest additional money into developing new test items every year. The *FCAT* administration had teachers sign an agreement that threatened them with a $1000 fine if they read the test questions. This policy had significant consequences in terms of how the teachers taught science. Teachers were confused over what knowledge or skills their students were called upon to demonstrate on the *FCAT* because of their limited access to the test questions. This confusion caused teachers to focus on students’ acquisition of canonical knowledge, rather than scaffolding instruction around students’ acquisition of critical thinking and scientific inquiry skills (Gallagher, 2006; Marton & Saljo, 1984; NRC, 1996).

**Validity of FCAT**

Florida Department of Education designed the *FCAT* to monitor the performance of public schools by annually measuring student’s knowledge and ability in reading, writing, mathematics and science (Florida Department of Education, 2005). Although politicians establish a strong relationship between students’ learning and their *FCAT* scores (Greene & Winters, 2003), my participants conceptions of the *FCAT* did not reflect this ideal.

**Kevin’s Conceptions**

Kevin emphasized that the *FCAT* did not cover the range and depth of knowledge that students learned in school. Kevin said:

> Science is very vast and it covers a lot of different topics and a lot of different principles. The test covers a lot of detailed topics on a lot of various subjects that some teachers might have covered but not every teacher, some teachers might have covered this other topic but not every teacher has, and so, if a student who was not taught that other topic--when he or she comes up to that question that has not been covered, they are stuck. (Kevin interview, March, 12, 2005)
Kevin challenged the validity of the FCAT by acknowledging the influence of teachers’ instructional goals and practices on students’ test scores. For instance, he said, not all students learn the same material in science classrooms. Furthermore, this quote suggests that Kevin recognized science as a collection of facts and basic principles not a way of thinking and solving complex problems. Kevin’s answer showed that he believed students were dependent upon the teacher and the teacher was responsible for providing students with the correct answers, rather than approaching science learning as a way of knowing. The test then reinforced Kevin’s beliefs about the need to cover more chemistry topics and acted against suggested science education reform efforts.

**Ryan’s Conceptions**

Ryan expressed his dissatisfaction with the way the FCAT science measured students’ learning in science. Ryan maintained that in its current format the FCAT did not measure students’ learning of science. Ryan said:

I wish for the FCAT Science though—I wish they would not test general subject — because this is not what the kids are actually taking — I wish they would specialize — ok here is FCAT for earth space and environmental, and here is FCAT for physical science, FCAT for chemistry for your biology and that would make a lot more sense to test students that way — I mean my students are learning a lot but they are not necessarily learning what is on the FCAT always. (Ryan interview, March, 02, 2005)

Ryan argued that students should be tested on the subjects that they learned in school rather than on their knowledge of general science. Although Ryan thought that his students learned extensively in his class, what his students learned in his class did not necessarily matched the content that he thought was on the FCAT. Taking into account that Ryan did not have access to test questions, he based his argument on the knowledge that he gained by looking at released sample test questions.

Although the FCAT administrators maintain the FCAT encourages teachers to focus on teaching critical thinking skills (Florida Department of Education, 2005), teachers’ limited knowledge of the test questions led them to teach science for students’ acquisition of special knowledge outlined in SSS. It follows that the manner in which FCAT is administered, pressured science teachers to teach science to promote science literacy goals that were counter to science education reform recommendations (AAAS,
Melody’s Conceptions

Melody maintained that the test scores alone did not reflect how much science a student knows. She argued the FCAT scores reflected students’ level of motivation, their test taking skills and their socioeconomic backgrounds. When I asked Melody to tell me what she thought the FCAT measured, she gave me the following answer:

I think it can tell us a lot about how well a child is taking the test and definitely about where they live and about how motivated they are and I think in the right format it really puts out what they are learning. I think we need to couple it [a student’s test score] with their grades, a teacher recommendation or a portfolio if we want to make it [assessment] fair and meaningful. (Melody interview, February, 18, 2005)

Melody maintained that test scores provided limited information on students’ learning of science. Melody maintained that in order for the FCAT to provide valid information about a student’s learning, test scores should be coupled with other evidence of students’ learning such as student’s grades, a teacher recommendation, and a student portfolio.

Summary

A significant number of scholars in educational research (Berliner & Bidle, 1995; Berliner, 2006; Brickhouse, 2006; Darling-Hammond, 2003; Haberman, 1991; Popham, 2006; Sacks, 1999; Shepard, 2000; Stiggins, 2001) maintain that test scores are not reliable measure of what students learn in schools. The perceptions of the teachers in this study reinforce the previous findings and suggest that the FCAT test scores provide educators with false readings about students’ learning in science. For instance, my participants maintained that FCAT is a poorly structured test that fails to capture the breadth of knowledge and skills that the students are taught in public schools. Furthermore, my participants did not believe that the FCAT held potential to bring about improvements in students’ learning in science. Although the teachers did not think that the FCAT held the potential to bring about improvements in students’ learning of science, they presented a supportive view about the administration of the FCAT.

These teachers’ expressed reactions to the FCAT were consistent with how they viewed science teaching and learning. For instance, teachers reduced science education to
increased high test scores and asked for the changes that would ensure high test scores. They expected students to absorb knowledge presented to them, retain that knowledge and prove to the administrators that they could hold that information for a certain period of time by receiving a passing score on the *FCAT*. Although my participants’ answers primarily focused on the problems highlighting the content of the *FCAT*, their answers also suggests that teachers had limited knowledge of the content of the *FCAT*.

The Influence of the *FCAT* on Science Teaching

The primary responsibility of any science teacher is to promote students’ learning in science using various pedagogical and assessment techniques (NRC, 1996). Science education reform documents emphasize that science teachers should foster students’ critical thinking skills and their skills of doing scientific inquiry, not just their acquisition of scientific facts and principles (AAAS, 1993; NAS, 2006; NRC, 1996). However, a growing number of educators argue that standardized tests pressure teachers to teach to the test, which limit students’ learning to acquisition of prescribed knowledge (Berliner, 2006; Brickhouse, 2006; Popham, 2001; Southerland & Hutner, in press). These scholars maintain that reducing science teaching to students’ acquisition of canonical knowledge is counter to the science literacy goals advocated by science education reform. Because of its centrality to science education reform, I discuss my participants’ conceptions about the influences of the *FCAT* on their teaching in the following section.

*Kevin’s Conceptions*

Although Kevin reported the influence of *FCAT* on his assessment practices, he did not think that he had made significant changes to his assessment practices in order to address students’ learning needs for passing the *FCAT*.

I have tried to make a little bit of my paperwork *FCAT*ish. Giving them some defined areas a little bit more and higher level thinking so they can learn how to adjust and work with things in different settings so in case they see a different example that I did not cover they can still relate to chemical principles and work with them. (Kevin interview, May, 06, 2005)

Kevin’s answer indicates that the *FCAT* motivated Kevin to make some positive improvements to his teaching. For instance, his interpretation of the knowledge and skills that the *FCAT* measured encouraged him to teach his students higher order thinking skills and the application of chemistry to unfamiliar contexts. Kevin maintained that he felt
confident about his teaching, therefore, he did not feel the need to make changes to his instruction. Kevin’s confidence in his teaching came from his commitment to teaching to the SSS.

*Ryan’s Conceptions*

Ryan rejected teaching the *FCAT* strategies because of his confidence in his students’ intellectual ability, knowledge and skills for handling the challenges presented to them on the *FCAT*. Although Ryan expressed that the *FCAT* did not influence his instruction he acknowledged that teachers who taught the standard level classes were pressured by the administration to make instructional and assessment decisions based on students’ learning needs for passing the *FCAT*. Ryan said:

> I don’t pay too much attention to the *FCAT*. That has been the focus especially with low-end students since I don’t teach low end so I have not really been there—I mean all my students when I look at their *FCAT* scores every student passed the *FCAT*. For 120 students and in each individual subsection only one student in all of my classes failed one section so to me the *FCAT* for me to get the students to do better on the *FCAT* is not a concern. I guess some teachers are trying to cover more *FCAT* material in their classroom. I know Ms. W is doing the “*FCAT* dailies” at least she was doing it last year and I now hear that Mrs. B is doing the “*FCAT* dailies,” too. (Ryan interview, March, 02, 2005)

This evidence suggests that because teachers who teach the lower-end students focus on teaching test taking skills, lower-end students may be experiencing poor learning in science. Furthermore, limiting students’ learning in science to the basic knowledge tested on the *FCAT* does not reinforce the goals of science education reform, which advocates teaching science for conceptual understanding, applications of scientific content in real life context, and students’ acquisition of critical thinking and inquiry skills (AAAS, 1993; NRC, 1996).

*Melody’s Conceptions*

Melody rejected establishing a connection between the quality of her teaching and her students’ failing *FCAT* scores. Melody viewed her job as teaching to SSS, and she felt confident about fulfilling her responsibilities in terms of teaching the science content outlined in the SSS. Therefore, she did not feel the need to make significant changes to her instruction. Melody related students’ failing test scores to some external factors that she did not think she could control. Melody said:
I try not to take it personally if students don’t get a good score on the FCAT. I think there is a problem somewhere inbetween that we have to figure out as a team. Like all the teachers, the administration and guidance counselors everybody together and try to see where is the problem. Some of the students cannot read and write — some of the kids don’t know or have not had experience with taking tests. Some of the kids can’t read well enough to take a test. It is really sad to see that in high school but that is what is out there. (Melody interview, February, 18, 2005)

Melody expressed a firm relationship between her students’ reading comprehension skills, their test taking skills and their test scores on the FCAT. Melody incorporated reading comprehensions skills in her teaching of science to address students’ learning needs for passing the FCAT. However, Melody maintained that addressing students’ various learning needs including their reading comprehension skills required all stakeholders to collaborate. This involves identifying the sources of students’ learning problems and developing a responsive plan.

Melody initially maintained that the FCAT did not have influence on her teaching. However, as the interview unfolded Melody started to mention that the FCAT influenced her lesson plans as well as her assessment practices. When I asked Melody to tell me whether she had made any changes to her assessment practices she said:

I do when I am designing my tests. I believe in essay questions I think that is important for them to learn how to put things together and be able to answer essay questions so instead of giving them essay questions I give them a box that looks like the FCAT box for essay. I make sure that they have multiple choice questions and short response questions they have a test that kind of mimics the look of the FCAT. So you know the face of it the look of it, kind of looks like the FCAT. I don’t think it has changed my test tremendously. Test taking strategies. I spend a lot of time in class really teaching them test-taking strategies and dissecting those short response answers and extended responses and teach them how to answer them. I don’t feel like that I have changed a whole a lot I think I have changed my language into the FCAT language so that it connects better. (Melody interview, February, 18, 2005)

Melody explicitly taught her students test-taking strategies and changed the format and the content of her assessment items to mimic the FCAT questions. Melody made changes to her assessment practices to ensure her students’ success on the FCAT. As Melody was mostly concerned about her students’ graduation from high school, she focused on teaching her students test taking skills to ensure their graduation from high
school. However, she felt that the changes that she made to her tests were not significant.

