# Preparing Teachers and Students for Democracy: Teacher and Student Learning and Authentic Intellectual Work

Author: Ann Marie Gleeson

Persistent link: http://hdl.handle.net/2345/2414

This work is posted on eScholarship@BC,
Boston College University Libraries.

# BOSTON COLLEGE
## Lynch School of Education

Department of
Teacher Education, Special Education,
Curriculum and Instruction

Program of Curriculum and Instruction

## PREPARING TEACHERS AND STUDENTS FOR DEMOCRACY: TEACHER AND STUDENT LEARNING AND AUTHENTIC INTELLECTUAL WORK

Dissertation
by

ANN MARIE GLEESON

submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

May, 2011

**ABSTRACT**

Preparing Teachers and Students for Democracy: Teacher and Student Learning and
Authentic Intellectual Work

By Ann Marie Gleeson

Marilyn Cochran-Smith, Ph.D., Chair


Preparing students to participate in a democratic society means cultivating
citizens who are capable of making informed, rational decisions about complex issues
related to the common good. In order to do this, teachers need to provide students
learning opportunities that promote critical thinking and involve in-depth examination of
meaningful content. Drawing on Gutmann's (1987) theory of democratic education, this
dissertation examines how beginning teachers who were prepared in a teacher education
program that emphasized social justice and democratic practices think about and engage
their students in this type of work. Specifically, using Newmann's (1996) framework of
"authentic intellectual work" as an indicator of knowledge consistent with democratic
education, this dissertation examines the extent to which the learning opportunities
teachers create and the work that students produce demonstrate authentic intellectual
work and examines the degree to which teachers' understandings of student learning
align with authentic intellectual work.

This qualitatively-oriented mixed methods study (Creswell, Plano Clark,
Gutmann, & Hanson, 2003; Morse & Niehaus, 2009) used quantitative and qualitative
methods concurrently to examine 11 beginning teachers' experiences during the
preservice period and first two years of teaching. Using the Teacher Assessment/Pupil

Learning protocol, this study draws upon quantitative methods to evaluate teachers' assessments/assignments (*n*=53) and students' work (*n*=481) on these assignments and qualitative methods to analyze interviews (*n*=54).

Findings suggest that these beginning teachers and their students engaged in "moderate" levels of authentic intellectual work, although this varied widely. The quality of assessments was positively correlated to the quality of student learning. The degree to which teachers fostered authentic learning opportunities is complicated by teachers' beliefs about assessment and student learning and particular contextual factors such as time, accountability frameworks, classroom management, student ability, and content area. Teachers whose goals for learning aligned with authentic intellectual work were more likely to construct more authentic learning opportunities. This dissertation argues that evaluations of teacher performance and student learning must account for the quality of learning and utilize multiple measures of evaluation.

# ACKNOWLEDGEMENTS

To my friends Meredyth Waterman, Kristin Berbae, Emily Presseau, Elissa Zizmor, Lianna Pizzo, and Sharon Kelly, thank you for everything, but especially for believing in me, for keeping me sane, making me laugh, providing me with food and shelter, and offering a shoulder to cry on when I needed one. Hopefully I now have time to see you all again!

To my dear friend, Anne Pfitzner Gatling, I would not have finished this dissertation without you. When times were the darkest, you were there. Thank you for making me leave an empty apartment and start doing work again. We have both earned a cupcake! Thank you (and Michael, Isabella, and Caroline) for being part of my family.

To my siblings and their spouses, JoAnn, Chris, Chrissy, Bruce, Kathy, Scott, and John, and my nieces and nephew, Danielle, Lindsay, Bridget, Brittany, Katelyn, Aurora, and Craig, this dissertation should finally answer your question, "When is Annie going to finish school?" Thank you for being the most amazing family whose love and faith in me is unparalleled.

Most importantly, to my Mom and Dad, words cannot express what you mean to me. Thank you for making all of this possible and for your unwavering love and support. This is for you.

And finally, for Michelle. Although you're not here to finish the journey with me, you are on every page of this work.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

**CHAPTER ONE: PREPARING TEACHERS AND STUDENTS FOR DEMOCRACY**

"American democracy is at risk," one Brookings Institution Report begins (Macedo, 2005, p. 1). Concerns over Americans' disengagement with civic, political, community, and voluntary organizations (Carnegie Corporation of New York & CIRCLE, 2003), lack of knowledge about democracy (Westheimer & Kahne, 2003), the global political situation in the wake of September 11th (Finn, 2002; Ravitch, 2001), and the growing diversity of the U.S. population (Banks, 2007) have fueled recent debates regarding the preparation of American citizens to participate in contemporary democratic society. At the center of these debates are competing notions and assumptions regarding the knowledge and skills necessary to propagate democratic society. Citizens in a democracy are responsible for making important and complex decisions about political, social, and economic issues, many of which have global and long-lasting ramifications. Individuals disagree about how these issues should be handled, and fundamental to democracy is the ability to come to consensus agreements (Gutmann, 1987). In order to make good decisions regarding these issues, citizens need to know how to critically analyze information (Halpern, 1997; Williams, 2005).