**Summary**

Although all participants expressed a strong support for the administration of the *FCAT*, distinct patterns emerged when they talked about their experiences with the *FCAT*. For instance, the group of students that the teachers taught determined the nature of the influence that the *FCAT* had on teachers’ instructional and assessment decisions. For instance, Ryan did not feel the administrative pressure to make changes to his instruction, while both Melody and other teachers that taught the standard students felt the administrative pressure to make changes to their assessments to improve students’ *FCAT* scores. Ryan felt adequate and successful in his teaching because of his students’ high test scores.

This suggests that *FCAT* policies made the administrators and teachers believe that educational improvement in science is all about increasing students’ test scores rather than about promoting the goals of science education reform. However, we know that the project of science education reform is multifaceted; it involves changes in curriculum, pedagogy and assessment (AAAS, 1993; NRC, 1996; Collins, 1998; Southerland & Hutner, in press).

In addition to the administrative pressures, teachers’ own conceptions of teaching and learning led teachers to teach for content coverage. For instance, all of my participants described teaching to the *SSS* as their primary responsibility. Teachers expressed their firm confidence in their teaching skills, content knowledge and the way they managed assessment of students’ learning in science. They firmly believed in students’ abilities and potential to learn, however, they expressed the reality that a lot of students were still struggling for various and complex reasons which they thought teachers alone could not afford to address. They felt confident that if students cooperated with them on meeting the learning objectives, turned in all the assignments on time and sought help when they needed, all students had the potential to pass the *FCAT*. However, my conversations with these teachers suggest that students’ commitment to learning was not present in spite of the high stakes involved and incentives that the school offered for better test performance. For instance, school administration promised to give away parking permits to increase students’ motivation to score high on the *FCAT*.
Although the *FCAT* encourages schools to use the punishment and reward system to motivate students for passing the *FCAT*, research repeatedly reveals that the punishment and reward based metaphor of assessment fails to motivate students for measuring up to the higher standards of achievement (Delandshere, 2002; Linn, 2000; Shepard, 2000). Furthermore, the findings from this multiple case study suggest that the *FCAT* fails to reinforce science education reform goals as it focuses on increasing students’ test scores on a test to which teachers do not have access. This urges us to look for alternative accountability measures for increasing students’ learning of science and for ensuring students’ acquisition of science literacy goals outlined in science education reform documents such as *NSES* (NRC, 1996).

I discuss factors that challenged the enactment of assessment reform in science classrooms and elaborate on ways to reduce the influence of those factors on students’ learning of science in the following final chapter.
CHAPTER NINE

DISCUSSION

The purpose of this chapter is to make meaning of the findings that I presented in the previous chapters, and to formulate an argument that will answer my research questions (see Table 1) as well as provide new directions for future research in science education.

In my discussion, I first elaborate on the nested relationship between teachers’ epistemic views of science, their pedagogical conceptions and their conceptions of assessment. Secondly, I discuss the structural, cultural and political attributes that influenced teachers’ conceptions and practices of assessment. Then, I elaborate on the ways in which teachers may be able to empower themselves for changing the culture of assessment in science classrooms. Finally, I focus on the recommendations for future research, for professional development of classroom teachers and policy makers. I end this chapter by acknowledging the limitations of my study.

Recapping the Purpose of Study

The purpose of this study is to characterize the professional and structural influences on the enactment of assessment reform in science classrooms. More precisely, I examine the ways in which science teachers’ epistemic views of science and pedagogical conceptions influence their assessment conceptions and practices. Furthermore, I examine teachers’ assessment conceptions and their assessment practices in light of the structural; political and cultural attributes of the school system.

Professional Attributes

I summarize the findings on my participants’ conceptions of science, pedagogy and assessment in the following tables in order to discuss the professional attributes of the teaching profession that presented challenges to the enactment of assessment reform in science classrooms.

These tables are the same tables that I used in the findings chapters (Chapter 4, 5, 6 and 7) of my dissertation. Therefore, I did not change the numbering system that I used for the tables. Also, representing these tables makes it easier for the reader to establish a
connection between my findings and my discussion of my findings.

Table 11. Teachers’ Essential Epistemic Views of Science.

<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptor</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>Science provides correct answers, or science represents the truth.</td>
<td>Accurate description; neutral or objective observations; objective interpretations; truths.</td>
<td>Kevin</td>
</tr>
<tr>
<td>Process</td>
<td>Scientific knowledge is discovered through ‘the’ scientific method or by following codified procedures.</td>
<td>The scientific method; codified procedures; process of discovery; following scientific rules.</td>
<td>Ryan</td>
</tr>
<tr>
<td>Constructivest</td>
<td>Science is a way of knowing, and it is invented through scientists’ agreed conventions and paradigms.</td>
<td>Invented reality; imaginative acts; theory-laden observations; constructed through social negotiations.</td>
<td>Melody</td>
</tr>
</tbody>
</table>

Adopted from (Tsai, 2002, p. 775).

Table 12. Teachers’ Essential Views of School Science.

<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptor</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>Science needs to be part of curriculum because it contributes to our economical productivity.</td>
<td>Science helps us sort the best and motivated students and enables us to prepare them for future careers in science. The rest only need to acquire the scientific knowledge necessary to understand the media reports on science</td>
<td>Kevin</td>
</tr>
<tr>
<td>Process</td>
<td>Science needs to be part of school curriculum because it helps students learn about the scientific procedures necessary to solve complex problems</td>
<td>Science enables students to use the scientific method to conduct scientific experiments. Science empowers students to weigh the credibility of competing claims using the principles of science</td>
<td>Ryan</td>
</tr>
<tr>
<td>Constructivist</td>
<td>Science needs to be part of school curriculum because it helps students develop a scientific habit of mind</td>
<td>Science empowers students to become critical thinkers who are able to make informed decisions both about the natural and social phenomena</td>
<td>Melody</td>
</tr>
</tbody>
</table>

Adopted from (Tsai, 2002, p.775).
<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptors</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Science teachers' beliefs of teaching science</strong></td>
<td>Traditional</td>
<td>Science is best taught by transferring knowledge from teacher to students.</td>
<td>Transferring of knowledge; giving firm answers; providing clear definition; giving accurate explanations; practicing tutorial problems; presenting the scientific truths or facts.</td>
</tr>
<tr>
<td></td>
<td>Process</td>
<td>Science is best taught by focusing on the processes of science or problem-solving procedures.</td>
<td>Teaching the scientific method; following problem-solving procedures; experiencing the processes of (self) discovery; working on the processes of verification.</td>
</tr>
<tr>
<td></td>
<td>Constructivist</td>
<td>Science is best taught by helping students construct knowledge.</td>
<td>Helping students make interpretations; providing authentic experiences; interacting with students; encouraging discussion and cooperative learning; paying attentions to students' prior knowledge or misconceptions.</td>
</tr>
</tbody>
</table>

| **Science teachers' beliefs on learning of science** | Traditional | Learning science is acquiring or ‘reproducing’ knowledge from credible sources | Transferring of knowledge; memorizing formula, definition, keywords and scientific facts; copying what teachers do; hard work on practicing tutorial problems; passive listening; finding the right answer; accurate calculation. | Kevin |
|          | Process     | Learning science is focusing on the processes of science or problem-solving procedures. | Understanding the scientific method; following problem-solving procedures; learning through the processes of (self) discovery; working on the processes of verification. | Ryan |
|          | Constructivist | Learning science is constructing personal understanding. | Making interpretations; exploring or coping with authentic experiences; discussing with peers and teacher; relating to prior knowledge or (personal or daily) experiences. | Melody |

Adopted from (Tsai, 2002, p.774).
Table 14. Teachers’ Essential Conceptions of Assessment.

<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptor</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td><strong>The purpose of assessment is to monitor students’ progress in order to meet the accountability purposes.</strong></td>
<td>Teachers assess students’ work to record their academic progress for certification purposes; passing the class and grade reporting. Assessment of students is limited to acquisition of prescribed knowledge.</td>
<td>Kevin</td>
</tr>
<tr>
<td>Process</td>
<td><strong>Although the sophisticated purposes of assessment are well understood, the purpose of assessment is to assist students to acquire knowledge and skills deemed important by curriculum demands.</strong></td>
<td>Assessment is viewed as a vehicle to monitor students’ performance for acquiring minimum knowledge and skills required by SSS with limited reference to real world application of scientific knowledge.</td>
<td>Ryan, Melody</td>
</tr>
<tr>
<td>Constructivist</td>
<td><strong>Assessment focuses on the improvement of both the instruction and students’ learning growth.</strong></td>
<td>The purpose of assessment is to monitor students’ progress for the purposes of helping them identify their weaknesses and strengths. Assessment involves giving students feedback for the subsequent actions for their learning growth. Assessment forces students to make connections between the science content and its real life applications. The teacher uses assessment results to craft instruction; both for adjustments to the type of knowledge and skills promoted and the instructional methods used to facilitate students’ learning</td>
<td></td>
</tr>
</tbody>
</table>

Adopted from (Tsai, 2002, p.774).

The Nested Relationship between Teachers’ Epistemic, Pedagogical and Assessment Conceptions

The findings summarized in the above tables suggest a nested relationship between science teachers’ epistemic views of science (Table 4) and their conceptions of school science (Table 5). However, the nested relationship between teachers’ epistemic views of science (Table 4), and their pedagogical conceptions (Table 6) and their conceptions of assessment (Table 7) is weak.

Although the sophistication level of Kevin and Ryan’s epistemic conceptions of science and school science matched the sophistication level of their pedagogical and
assessment conceptions, this was not the case for Melody. For instance, while Melody held constructivist epistemic views of school science, her pedagogical conceptions and her conceptions of assessment stayed at the process level. Furthermore, although teachers’ conceptions reflected the attributes of more than one category, I placed them in one single category partly because the information that teachers provided on other categories were relatively weak. Taking into account the complexity involved in characterizing teachers’ internally held conceptions, claiming a significant linear relationship between teachers’ epistemic views of science, and their pedagogical and assessment conceptions would be a naïve argument.

In addition to the complexity of capturing teachers’ internally held conceptions, the structural influences also account for inconsistency between teachers’ epistemic, pedagogical and assessment conceptions. For instance, Melody’s pedagogical and assessment conceptions did not reach her epistemic sophistication partly because Melody’s pedagogical and assessment conceptions were challenged by the contextual and political influences. Melody primarily taught the standard students, therefore, Melody often felt pressured to entertain teaching goals that did not reflect her epistemic sophistication. For instance, Melody responded to the FCAT objectives in her classes in order to avoid the punishment that was attached to students’ low FCAT scores. Similarly, Ryan’s pedagogical conceptions did not perfectly reflect his epistemic views of science because of the perceived curriculum demands. For instance, although Ryan held authoritarian or traditional pedagogical conceptions, the demands of the IB program influenced Ryan to assess relatively higher order skills in his assessment practices.

Furthermore, the inconsistency between teachers’ epistemic views of science, and their pedagogical and assessment conceptions suggest that teachers did not understand the nested relationship between science literacy goals that the science education reform documents such as NSES promote, the pedagogical sophistication needed to reinforce these goals and the assessment methods necessary to ascertain students’ acquisition of these science literacy goals.

Although teachers’ pedagogical and assessment conceptions did not significantly reinforce their epistemic views of science, there was a nested relationship between teachers’ pedagogical conceptions and their conceptions of assessment. For instance,
Kevin’s conceptions of assessment remained traditional partly because his pedagogical conceptions reflected the attributes of behaviorist pedagogy. For Kevin, scientific knowledge consisted of a discrete body of knowledge that could be transferred from teacher to the learner by means of straightforward lecturing and acquired by the students through rote memorization. Consistent with his pedagogical conceptions, assessment for Kevin was just a test of how much of the scientific facts and laws his students had absorbed. Although Kevin reduced the purposes of assessment to a failing or passing grade because of his naïve pedagogical conceptions, the local, school-based policies that required teachers to report students’ progress in the form of a letter grade also had an influence on Kevin’s assessment decisions. For instance, the school policies did not require Kevin to measure students’ depth of understanding rather they only required Kevin to measure students’ acquisition of scientific facts and laws. These policies were outlined in the teacher handbook. For instance, the county policy guiding the teachers’ assessment practices focused on the accountability aspect of assessment instead of pressuring teachers to reinforce higher order thinking and inquiry skills in their assessment of students’ learning. Because accountability took precedence over the type of knowledge and skills that the teachers needed to assess, teachers met their responsibility by reporting a letter grade to the administration through progress reports.

Similarly, both Ryan’s and Melody’s conceptions of assessment remained at the process level because their pedagogical conceptions reflected the attributes of the process level. Evidence presented in the previous chapters indicate that while both Melody and Ryan had a basic understanding of constructivist pedagogy and the accompanying assessment conceptions, they expressed some beliefs that were consistent with behaviorist pedagogy. Therefore, I placed both Melody and Ryan in the process category. Although both Ryan and Melody were cognizant of the assessment goals that the prominent science education reform documents advocate (AAAS, 1993; NRC, 2001), their assessment conceptions did not reflect the attributes of the sophistication needed to reinforce the science literacy goals advocated in science education reform documents. For instance, teachers’ conceptions of assessment frequently required students to provide correct answers that matched the textbook knowledge instead of challenging students with adequate feedback that they need for continuous learning (Davis et al., in press).
The prominent science education reform documents (AAAS, 1997; NRC, 2001, NAS, 2006) suggest that assessment should focus on students’ abilities to identify patterns and establish relationships, encourage students to think critically and search for meaning (AAAS, 1993; NRC, 1996). If the teacher is merely interested in checking whether students are able to complete certain learning tasks and know certain factual knowledge, designing learning activities around attributes of higher order thinking, inquiry skills and a search for meaning may be of little concern (AAAS, 1993; NRC, 1996; NRC, 2001; 2006). Furthermore, for assessment to reinforce the goals of science education reform, assessment should not be about measuring how much knowledge students have accumulated but about how much growth in students’ learning has taken place over time (Davis et al., in press; NRC, 2001; Shepard, 2000; Willis, 1993).

In summary findings suggest that teachers’ assessment conceptions did not reflect the attributes of the reform-based ideas of assessment due to the collective influence of their naive epistemic and pedagogical conceptions and the political pressures of the school system.

After elaborating on the nested relationship between teachers’ epistemic, pedagogical and assessment conceptions, I discuss how my participants’ conceptions of assessment translated into their assessment practices. However, I find it useful to introduce the Table 10, which summarizes my participants’ assessment practices.

<table>
<thead>
<tr>
<th>Category</th>
<th>Perspective</th>
<th>Descriptor</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>Assessment is limited to rote memorization of facts and a test of students’ acquisition of canonical knowledge. The role of assessment is limited to accountability purposes.</td>
<td>Providing answers that match expert knowledge, assessment of students’ learning involves checking their scientific vocabulary, the memorization of formulas and theories</td>
<td>Kevin</td>
</tr>
<tr>
<td>Process</td>
<td>Assessment is limited to students’ ability to solve mathematical problems following codified procedures and to their ability to reproduce existing knowledge</td>
<td>Students answer the test questions by following problem-solving procedures; test questions engage students in the processes of verification of the existing scientific theories. Feedback is given to the students is limited to grades.</td>
<td>Melody Ryan</td>
</tr>
</tbody>
</table>

Table 15. Teachers’ Essential Practices of Assessment.
Constructivist Assessment of students’ learning of science involves engaging students in the process of doing science and construction of knowledge.

Assessment tasks are designed to engage students in the construction of knowledge, assessment tasks encourage students to make connections between what they learn in class and their everyday experiences. Feedback is given to the students for the purposes of engaging students in further thinking to assist them in their quest to construct knowledge.

Adopted from (Tsai, 2002, p.775).

Teachers’ Conceptions and Their Assessment Practices

A critical interpretation of “Table 10” indicates that science teachers’ assessment practices reflected the sophistication level of their pedagogical and assessment conceptions. However, teachers’ assessment practices did not matched their epistemic sophistication. For instance, while both Melody’s and Ryan’s assessment practices reflected the attributes of process level, Kevin’ assessment practices remained at the traditional level. Kevin designed his assessment practices to ascertain students’ acquisition of canonical knowledge rather than challenging his students to develop skills of inquiry and engaging them in a search for meaning (Marshall, 2006; NAS, 2006). Furthermore, he expected his students to match their learning with his expert knowledge or knowledge that could be found in the textbook, which is consistent with behaviorist pedagogy (Shepard, 2000).

Although Melody and Ryan also expected their students to match their learning with the teachers’ expert knowledge, their assessment practices were more sophisticated than Kevin’s assessment practices. For instance, both Melody and Ryan incorporated aspects of inquiry in their assessment practices. However, they were not able to challenge their students to develop the scientific inquiry skills partly because of their naïve understanding of inquiry. Without a solid understanding of the purposes and intents underlying inquiry learning, assessment practices cannot reflect or support these purposes. These teachers’ limited experiences with learning through inquiry and their limited knowledge of alternative assessment practices led them to reduce the purposes of
assessment to students’ acquisition of canonical knowledge. Furthermore, teachers’ assessment practices did not reflect teachers’ understanding of science partly because the curriculum demands and the political pressures took precedence over teachers’ epistemic views of science. Because curriculum demands and the school politics reduced assessment to accountability purposes, teachers adhered to the traditional means for assessing students’ learning of science. However, the political pressures can also help teachers develop sophisticated pedagogical conceptions. For instance, because the IB curriculum pressured teachers to assess students’ inquiry skills and their critical thinking skills rather than students’ acquisition of scientific facts and laws, Ryan held relatively more sophisticated assessment conceptions than Kevin did. Ryan was unlikely to ask more sophisticated questions in his assessment practices than Kevin did if he was to teach standard level classes. It follows that the pressure to assess students’ inquiry skills helped Ryan develop relatively sophisticated conceptions of inquiry.

In summary, teachers adhered to the traditional assessment methods to evaluate students’ learning as a result of two factors. These factors are; 1) science education reform documents lacked institutional power to influence teachers’ assessment practices and 2) teachers failed to structure their assessment practices around the goals of science education reform documents because of their naïve pedagogical content knowledge (Barnet & Hodson, 2001; Gess-Newsome & Lederman, 1999; Shulman, 1986). Because the absence of these two factors limited the purpose of science teaching to students’ acquisition of canonical knowledge, the traditional assessment methods sufficed the purpose of assessment. Therefore, teachers did not seek alternative forms of assessment to support students’ learning in science. It follows that in order for change in the culture of assessment to take place, science educators not only should help teachers understand the nested relationship between the epistemology of science, science literacy goals, pedagogy and assessment, but they should also help them reassess their taken-for-granted assumptions about teaching, learning and assessment. Furthermore, science educators must help science teachers to acquire the knowledge and skills necessary to teach science for understanding and assess students’ learning for continuous improvement (Davis et al., in press; NAS, 2006; NRC, 2001) rather than using assessment merely for grading purposes.
After elaborating on the professional limitations on the enactment of assessment reform in science classrooms, I focus on the structural influences in the following sections.

Structural Attributes

I focus on cultural and political structures that presented challenges to the enactment of assessment reform in science classrooms in the following passages. I discuss the influences of cultural structures followed by the influences of political structures.

Cultural Structures: The Metaphor of Schools as Workplaces

Schools as the workplace showed up as one of the dominant norms of the Southern High School culture (Marshall, 1988; Tobin, 1990), a schema that served as a constraining power limiting the enactment of assessment reform in science classrooms. Teachers adhered to an authoritarian view of teaching and assessment in which they expected students to match their learning to the teachers’ expert knowledge or knowledge that could be found in the textbook. The metaphor of the teacher as an authoritarian figure in the classroom is an extension of school as workplace metaphor (Marshall, 1988; Tobin, 1990). This workplace metaphor shaped the nature of power relationships amongst administrators, teachers and students, thus, maintained status quo in the assessment of students’ learning.

Marshall (1988) argues that the American educational system grew during the industrial revolution, thus, incorporated many aspects of how industry operates. She argues that science curricula and the responsibilities of the teaching profession have been designed to fit the workplace metaphor (Marshall, 1988; Tobin, 1990). Thus, teachers design assessments to punish those who lack motivation to do the work and reward those who are able to finish all the assignments on time and match their learning with the textbook knowledge (Davis et al., in press; Marshall, 2006; Shepard, 2000).

The Southern High School’s culture reflected the attributes of school as workplaces metaphor (Marshall, 1988). For instance, classroom management showed up as the most crucial component of the teaching profession and consumed a great deal of teacher thinking and energy as they designed and enacted their lesson plans. Furthermore,
instead of allowing students to interact with one another and placing responsibility for learning on the student, teachers arranged seating charts, designed learning tasks so that students had limited interaction with one another. Teachers designed lesson plans in order to maintain control over student thinking and generate predictable behaviors. Science teaching was reduced to keeping students busy with listening to the lectures, taking notes and completing worksheets rather than providing students with the challenge necessary to think critically and search for meaning by doing science (Marshall, 2006; NRC, 1996).