The need to prepare citizens capable of making these decisions in an increasingly pluralistic society has prompted a renewed interest in democratic education where students engage in dialogue, learn rational decision-making, and are exposed to multiple and competing perspectives (Michelli & Keiser, 2005b). Part of preparing students for democracy, Michelli (2005) argues, "means preparing them to see the problematic and to act on it. We need them to be active, not passive; engaged, not bored" (p. 7). Critical

thinking, therefore, is imperative to democratic education. National education policy initiatives further support the need for fostering critical thinking with goals to produce students "who demonstrate an advanced ability to think critically, communicate effectively, and solve problems" (National Education Goals Panel, 1991, p. 237).

Research suggests that students are not typically engaged in learning opportunities that emphasize critical thought (Bryk, Nagaoka, & Newmann, 2000; Gore, Griffiths, & Ladwig, 2004; Newmann, Marks, & Gamoran, 1996). Preparing k-12 students to become complex thinkers requires teachers who promote learning opportunities that foster reasoning and decision-making skills and also encourage students to examine problems and issues like those they will encounter in society and the workplace. Because schools will influence the ways in which future teachers prepare citizens to make decisions, critical thinking is "indispensable in the field of teacher education" (Williams, 2005, p. 164). Teacher educators need to prepare teachers to conduct discussions in communities of inquiry, foster supportive environments where students can experiment with new ideas with their classmates, encourage students to support and defend positions, and help students respect and become flexible towards new positions (Michelli, 2005). This also includes developing teacher candidates' reflective thinking so that they are able to "help students become thoughtful, caring, and reflective citizens in a multicultural world society" (Banks, 2001, p. 5). The research on critical thinking in teacher education is sparse  (Williams, 2005), and few preparation programs take a systematic approach to teaching critical thinking, although there is evidence that this is changing (Martin, 2005).

There are calls to place teacher education at the forefront of research in creating and engaging in practices that foster analytic thinking (Kay & Honey, 2006).

The current climate of accountability in education makes it difficult for teachers and teacher education to engage in practices that support democratic and critical teaching and learning (Michelli & Keiser, 2005b). Some scholars argue that test preparation and the focus on math, science, and reading, has narrowed curriculum and removed content related to democracy, social justice, and critical thinking (Kahne & Westheimer, 2003) because policy makers often see material not directly related to basic reading and math skills "as disposable and inappropriate" (Michelli & Keiser, 2005a, p. xix). Although teaching for democracy and acquiring content knowledge are "not mutually exclusive goals," the two concepts are often at odds as teachers are pressured to focus on test-related content  (Michelli, 2005, p. 20) and lower- order skills (Martin, 2005). Some have argued that accountability frameworks that focus on standardized testing serve as a barrier to teaching for democracy (Carlson, 2008; Michelli, 2005) in the way that teachers only teach what is on the test and that the format of standardized tests, such as multiple choice questions, limit the ability of students to interrogate multiple viewpoints (Kovacs, 2009).

Similar constraints exist in teacher education. As teacher education is increasingly held accountable for teacher candidate and k-12 student learning (Cochran-Smith, 2001), it is a challenge for teacher educators to maintain democratic program assessments that account for such things as critical thinking, authentic learning, sociocultural influences on teaching and learning, multiple evaluative measures, and learning over time (Cochran-

Smith, Gleeson, & Mitchell, 2010). Linking teacher preparation to pupil learning, particularly in a way that accounts for the previous items, requires measures that are time-, labor-, and cost-intensive (Cochran-Smith, 2005; Cochran-Smith & Fries, 2005; Kennedy, 1999). Cochran-Smith, Gleeson, and Mitchell (2010) categorize attempts to connect teacher education to pupil learning into four types: (1) correlating preparation programs and entry pathways with pupils' achievement scores; (2) evaluating programs in terms of candidates' demonstration of classroom behaviors correlated with pupils' test scores; (3) assessing the learning opportunities teacher candidates create along with pupils' performance; and (4) assessing the learning opportunities program graduates create along with their pupils' performance. Although the four approaches are described in terms of their consistency with social justice and pupil learning, the framework also applies to democratic education.