This authoritarian view of teaching in turn shaped teachers’ assessment practices as well. Consistent with Marshall’s (1988) argument, the assessment of student learning was frequently limited to task completion and restatement of established scientific facts and theories. For instance, tests, homework assignments and quizzes served as a system that rewarded hard workers, those who were able to match their knowledge to what the teacher told them and punished the lazy ones who failed to retain what the teacher told them. Several scholars maintain that the teachers who adhere to a work metaphor in their assessment practices view errors as evidence of inefficiency and lack of students’ motivation rather than evidence of misconceptions that the teachers could use to enhance students’ learning (Marshall, 2006; Shepard, 2000). Although changing this metaphor can lead to significant improvements in science education (Tobin, 1990), the road to change is full of challenges. These challenges deal with changes in curriculum goals, pedagogy used to support students’ learning and assessments used to ensure students’ acquisition of science literacy goals (Southerland & Hutner, in press). Unless teachers and administrators pay sufficient attention to these three components simultaneously, the traditional assessment practices that reduce students’ learning to knowledge acquisition are likely to dominate science classrooms across the nation’s public schools (Abell & Volkman, 2006; Brickhouse, 2006; DeBoer, 2002; Delandshere, 2002; Southerland & Hutner, in press).

Political Structures

The fundamental political structures that influenced teachers’ instructional and assessment practices in the context of this study are SSS and the FCAT (Florida Department of Education, 2004) as an extension of NCLB (U.S. Department of Education, 2002).
My analysis of these documents reveals that the nature of learning objectives that the policy documents, *SSS* and *FCAT*, promote, and pedagogical discourses and assessment practices that these documents generate contest the science literacy goals advocated in science education reform documents. Although *NSES* (NRC, 1996) defines the purpose of science teaching as helping students view science as a way of knowing and understanding the course of natural world, and helping them apply what they know about science to the real life and unfamiliar contexts, *SSS* exclusively focuses on students’ acquisition of canonical knowledge. For instance, every benchmark in *SSS* starts with “student knows...” without encouraging teachers to teach science for students’ acquisition of science literacy goals that the prominent science education reform documents such as *NSES* promote. Thus, instead of helping students acquire inquiry skills and view learning as a search for meaning, *SSS* limits students’ learning of science to the accumulation of a range of scientific knowledge, which was tested on the *FCAT*. Although *SSS* fails to reinforce the science literacy goals outlined in *NSES* (NRC, 1996), it holds the institutional power, thus, greatly influenced what teachers taught and how they assessed students’ learning in science classrooms.

*FCAT* as a political extension of *SSS* played a significant role in assessment of students’ learning at Southern High School. The school administration required teachers to make references to *SSS* in their lesson plans partly because of the *FCAT*. Because of the *FCA*-related pressures, teachers felt their fundamental responsibility against their contract was to teach science for students’ acquisition of the learning objectives outlined in *SSS*. Teachers’ commitment to teaching to the *SSS* led to an array of assessment practices that failed to reinforce science literacy goals that the prominent science education reform documents such as *NSES* (NRC, 1996) promote. Instead of fostering students’ acquisition of skills of inquiry and encouraging them to search for meaning, assessment practices required students to reproduce bodies of information selected from the “elite culture”; knowledge of scientific vocabulary, theories and laws (Willis, 1993, p. 384) that was tested on the *FCAT*.

Although the *FCAT* held political power to shape the culture of teaching, learning and assessment, evidence presented in the previous chapter on teachers’ conceptions of the *FCAT* challenges the changes that the *NCLB Act* and the *FCAT* promise to bring to
students’ learning in public schools on three different accounts. First, teachers severely challenged the technical quality of the FCAT maintaining that FCAT failed to measure what students learned in science classrooms. Secondly, the FCAT encouraged the school administration to make curriculum changes not based on the call for reform (AAAS, 1993; NRC, 1996), but based on the practical needs for meeting the NCLB’s achievement benchmarks. Furthermore, these changes led teachers to focus on teaching science for the acquisition of canonical knowledge more effectively. Finally, the pressure of the FCAT led teachers to accommodate other curriculum interests (e.g., teaching for reading comprehension) in science classrooms. All of these factors combined and produced learning outcomes that did not reinforce the science literacy goals outlined in science education reform documents. I elaborate on these three implications of the FCAT for science education reform in the following section.

**Misinterpretation of the Call for Change**

The fundamental fallacy of the FCAT is that teachers are expected to know the types of changes that they need to make in their instructions through their encounters with the sample test questions; thus, to produce significant improvements in student learning. However, evidence shown in the previous chapter indicates that teachers were confused about the type of knowledge and skills that they needed to teach to enhance students’ learning. The fundamental motivation behind the policy that denies teachers access to the test questions is to discourage teachers from reducing their teaching to the content of the FCAT. However, this policy leads to dysfunctional responses such as reducing science teaching to students’ acquisition of prescribed knowledge.

It follows that in order for the FCAT to inform science instruction and lead to improvements in students’ learning in science, the restrictions on teachers’ access to the test questions must be lifted. Furthermore, teachers should be involved in the assessment of the FCAT so they are able to identify their students’ strengths and weaknesses; thus, make responsive adjustments to their instructions and assessment to accommodate the learning needs of struggling students. Unlimited access to the FCAT test questions also holds potential to bring the quality, type of knowledge and skills that the FCAT measure under close scrutiny, thus, providing us with the opportunity to better the FCAT and bring about real improvements to students’ learning of science.
Critics argue that standardized testing policy narrows the scope of curriculum taught in public schools (Firestone, Mayrowetz & Fairman, 1998; Grant, 2001; Popham, 2006; Stiggins, 2004). This multiple case study reinforced the previous findings. The findings from this study suggest that the FCAT policies encouraged teachers to teach for students’ acquisition of factual knowledge rather than teaching to foster the knowledge and skills advocated by the national science education reform documents (AAAS, 1993; NRC, 1996). For instance, schools now offer more courses that match the content of the standardized tests. Furthermore, because test scores are now considered evidence of success for schools, the administration pressured teachers more than ever to pay more attention to the standards as they plan and deliver instruction in their classes (Delandshere, 2002; Popham, 2006; Settlage & Meadows, 2002; Stiggins, 2004). The accountability associated with the test scores, thus, not only compels attention to the teaching of standards, it frames and defines what is acceptable and valued in teachers’ pedagogical knowledge of teaching. As a result of this focus on test scores, many teachers’ pedagogical practices emphasize directive teaching, rather than fostering the acquisition of skills and knowledge deemed important in science education reform documents such as NSES (NRC, 1996). Furthermore, teachers are now motivated to design assessment materials that mimic the standardized test questions, rather than designing and using assessment materials to ascertain students’ acquisition of critical thinking and inquiry skills.

Assessment: Measurement vs Creating Conditions for Improvement

The fact that standardizing testing has not brought about improvements in students’ learning in public schools (Darling-Hammond, 2003; Delandshere, 2002; Popham, 2006; Popkewitz, 2000; Shepard, 2000; Stiggins, 2004) suggests that this traditional and popular approach to educational improvement is based on a fundamentally flawed set of beliefs about the role of assessment in educational reform. Scholars argue testing has its roots in behaviorist approach to learning, which heavily focuses on objectives and outcomes rather than focusing on the conditions and the processes that enable or limit students’ learning in science (Davis et al., in press; Delandshere, 2002; Shepard, 2000; Stiggins, 2004). The fundamental fallacy that lies in the promises of
measurement-based assessments is that setting standards high and testing students on these standards with limited feedback to students will lead to significant improvements in students’ learning (Berliner, 2006; Brickhouse, 2006; Darling-Hammond, 2003; Stiggins, 2004; Wiliam, 2001). However, recent research reveals that in order for students to live up to the higher academic expectations, the learning environment, instructional means and the assessment formats used must be conducive to students’ learning. Because the conditions are not set up for improving the learning of lower-achieving students, test results show scarce improvements in students’ achievement.

A Punitive Approach to Assessment

Many disturbing outcomes of standardized-based reform efforts have been associated with a punitive approach to interpreting test scores (Darling-Hammond, 2003; Popham, 2006; Stiggins, 2004). For instance, although the technical improvements have been observed in the design of standardized tests (NRC, 2005), pursuing a punitive approach to school improvement linked to test scores leads to dysfunctional responses that fail to produce projected results that the policy makers want to see.

A large body of research reports the inadequacy of standardized tests for bringing about improvements in students’ learning in science. Sacks (1999) has long advocated that “standardized tests have thwarted rather than helped educational reform and that they continue to be remarkably biased and inaccurate assessments of the abilities of many Americans” (p. 3). More importantly, he argues that standardized tests favor society's historically privileged elites under the guise of merit. For instance, Sacks (1999) argues that students’ test scores tend to significantly correlate with the income and education of their parents. This link between social class and students’ achievement on the standardized tests has been voiced by other scholars as well. Berliner (2006) finds that poverty places severe limitations on what can be accomplished through standardized test-based accountability measures. That is partly because while the wealthy and well-educated parents pass on the benefits of their social status (knowledge and resources) to their children, students coming from low income families do not have access to those resources (Sacks, 1999).

It follows that test scores are very sensitive to the characteristics of student population, so, schools in certain neighborhoods are in financial trouble. To compound
greater numbers of qualified teachers are now leaving low-performing schools to avoid the sanctions, making these students vulnerable to further failure (Moore, 2004; Popham, 2006; Stiggins, 2004). These consequences harm the performance of the schools in neighborhoods that have high number of African American, Hispanic students and students coming from homes located in socioeconomically deprived neighborhoods (National Commission on Teaching and America’s Future, 1996).

American Educational Research Association (AERA) (2004) states that as the result of the sanctions associated with many state-wide standardized tests, students may be placed at increased risk of educational failure; teachers may be blamed or punished for inequitable resources over which they have scarce control (AERA, 2004).

Critics argue that sanctions are producing greater failure than helping low performing students to score higher on tests (Darling-Hammond, 2003; Sacks, 1999; Stiggins, 2004). While policymakers are busy with raising the bars for academic achievement, drop out rates are going up and fewer students are now graduating from high schools (Stiggins, 2004), distancing these marginalized groups further from access to cultural capital. Thus, instead of empowering students to enter the social system, education serves as an instrument of marginalization for students coming from minority and socioeconomically deprived backgrounds. The Orlando Sentinel reported in 2005 that drop out rates are increasing in Central Florida because schools that face closure are letting low-achieving students out easier than ever in order to avoid the punishment that the schools receive because of low test scores (National Center for Fair and Open Testing, 2005). Standardized tests have begun to act as gatekeepers to deny access to the cultural capital for those students who are coming from generational poverty.

A view of educational reform that forms policy structures that enable equitable distribution of resources and focus on teacher quality and development is likely to create conditions necessary for responsive practice and desired outcomes than a punitive approach to educational reform. Darling-Hammond’s (2003) analysis of several reports looking at the influence of the standardized testing on student learning and concludes that unless students experience “good teaching around appropriate curriculum with adequate resources” they are unlikely to measure up to the standards of achievement defined in accountability systems” (p. 6).
However, as evidence in this study and in the previous studies indicates, the cultural and political structures continue to influence teaching, learning and assessment. Furthermore, since the political will to design and administer standardized testing is unlikely to go away in the near future (Popham, 2006), science educators must become political in helping teachers to reduce the influence of standardized testing on their teaching and students’ learning.

They should help teachers develop emancipatory power strategies to weigh the implications of the actions that they are asked to perform by the political structures for students’ learning and act accordingly. This involves helping teachers develop sophisticated conceptions of science, pedagogy and assessment (Southerland & Hutner, in press). Teachers with such sophistication are likely to reduce the influence of the political pressures on their pedagogical and assessment practices, thus, teach science for understanding (Blanchard, 2006; Yalaki, 2004).

I now elaborate on influences of political cultural structures on day to day operations of public schools with specific attention to the teacher’s role in managing the political and cultural pressures on their assessment practices.

Structures, Teacher Agency and Assessment Reform

Political structures are central to the production and reproduction of purposes, practices and outcomes of schooling; thus, are essential to the sustainability of traditional school norms and expectations and for the possibility of change in schools (Gilmer, 2004; Sewell, 1992). Structures are powerful in terms of their ability to limit and enable the successful enactment of science education reform by exerting control over the teacher agency and by mobilizing resources in certain ways (Brickhouse, 2006; Collins, 1998; Gilmer, 2004; Moore, 2004; Sewell, 1992). Structures are viewed as a complex set of beliefs and rules manifested in a particular space and time for achieving specific goals and objectives (Giddens, 1979; Sewell, 1992). Teacher agency refers to the capacity for freedom of action in light of or despite political and cultural structures (Giddens, 1979; Sewell, 1992).

Sewell (1992) argues, while the influence of the structures on agency is a fact that cannot be ignored, individual agents also hold the power to reduce the influences structures on his/her actions. Furthermore, he maintains the agent has potential to gain
power, thus, influence the structures to evolve. For instance, teachers filter the political structures that create conditions for teaching, learning and assessment, using their own personal beliefs about the curriculum. Teachers’ personal beliefs about curriculum involve their epistemic, pedagogical and assessment conceptions as well as their views of their students and of the context in which they teach (Gudmundsdottir, 1995). Thus, teachers’ personal curriculum theory (Gudmundsdottir, 1995) suggests that the teacher agency determines the educational value and weighs the political significance of what these structures offer in the context of his/her profession. This understanding then acknowledges the role of the teacher agency in maintaining status quo or bringing about change in the educational system. For instance, this understanding of the teacher agency suggests that teachers, who hold sophisticated epistemic views of science, pedagogical conceptions and assessment conceptions, are likely to bring about change in the culture of assessment in science classrooms.

Traditionally, society views teachers as the obedient professionals who perform an array tasks to ensure students’ acquisition of certain curriculum goals that they are asked by their contracts to reinforce in the classrooms (Davis, 1996; Jeanpierre, Oberhauser & Freeman, 2005; Osborne, 1998; Sweeney, Bula & Cornett, 2001; van Driel, Beijaard & Verloop, 2001). Therefore, this traditional view significantly undermines teachers’ role in bringing about change in schools (Bybee, 1993; Cuban, 1990; Davis, 1996). If the teachers are to lead the change in the culture of pedagogy and the accompanying assessment practices in science classrooms (Gess-Newsome & Lederman, 1999; van Driel, Beijaard & Verloop, 2001), our view of teacher agency in educational reform must change (Cuban, 1990; Gilmer, 2004; Sewell, 1992).

Although the awareness about the power of teacher agency over the cultural and political structures of the schools is essential, this awareness alone cannot bring about assessment reform in science classrooms. In order to bring about change in the assessment of students’ learning in a lasting way, teachers must empower themselves by developing sophisticated epistemic, pedagogical and assessment conceptions (Marshall, 2006; Tobin & Dawson, 1992; Southerland & Hutner, in press; Tobin & McRobbie, 1996). Furthermore, if the teachers are to pave the way for change in the culture of assessment in science classrooms, they need to develop a better understanding of the
nested relationship between their epistemic, pedagogical and assessment conceptions. Such understanding may assist them to realize the conflicts between their epistemic, pedagogical and assessment conceptions and subsequently help them take necessary actions to reinforce science education reform goals in their assessment. Teachers may be able to develop such an understanding through their active involvement in autobiographical (Gilmer, 2004) or action research (Butler, 1999; Davis, 1996; Schön, 1983).

If the teachers are not challenged to develop sophisticated conceptions of science, pedagogy and assessment, they are likely to reinforce traditional curricular goals, teach science through traditional pedagogical means and assess students’ learning through traditional assessment methods (DeBoer, 2002; Settlage & Meadows, 2002; Southerland & Hutner, in press). It follows that science education programs must explicitly help pre-service teachers and those teachers who are already teaching to understand the nested relationship between curricular goals, pedagogy and assessment and develop the level of sophistication necessary to reduce the influences of political and cultural structures of the public schools their assessment practices. Although developing such an understanding may be hard for the new graduates as they have limited knowledge of the context in which they may be teaching, teachers with several years of experience are likely to develop such sophistication if given professional development opportunities. I elaborate on how science educators may be able to help science teachers to develop a sophisticated professional knowledge base that will assist them to transform the culture of assessment in the following passages.

Teacher Empowerment and Professional Development

The traditional conceptualization of effective teachers is based on the assumption that the teachers, who are confident in their knowledge of subject matter and their knowledge of pedagogy, are able to create conducive learning environments for students (Osborne, 1998; Davis, 1996; Sweeney, Bula & Cornett, 2001; van Driel, Beijaard & Verloop, 1998; Wallace & Louden, 1992). Therefore, traditional teacher training programs focus on transmitting expert content and pedagogical knowledge to pre-service teachers and expect them to successfully apply what they know about science and teaching into their practice (Davis, 1996; van Driel et al., 1998). However, the dynamics
of high school science classrooms challenge this assumption.

The findings from this study indicate that teachers’ knowledge of content, pedagogy and assessment becomes problematic because of the tension between the need to address students’ learning needs and the pressures to address political demands of the school administration. Furthermore, teachers’ knowledge base is challenged because of teachers’ tendency to fit in the established norms of the school culture, which attempt to ensure students’ acquisition of curricular goals by maintaining control over students’ thinking and behaviors.

Ryan’s enculturation into the teaching profession serves a great example for describing the influence of the school culture on teachers’ pedagogical and assessment practices. For instance, although Ryan held science teaching purposes that were partially consistent with the goals of science education reform (NRC, 1996), his commitment to fit in the existing school culture convinced him to reinforce a set of science literacy goals that contested the goals of science education reform. Ryan’s pedagogical and assessment practices reflected the beliefs of a more experienced IB teacher. The experienced IB teacher was able to convince Ryan that he needed to control students’ behaviors in order to increase his chances of success as a teacher. As the success of teacher was primarily measured by teachers’ ability to maintain control in the classroom, scaffolding instruction and assessment practices around keeping students busy became significantly important for Ryan. Thus, to survive in such culture, Ryan adopted a set of pedagogical and assessment goals that the more experienced IB teacher wanted him to promote.

This multiple case study indicates that because of structural influences, most teachers are in the position of being monitored and are expected to complete tasks simply because they are told to do so, not because they believe in the merit of the outcomes that their practices may generate (Barnett & Hodson, 2001; Gess-Newsome & Lederman, 1999; Osborne, 1998; Sweeney et al., 2001; Marshall, 2006). Submitting to the existing pedagogical and assessment culture can be quite problematic for the enactment of reform-based ideas of assessment in science classrooms. However, the influences of the traditional school culture and political structures on teachers’ assessment practices can be reduced by empowering teachers with knowledge and skills necessary to implement reform-based ideas of assessment. That is partly because the empowered teachers are
likely to critically examine the outcomes of the tasks that they are asked to perform, weigh their significance for the learner and act accordingly (Davis, 1996; Osborne, 1998). By empowering teachers to take a critical look at the goals that they are asked to promote and the tasks they are asked to perform, we may be able to reduce the political and cultural influences on teachers’ assessment practices, thus, make assessment reform in science classrooms a reality. The question that one must answer then is how do teachers become empowered? What types of knowledge and skills are needed for a teacher to become empowered?

The theory of pedagogical content knowledge (PCK) (Gess-Newsome & Lederman, 1999; Shulman, 1986) proves to be a useful framework for helping teachers to recognize inconsistencies in their beliefs and empower themselves to bring about change in the culture of assessment in science classrooms.

_Pedagogical Content Knowledge and Teacher Empowerment._

Many scholars in science education (Barnett & Hodson, 2001; Gess-Newsome & Lederman, 1999; Grosman, 1990; Magnusson, Krajcik, & Borko, 1999; Marks, 1990; Shulman, 1986; 1987; Southerland & Gess-Newsome, 1999; Wallace & Louden, 1992) maintain that science teachers, who possess sophisticated pedagogical content knowledge (PCK) are likely to teach science in a way that is instructionally effective and personally meaningful for students (Gess-Newsome & Lederman, 1999; Magnussonet al., 1999; Osborne, 1998; Shulman, 1986).

Shulman (1986) describes pedagogical content knowledge (PCK) as an amalgam of pedagogy and content. He maintains that the teaching of scientific content must be reorganized in a way that will take into account curricular goals, students’ psychological development and classroom dynamics. It follows that a sophisticated pedagogical content knowledge base involves knowing how to organize, sequence, and present the scientific content to the students in a meaningful fashion (Gess-Newsome & Lederman, 1999). This definition of PCK has helped science educators’ to better understand the unique ways in which teachers understand, contextualize and teach their subject matter for the last 20 years.

However, scholars in the field of PCK have recently started to acknowledge the influence of contextual influences on teachers’ PCK. Scholars, who highlight the
contextual component of teachers’ PCK base, maintain that teachers’ pedagogical knowledge is not fixed, limited to an objective, concrete body of pre-existing knowledge that teachers acquire through their educational training (Gess-Newsome & Lederman, 1999; Van Driel et al., 1998). Rather, the political and cultural structures of schooling influence teachers’ PCK base, therefore, are subject to change as the teachers engage in professional activities in the context of their profession (Barnet & Hodson, 2001; Osborne, 1998). This understanding views teaching as a learning process for achieving professional expertise, thus, highlights the socially constructed process that guides teachers’ professionalism (Bencze, 2000; Driver et al., 1994; Ernest, 1995; Fosnot, 1996; Gergen, 1995; Tsai, 2002). This view of teacher development is consistent with social constructivist view of learning (Tobin et al., 1996).

Interpreting the professional identity of a teacher in a social constructivist framework suggests that teachers are constantly developing responses to policies, actions and events taking place within their physical and psychological proximity (Tobin et al., 1996). In other words, teachers’ instructional and assessment decisions are products of their reconstruction of their priorities pertinent to institutional expectations, contextual demands, collegial cooperation, and students’ perceived learning needs, in addition to their personal beliefs about what should be prioritized in the curriculum. This array of personal and structural influences shapes how teachers go about teaching and assessing students’ learning in science (Barnett & Hodson, 2001; Gess-Newsome & Lederman, 1999).

Barnett and Hodson (2001) argue that the teachers’ PCK base serves as a foundation upon which they are able to teach and a vehicle through which they are able to learn about the most effective means of ensuring curricular goals. This view of teachers’ knowledge base suggests that teachers learn from their experiences as they teach (Osborne, 1998), and what they learn from their experiences shape how they go about teaching science and assessing students’ learning (Barnet & Hodson, 2001; Gess-Newsome & Lederman, 1999; Osborne, 1998; Van Driel, Jong, & Verloop, 2001). Thus, a sophisticated pedagogical content knowledge base not only enables teachers to deliver instruction in an effective manner and in a manner that is personally meaningful to the students, it also enables teachers to establish a measure of authority and to feel adequate,
comfortable and effective in the classroom by addressing contextual demands of the school and classroom (Gess-Newsome & Lederman, 1999; Osborne, 1998).

It follows that science educators can use PCK (Gess-Newsome & Lederman, 1999; Shulman, 1986) as a framework for helping teachers to restructure their conceptual ecology (Southerland et al., 2006), accommodate students’ learning needs and reduce the structural influences on their pedagogical and assessment practices. PCK is a powerful framework for helping teachers to empower themselves, partly because a sophisticated PCK will enable teachers to better understand the relationships between curricular goals, pedagogy and assessment (Gess-Newsome & Lederman, 1999). Furthermore, science educators can use PCK as a framework to help teachers develop sophisticated epistemic views of science, adopt sophisticated pedagogical conceptions, develop sophisticated conceptions of assessment, and transfer this sophistication into assessment practices that will ensure students’ acquisition of science literacy goals that science education reform documents promote. However, in order for the teachers to make sense of a sophisticated PCK base, they need to have sufficient experience in teaching. It follows that instead of transmitting expert knowledge into pre-service teachers, science teacher educators must consider alternative approaches to teacher education in their efforts to help teachers develop sophisticated PCK base (Gess-Newsome & Lederman, 1999; Shulman, 1986). For instance, online teacher portfolios (Unal, 2005) can be used for tracing the growth of teachers’ PCK as they started to engage in the context of their teaching and reflect on their experiences. In addition to helping science educators to trace teachers’ growth, online teacher portfolios can serve as a medium for teachers to learn from each other’s experiences and better respond to students’ learning needs.

Although teacher empowerment is central to change in schools, individual efforts to make change are unlikely to sustain change. In order to sustain change, collective ownership of reform ideas is a must. I elaborate on the need to make science education reform a collective effort in public high schools in the following section.

Collective Vision for Change

Many studies of culture acknowledge the role of agency in social relationships (Sewell, 1992; Sewell, 1999). However, many of these studies view agency as an individual entity, isolated from the social group (Sewell, 1999). Sewell (1992) criticizes
this position and defines agency as a construct that “entails an ability to coordinate one’s actions with others or against others to form collective projects, to persuade, to coerce, and to monitor the simultaneous effects of one’s own and other’s activities” (p. 21).

Sewell’s (1992) view of agency recommends that change in the culture of assessment in science classrooms is likely to take place if collective power of teacher agency, student agency and administration work towards accomplishing similar goals and purposes. Achieving collective agency amongst the stakeholders informs us about political structure’s vulnerability to change; thus, about the successful enactment of assessment reform in science classrooms. At the same time, it challenges science educators to develop effective models for facilitating the formation of a collective agency in science departments at the high school level that will bring about significant improvements to science instruction and to the accompanying assessment goals and practices.

Teachers worked individually, used different resources and followed a different set of rules to scaffold instruction and design assessments in the context of my study. Teachers’ own pedagogical dispositions or the pedagogical disposition of their mentors, peers and administrators, influenced teachers’ instructional and assessment practices. Furthermore, the learning goals that both mentor teachers and the administrators promoted did not help students acquire the type of knowledge and skills that the prominent science education reform documents such as *NSES* promote (NRC, 1996).

The administrators worked aggressively to meet the mandates of the *NCLB Act* or to avoid the punishment of the low *FCAT* scores. Teachers were primarily concerned about the efficiency of their instruction to maintain their status of effective teacher in order to receive positive evaluations. This reduced the purpose of science teaching to addressing the institutional demands rather than helping students acquire inquiry and critical thinking skills, the skills that the science education reform documents promote.

Furthermore, the science department lacked a shared vision for reforming the teachers’ instructional and assessment practices. The department chair was not aware of the purposes of the science education reform documents: *Benchmarks for Scientific Literacy* (AAAS, 1993) and *NSES* (NRC, 1996). Therefore, the department chair failed to reinforce the goals of science education reform documents. Instead, he reinforced the goals of the *FCAT* and encouraged teachers to help the administrator’s ambitious goal to
ensure high test scores during our department meetings. He wanted to make sure that the teachers who taught the lower-end students had enough resources in reading and mathematics to prepare their students for passing the *FCAT*.

It follows that although the empowered teacher agency holds potential to induce change in science classrooms, individual efforts to bring about reform in science classrooms is unlikely to ascertain success for all students in science. In order for change to take place and for science teaching to impact the learning of all students, collective ownership of science reform agenda among science teachers is a must. Furthermore, this collective effort must be organized in ways that will hold significant power to influence the learning of all students, thus, assist them to acquire science literacy goals advocated in science education reform documents: *Benchmarks for Scientific Literacy* (AAAS, 1993) and *NSES* (NRC, 1996). Also, ensuring students’ acquisition of reform-based science literacy goals requires strong leaders who will provide opportunities for the teachers to acquire pedagogical knowledge and skills necessary for teaching science to ensure students’ acquisition of science literacy goals advocated by science education reform documents. In the absence of such leadership, teachers are likely to reinforce the traditional science literacy goals that the political structures such as *FCAT* ask the teachers to reinforce.

**Summary**

This study highlights the need for understanding the ways in which teachers construct their conceptions of teaching, learning and assessment in the presence of various dynamics of school and classroom culture. This understanding is essential in terms of our ability to empower teachers for teaching science to reinforce the goals of science education reform documents: *Benchmarks for Scientific Literacy* (AAAS, 1993) and the *NSES* (NRC, 1996). The findings from this study also suggest that the project of teacher empowerment involves making teachers’ epistemic, pedagogical and assessment conceptions problematic. The better the teachers recognize the limitations of their PCK base (Gess-Newsome & Lederman, 1999; Shulman, 1986), the more likely they are able to acquire necessary knowledge and skills to address students’ various learning needs (Davis, 1996), and teach science and assess students’ learning to ensure students’ acquisition of science literacy goals. Furthermore, because teachers shape their
understandings and practices of assessment by complex interactions between their epistemic, pedagogical and assessment conceptions (Anyon, 1995), calls for assessment reform in science classrooms must acknowledge the influence of teachers’ epistemic and pedagogical conceptions on teachers’ assessment practices.

This study highlights that although teachers’ empowerment is central to science education reform, both the possibility and the pace of educational reform are largely dependent on the political will, the norms and expectations of the school culture (Fullan, 1982; Lipman, 1997; Sarason, 1996). Willis (1993) maintains that the operation of schooling and the possibility of reform in schools are influenced by the balance of cultural, political and economic forces that reflect the wider societal trends. It follows that the arguments for the role of teacher in bringing about educational reform should be examined in light of the wider political and cultural factors influencing the operation of public schools.

Implications of the Study

This study contributes to a large body of literature that acknowledges the complexity of the teaching profession and adds to our understanding of how others can be prepared to engage in this profession in a more effective and educationally responsive manner.

Understanding the implementation of assessment reform in science classroom is a multifaceted and a complex process. This is partly because assessment interacts with various professional, cultural and political factors. Therefore, understanding and facilitating the process of change in assessment cannot be understood by merely looking at the cases of three teachers, however, the insight revealed by the cases of these three teachers has significant implications for science education research, for science education scholars who are involved in professional development of secondary school science teachers, for the classroom teachers and for the politicians formulating policy initiatives to bring about improvements in students’ learning. I elaborate on the implications of this study for various stakeholders in the following passages.

*Implications for Science Educators*

A large body of research in science education looks at the professional attributes
of teaching. This involves research looking into teachers’ epistemic sophistication and their pedagogical conceptions. However, research that looks into understanding teachers’ conceptions and practices of assessment is not abundant in science education literature (Genc, 2005). This multiple case study highlights that research that examines science teachers’ conceptions and practices of assessment is critical to the successful enactment of science education reform agenda. This study also highlights that any research that attempts to understand science teachers’ conceptions and practices of assessment without an explicit reference to their epistemic views of science and their pedagogical conceptions is unlikely to produce insight and knowledge necessary to challenge the traditional culture of assessment in science classrooms. The insights that this line of research may produce are likely to help science educators develop effective models to orient teachers towards adopting a nurturing rather than a punitive approach to assessment in their assessment of students’ learning in science. Furthermore, acknowledging the influence of teachers’ epistemic views of science and their pedagogical conceptions on teachers’ conceptions and practices of assessment provides us with the opportunity to challenge teachers to teach science for ensuring students’ acquisition of science literacy goals advocated in science education reform documents (AAAS, 1993; NRC 1996).

**Implications for the Professional Preparation of Science Teachers**

The traditional conceptualization of the teaching profession has led to the development of a teacher preparation system in which experts (e.g., university professors) offer a body of knowledge and an array of skills to the pre-service teachers, which, in principle are transferable to the classrooms (Briscoe & Wells, 2002; Davis, 1996; Korthagen & Kessels, 1999; Osborne, 1998). These experts are often able to convince pre-service teachers with reasonable arguments about the transferability of the knowledge and skills to the classroom settings (Bullough & Gitlin, 1994). However, once teachers are in the classrooms, the knowledge and skills that they acquire through their formal training becomes quite problematic (Davis, 1996; Osborne, 1998).

As new science teachers enter into the context of the teaching profession in a public school environment, they are often at a loss and find themselves in a constant fight for developing a sense of survival, success in teaching and a sense of belonging (Briscoe
& Wells, 2002; Goodlad, 1990). As the teachers start to come in contact with the students and engage in collegial conversations, they start to recognize certain patterns and practices that are socially acceptable or highly valued in the context of their profession (Wubbels, Korthagen & Brekelmans, 1997). These teachers often show a tendency to unquestionably embrace the norms, values and expectations of the school culture that they identify through their enculturation into the profession of teaching for their survival (Osborne, 1998). This confirmatory tactic enables schools to maintain status quo by pressuring the teachers to reproduce the established norms, values and expectations (Davis, 1996; Osborne, 1998).

The findings from this study reinforce the previous findings and suggest that teachers were introduced into a culture of teaching that encouraged them to maintain status quo in teaching of science and assessment of students’ learning. For instance, the administration and mentors instructed teachers to maintain control in the classrooms rather than assisting them to create a nurturing learning environment for students. Furthermore, the pedagogical practices that the teachers were instructed to reinforce shaped how the teachers went about assessing students’ learning. For example, because pedagogical practices focused on maintaining control in the classroom, the teachers designed assessment systems to reflect the attributes of a punishment and reward metaphor. This punishment and reward metaphor limited the purpose of assessment to accountability and students’ acquisition of the expert knowledge that the teacher was able to deliver through lectures.

It follows that teacher educators must pay special attention to the process of teachers’ enculturation into the school context as they prepare teachers of science for the 21st century (Goodlad & McMannon, 2004). Most science education professors introduce pre-service teachers to reform-based ideas and think highly of what their students are capable of accomplishing. On the other hand, upon their engagements with students most teachers come to an awareness that their education does not offer them ideas and knowledge that they can immediately use to design learning activities in their teaching to produce desired outcomes (Davis, 1996; Fullan, 1993). The tension between what science educators think their students can accomplish and what teachers are able to accomplish calls for a shift from knowledge-based teacher training to practice-based teacher training.
in pre-service teacher education.

If the teacher education programs fail to produce teachers who are equipped with knowledge and skills that will enable them to engage in the critical judgment of their actions and the outcomes that these actions generate, they are most likely to fall prey to the influence of the political structures and cultural values that can only afford to reinforce the traditional curricular goals and assessment practices in science classrooms. When these traditional political and cultural structures become the mere resources, which inform teachers’ instructional and assessment decisions, teachers become actors whose instructional and assessment practices remain limited to reciting preexisting scripts (Fullan, 1993; Goodlad & McMannon, 2004). If the teachers are to pave the way for a change in the culture of assessment in science classrooms, teachers must be assisted to acquire the knowledge and skills necessary for them to challenge the fundamental assumptions of the traditional school cultures and political demands of schools. Without a critical look at curricular goals and the pedagogical practices that the schools promote, there is little cause for optimism that funneling new ideas into pre-service teacher education will bring about assessment reform in science classrooms. Science educators may be able to assist science teachers to become critical by empowering them with sophisticated PCK (Gess-Newsome & Lederman, 1999; Shulman, 1986). A sophisticated PCK will enable teachers to better understand the nested relationship between the nature of science, curriculum goals, pedagogy and assessment. Such an understanding is likely to help teachers change the culture of assessment in science classrooms.

*Implications for Teachers*

An interesting finding coming from this study is that teachers’ assessment decisions focused on ensuring students’ acquisition of canonical knowledge rather than focusing on identifying students’ learning needs and helping them learn science continuously. Teachers need to shift their focus away from transmitting their expert knowledge to students and view students as individuals searching for meaning and understanding as they make instructional and assessment decisions. Conceptual change model (Marion, Hewson, Tabachnick & Blomker, 1999; Strike & Posner, 1982) may serve as a promising metaphor for helping teachers to make that shift. Teachers, who are able to apply the conceptual change model in analyzing their strengths and weaknesses
are most likely to take a critical look at what they know about assessment, challenge
unfruitful assessment practices for students’ learning and adopt assessment conceptions
that are likely to ensure students’ acquisition of science literacy goals outlined in science
education reform documents such as *NSES* (NRC, 1996). Also, in order to reduce the
constraining effects of school cultures and the political structures within the school,
teachers need to develop strategies of gaining power for negotiating priorities within the
context of their profession.

As standardized state achievement tests dominate science classrooms across the
nation and the strong presence of political will to administer them suggests that they are
unlikely to go away in the near future (Brickhouse, 2006; Delandshere, 2002;
Southerland & Hutner, in press; Stiggins, 2004), the notion of teacher empowerment
becomes central to our understanding of the role of the teacher agency in educational
reform. However, one must realize that empowerment is not an object that can be given
to the teachers. As the empowerment depends upon a teacher’s willingness to take on the
challenge, it is a characteristic/disposition that needs to be developed by the teacher
(Davis, 1996). Empowering teachers to become critical analyzers of their own beliefs
about nature of science, about teaching, about their teaching conditions and of the
cultural and of political influences on their teaching is promising for reducing the effects
of structural constraints on the enactment of assessment reform in science classrooms. In
the absence of such empowerment, teachers are likely to reinforce the traditional
curricular goals, adopt traditional pedagogical practices and design assessment to
measure traditional learning outcomes in science classrooms.

Davis (1996) maintains that science educators can assist science teachers to
become empowered by providing opportunities for them in which they can empower
themselves through learning more content, pedagogy or assessment strategies. For
instance, teachers can develop sophisticated epistemological views of science by
collaborating with science faculty (Gilmer, Hahn & Spaid, 2002).

However, as Davis (1996) has discussed, empowerment is an individual choice
that teachers make rather than a set of expert knowledge that can be funneled into
teachers’ knowledge base. It follows that science educators can only provide the
opportunities and external structures to support teachers’ empowerment efforts. These
external structures may include strategies of reflection that will assist teachers to recognize potential inconsistencies between their epistemic, pedagogical and assessment conceptions (Davis, 1996). By assisting teachers to develop an awareness of the inconsistencies in their beliefs and by helping them question their reasons for the particular actions that they take, we may be able to assist teachers to bring coherences and dissonance forth. Thus, make assessment reform in science classrooms an attainable goal.

**Implications for Policy Makers**

Policy makers tend to reduce the project of educational reform to technical improvement to curricular goals and greater accountability measures for the school districts (Brickhouse, 2006; Collins, 1998; Darling-Hammond, 2003; Southerland & Hutner, in press; Stiggins, 2004). Literature suggests although technical improvements to curriculum are central to educational reform, without a knowledgeable and committed teaching task force, innovations in curriculum are unlikely to bring about change in schools (Carlone, 2003; Gallagher, 2006; Southerland & Hutner, in press). It is important not only to empower teachers to become technically sufficient and instructionally effective but also to create a supportive school culture as teachers start to implement new ideas in their classrooms (Carlone, 2003). However, several scholars maintain that due to recent large scale, accountability-based assessment measures, teachers are unlikely to provide nurturing learning environments for the students (Darling-Hammond, 2003, Shepard, 2000). Instead of creating a nurturing learning environment, teachers now focus on teaching test taking skills to ensure high test scores. Teachers’ focus on increasing their students’ test scores lead the teachers to use a punishment and reward metaphor in the assessment of students’ learning. This punishment and reward metaphor has significant implications 1) for students’ learning of science and 2) for reinforcing students’ acquisition of science literacy goals outlined in science education reform documents such as *NSES* (NRC, 196).

**Implications for the Administrators and Department Chairs**

Krajcik et al. (2000) maintain that the school administrators can play a significant role in helping teachers to enact reform based ideas in public schools. Furthermore, they argue that school leaders have the power to influence the nature of science instruction and
assessment methods used to support students’ learning in science. It follows that the school leaders can change the course of assessment in science classrooms by promoting professional collaboration and providing rich opportunities for teachers’ professional growth. These onsite leaders can help establish an atmosphere of collective inquiry into teaching science. If the teachers are encouraged to reflect on their experiences and view teaching as inquiry into learning about teaching, they are most likely to recognize the inconsistencies between their epistemic, pedagogical and assessment conceptions and their assessment practices. Furthermore, an environment of learning can help teachers better address the inconsistencies between various constructs that make up their conceptual ecology (Southerland et al., 2006) and their assessment practices.

Furthermore, such collaboration could generate dialogue about innovative assessment practices among teachers, thus, empower them for translating reform-based assessment ideas into their teaching of science. Such opportunities of empowerment are likely to help forge a shared culture of innovation among teachers and foster their commitment to change the traditional understandings and practices of assessment in science classrooms.

Although professional development opportunities hold potential to foster the integration of reform-based ideas of assessment in science classrooms, a traditional approach to professional development (e.g., transmission of knowledge) is likely unlikely to pave the way for change in the culture of assessment in science classrooms. Therefore, the school leaders must be aware of the limitations of the professional development programs offered by the outside organizations. The science department chairs should weigh the impact of the professional development programs offered by the outside sources on students’ learning act accordingly. However, for this to happen department chairs must be elected based on their knowledge and skills to lead the change rather than based on seniority.

**Limitations of the Study**

There are several limitations to this study. First, understanding the enactment of assessment reform requires a robust understanding of the relationships between teacher teachers’ beliefs and actions, policy structures and school culture. Therefore, the limited number of participants with whom I worked might not have provided enough information
for me to understand the complex nature of the interaction between the teacher agency and the political and cultural structures of the school. Secondly, because these three teachers had limited years of experience in teaching they had unstable and yet developing conceptions. Thus, teachers’ developing conceptions presented challenges to my ability to characterize teachers’ epistemic pedagogical and assessment conceptions. This limitation calls for further research that use alternative frameworks for understanding teachers’ internally held conceptions.

Also, although an interpretive methodology served as a powerful means for me to understand teachers’ conceptions and practices of assessment, it also opened the door for me to introduce my own biases into my discussion. My own engagement in this research may have put limitations on how I might have prioritized the aspects of teachers’ understanding and practices in my interpretations.

**Generalization**

This study characterized the fundamental attributes of only three science teachers’ professional knowledge and examined the influence of cultural and political structures on teachers’ assessment practices limited to a particular school and a state. Therefore, I do not advocate generalizing from cases of three science teachers with varying experiences, successes and failures with the assessment of students’ learning. I rather prefer my readers to reflect on these teacher profiles and take from these teacher profiles the attributes that most strongly resonate with their own local situations.

This study took place in a specific school context with unique political, cultural attributes as well as with unique student population that differs from other settings by their age, socioeconomic backgrounds and academic performance. Therefore, the results of this study cannot be generalizable all grade levels. However, the results of this study can be transferable (Glesne, 1999) to the public school contexts that share similar goals and host similar cultural dynamics. Finally, science teachers may be able to learn from this study by reflecting on the limitations of these three science teachers’ conceptions and practices of assessment.

**Reflecting on My Learning**

This study contributed to my learning as a researcher and as a classroom educator.
The unique challenge and feedback that my professors gave me helped me better understand the challenges of designing a research study. I learned that one must pay special attention to the design a research study. At the beginning of my writing, I heavily focused on enriching the content of my writing with limited attention to my research design. I soon learned that a researcher must ask oneself, what is it that I am trying to find? What is the best way to go about collecting relevant information? What types of participants are most likely to provide me with the most valuable information that I need to answer my questions? Furthermore, I learned that as one writes, one must always ask oneself: Will these data help me answer my research question?

Furthermore, I learned that no matter how good an argument sounds, one must be willing to drop one’s arguments if the data fail to support the argument. However, I also learned that my active participation in teaching encouraged me to focus on the plausibility of my findings for practicing teachers in my writing rather than focusing on the scientific quality of my argument. This tension between my tendency to address the needs of the practicing teachers and to meet the quality standards of the scientific community made the process of my writing relatively longer.
APPENDIX A
Sample Interview Questions

PART A. BACKGROUND QUESTIONS

1. How long have you been teaching? Have you always taught science?
2. Why did you choose to become a science teacher?
3. Can you tell me about your educational background a little bit?
4. How does your educational background influence the way you teach science?

PART B. EPISTEMIC VIEW OF SCIENCE QUESTIONS

1. What is your definition of science?
2. Why do you think science should be part of school curriculum?
3. Can you share your opinion about the intersection of science and religion with me?
4. How does the definition of science that you just gave me influence the way you teach science?

PART C. PEDAGOGICAL QUESTIONS

1. Do you think a person is born with the ability to learn or do you think intelligence is something that we acquire later in life?
2. How do you think learning takes place?
3. How do you go about planning your lessons?
4. What do you want to accomplish in your teaching?
5. How do you teach your classes?
6. What do you do to motivate your students to meet the learning objectives in our class?
7. What factors do you believe influence your students’ motivation for learning?
8. What is your definition of inquiry?
9. How do you go about integrating inquiry learning into your teaching?

PART D. ASSESSMENT QUESTIONS

1. What do you think assessment is for?
2. What do you think grading is for?
3. What are your assessment policies?
4. How do you communicate your assessment policies to your students?
5. What forms of assessment do you currently use in your instruction?
6. How do your assessment practices influence the relationship that you have with your students?
7. What do you know about alternative forms of assessments?
8. What do you think of alternative assessment formats in terms of their contribution to students’ learning of science?
9. How do you account for the cultural differences as you structure and practice your assessment ideas?
10. How often do you use alternative forms of assessments in your instruction?
11. What is preventing you from the use of alternative forms of assessments?
12. In your perception what expectations do you think the parents have of your assessment? And how does that display itself as you interact with parents?

PART E. FCAT QUESTIONS

1. What do you think of the FCAT?
2. What types of policy changes have you observed in this school because of the FCAT?
3. What is your reaction to those changes?
4. How do you think that the FCAT will bring about the expected improvements in students’ learning?
5. What is your response to the FCAT Science?
6. What types of changes have you made to your assessment practices to meet the targeted FCAT objectives by your school? If Yes, What motivated you to do those changes?
7. What types of responses have these changes generated among your students?
8. What do you think the FCAT scores show about students’ learning?
9. How do you feel about measuring teachers’ performance based on students’ FCAT scores?
10. What factors contribute to the design and implementation of your assessment practices?
APPENDIX B
Human Subjects Approval Memorandum
Office of the Vice President For Research
Human Subjects Committee
Tallahassee, Florida 32306-2763
(850) 644-8633 · FAX (850) 644-4392

APPROVAL MEMORANDUM

Date: 11/4/2004

To: Mehmet Aydeniz
8104 Cardinal Cove Circle
Sanford, FL, 32771

Dept.: MIDDLE AND SECONDARY EDUCATION

From: John Tomkowiak, Chair

Re: Use of Human Subjects in Research
Understanding the Challenges of Enacting Assessment Reform in Science Classrooms

The forms that you submitted to this office in regard to the use of human subjects in the proposal referenced above have been reviewed by the Human Subjects Committee at its meeting on 9/8/2004. Your project was approved by the Committee.

The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals which may be required.

If the project has not been completed by 9/7/2005 you must request renewed approval for continuation of the project.

You are advised that any change in protocol in this project must be approved by resubmission of the project to the Committee for approval. Also, the principal investigator must promptly report, in writing, any unexpected problems causing risks to research subjects or others.

By copy of this memorandum, the chairman of your department and/or your major professor is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols of such investigations as often as needed to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Protection from Research Risks. The Assurance Number is IRB00000446.

cc: Nancy Davis
HSC No. 2004.520
LETTER OF INFORMED CONSENT

Dear Science Teacher,

I am a graduate student working on my dissertation under the direction of Dr. Nancy Davis, in the Department of Middle and Secondary Education/Science Education Program, College of Education at the Florida State University. I am conducting a research study to investigate the contradictions of enacting assessment reform in science classrooms at Middle and High Schools.

Your participation in this study will involve permitting me to observe some of your classes and collect field notes during these observations, and participating in several open-ended interviews, which will be audio-taped and transcribed. Your participation in this study will continue from September 2004 until the end of May 2005. Your participation in this study is voluntary. If you choose not to participate or to withdraw from the study at any time, there will be no penalty. There are no experimental aspects of this study.

The results of this research study may be published however your name will not be released in this publication. I will refer to you with a pseudo name for your responses to the interview questions and in my observations of your classrooms. This pseudo name will not be linked to you in any way and I will not keep a master list of names and pseudo names. I will only have access to these documents and audiotapes. I will keep the observation field notes, interview transcripts, and audiotapes in a locked file cabinet and I will destroy them two years after the completion of this research, which is September 31, 2007. All confidentiality will be maintained to the extent allowed by law.

There are no foreseeable risks or discomforts if you agree to participate in this study. Although there may be no direct benefit to you, the possible benefit of your participation in this research is improving the quality of science teacher education programs and also professional development programs for science teachers through understanding the contradictions that exist in science classrooms that interfere with the enactment of assessment reform.

If you have any questions concerning this research study please call me at (850) 980-7159 or send e-mail to maaz7567@fsu.edu. You may also contact Dr. Nancy Davis at (850) 644 7804 or send e-mail to ndavis@garnet.acns.fsu.edu.

Sincerely,

Mehmet Aydeniz

I give my consent to participate in the above study. I understand that I will be tape recorded by the researcher. These tapes will be kept by the researcher in a locked filing cabinet. I understand that only the researcher will have access to these tapes and that they will be destroyed by September 31, 2007. I understand that all confidentiality will be maintained to the extent allowed by law.

(Date)

(Signature)

If you have any questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk, you can contact the Chair of the Human Subjects Committee, Institutional Review Board, through the Vice President for the Office of Research at (850) 644-8533.
Parental Consent Letter for Minors

Dear Parent:

I am a graduate student at The Florida State University. I am asking permission for your child to participate in a research project on assessment. This study will look at the challenges associated with the implementation of alternative assessment practices in science classrooms at your child’s school. This study has the approval and support of your child’s school along with approval from Human Subject’s Committee at The Florida State University.

Your child has been selected because of his/her enrollment in a science course. With your permission, she/he will participate in one-two interview sessions each approximately lasting 20-30 minutes. These interview sessions will be conducted by a researcher from the Florida State University. During each session, she/he will be asked several questions pertinent to her/his teachers’ assessment practices as well as questions related to her understanding of the role of assessment. These interviews will be audio-taped and kept in a locked cabinet for two years. The interview tapes will be destroyed by August 31 2006.

Your participation, as well as that of your child, in this study is voluntary. If you or your child chooses not to participate or to withdraw from the study at any time, there will be no penalty. This study will not affect your child’s grade or participation in academic activities in anyways. Interviews will be conducted at your child’s school. The results of this study may be published but no reports about this study will contain your child’s name. I will not release any information about your child without your permission.

Although there may be no direct benefit to your child, the possible benefit of your child’s participation is to help educators better understand the challenges associated with the implementation of alternative assessment practices in science classrooms.

If you have questions about this study, please contact Mehmet Aydeniz at 850-980-7159 at The Florida State University, or you may contact My Major professor Dr. Nancy Davis at The Florida State University at 850-6447604. Your cooperation will be greatly appreciated.

Sincerely,

Mehmet Aydeniz

I give consent for my child ___________ to participate in the above study.

Parent’s Name: ______________________________

Parent’s Signature ____________________________ (Date) ____________________________

If you have any questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk, you can contact the Chair of the Human Subjects Committee, Institutional Review Board, through the Vice President for the Office of Research at (850) 644-8633.
Florida State University
Office of Research
Human Subject's Committee
Written Child Assent Form

I have been informed that my parent(s) have given permission for me to participate, if I want to, in a study concerning alternative assessment of student learning. My participation in this project is voluntary and I have been told that I may stop my participation in this study at any time. If I choose not to participate, it will not affect my grade in any way.

Name: ________________________________
REFERENCES


Ohio.


http://PAREonline.net/getvn.asp?v=8&n=9
sociological context. NY: Harcourt.


Qualitative Research Interest Group at the University of Georgia.


Studies, 19(4), 317-328.


Wilson, V.A. (1998), Qualitative research: An introduction, purposes, methodology, criteria for judgment, and a rationale for mixed methodology. Database online.


BIOGRAPHICAL SKETCH

Mehmet Aydeniz was born in Sanliurfa, Turkey. Mehmet received his Bachelor’s degree in Science Education from Gazi University in 1997. He worked as a Middle School Science Teacher after graduating from Gazi University. Mehmet Aydeniz was awarded a full scholarship by the Turkish Government for pursuing his graduate studies in the United States. Mehmet attended an intensive English course for 8 months at the Middle East Technical University before coming to the United States.

Mehmet Aydeniz entered FSU as a graduate student in the Fall of 1999. Mehmet graduated with a Master’s Degree in Science Education from The Florida State University in 2001. Mehmet started his doctoral program in the Fall of 2001 in Science Education at the Florida State University.

Mehmet worked as a Lab Assistant in Department of Biochemistry and Chemistry, and as a chemistry and mathematics tutor at The Center for Academic and Retention and Enhancement at The Florida State University. Mehmet also worked as research assistant under supervision of Dr. Penny J. Gilmer for the Florida Collaborative Excellence in Teacher Preparation (FCETP), a National Science Foundation project for two years. Mehmet’s responsibilities involved analyzing research data, writing papers for conference presentations and maintaining the website for FCETP. Mehmet has been working as a chemistry teacher in central Florida since 2004.

Mehmet Aydeniz’s research focuses on science education reform, assessment of students’ learning and teacher development.