In the first two categories, pupil learning is linked to teacher education through pupil performance on standardized tests  or the extent to which teacher candidates demonstrate particular practices in the classroom that have been identified as related to pupil learning. Assessments in these categories are removed from daily classroom practices and may not reflect what really goes on. The last two categories are more closely related to actual classroom practices and the work of the teacher education program. Here, teacher candidates and teachers are assessed on how they analyze their pupils' learning or through an external evaluation of the learning opportunities that they create. For example, this perspective uses measures such as performance-based and authentic assessments (Darling-Hammond, 2006) and teacher work samples (Girod,

2002; Schalock, Schalock, & Girod, 1997) where teacher candidates use samples of pupil work that they have used in their teaching experiences to reflect on pupil learning. Teacher candidates' learning is then evaluated by the teacher education program. The degree to which the last two categories are more consistent to democratic education than the first two categories depends upon the types of work in which pupils are engaged. The connection to work in the classroom, however, is more likely to reveal the quality of learning experiences that pupils are engaged in on a day-to-day basis.

One possible way for teacher education to evaluate pupil learning, in a manner consistent with Cochran-Smith et al.'s (2010) last two categories and democratic education, is through the lens of Fred Newmann's framework of "authentic intellectual work" (Newmann & Associates, 1996). This framework, which is described in detail later in this study, values critical thinking and engagement in real world situations as learning outcomes, ideas consistent with democratic education. With an emphasis on the construction of knowledge, disciplined inquiry, and value beyond school, it provides criteria to measure teacher candidate and pupil learning in a way that takes into account the type of complex thinking that citizens need. The framework has been used to evaluate work that is used in classrooms and reflects what teachers do in their actual practice. Applying authentic intellectual work to teacher education presents an opportunity to assess pupil learning in a way that reflects teacher candidates' actual classroom practices and represents significant learning goals.

**The Purpose of the Study**

This study examines the extent to which teachers[1] who were prepared in a teacher

education program that emphasized social justice and democratic practices engage their

pupils in the type of work necessary for democratic education during the preservice

period and the early years of teaching. In this study, the framework of "authentic

intellectual work" is used as an indicator of knowledge that is consistent with democratic

education based on its emphasis on critical thinking and meaningful content. This

dissertation seeks to examine the extent to which the learning opportunities that teachers

create and the work that pupils produce demonstrate authentic intellectual work, and

examine if teachers think about principles of authentic intellectual work when they talk

about their goals for assessment and pupil learning. Specifically, this study addresses the

following questions:

1. What is the level of authentic intellectual work that teacher candidates/teachers create for their pupils?
2. What is the level of authentic intellectual work of pupils' responses to these learning opportunities?
3. What is the relationship between the quality of learning opportunities and the quality of pupil learning?
4. How does the way that teachers talk about their goals and understandings of assessment and pupil learning reflect the concepts of authentic intellectual work?
5. How does the way that teachers talk about assessment and pupil learning compare to the quality of authentic intellectual work on their assessment tasks and pupil work?
6. What conditions and contexts influence the level of authentic intellectual work teachers provide over time?

---

[1] From this point forward I will use the term "teacher candidate(s)/teacher(s)" to refer to the participants in this study since this study follows participants through their teacher education program (teacher candidates) and first two years of teaching (teachers). I will use the term "pupil(s)" to refer to the k-12 students taught by teacher candidates/teachers.

To do this, this study draws upon data collected during a longitudinal qualitative case studies project on the experiences of teacher candidates as they learned to teach in a teacher preparation program. Using a protocol called the Teacher Assessment/Pupil Learning Protocol, this study looks at assessments/assignments that teacher candidates used in their classrooms, as well as samples of pupils' work on those assessments, to evaluate the quality of authentic intellectual work. At the same time, this study uses qualitative interview data to examine the ways in which teacher candidates/teachers think about authentic intellectual work as they learn to teach and provides insight into the contexts and conditions which influence their perceptions and the quality of authentic intellectual work.

This study found that the beginning teachers who were prepared in a teacher education program that emphasized pupil learning and the construction of knowledge were, on the whole, able to provide learning opportunities for their pupils that were considered to be at a moderate level of authentic intellectual work. There were some teachers who engaged their pupils at higher levels of authentic intellectual work than others. Their pupils' responses to these learning opportunities varied, although, generally speaking, were also at a moderate level and were related to the quality of the assessment. Pupils who were given assessments of higher authentic intellectual quality were more likely to produce work of higher quality. The quality of authentic intellectual work in a classroom was influenced by teachers' understandings of assessment and pupil learning, as well as particular contextual factors such as content area, time, classroom management, and the type of assessments teachers used, particularly when teachers were

mandated to use certain assessments or prepare pupils for standardized assessments. My analysis suggests that teachers whose goals for assessments and pupil learning were more aligned with the underlying principles of authentic intellectual work were more likely to engage pupils in this type of learning.

**The Importance of the Study**

There are many studies related to various aspects of democratic education, teacher candidate learning, pupil learning, assessment, critical thinking, and authentic intellectual work. However, very few studies connect these concepts in relation to teacher preparation and learning to teach, and very few studies examine the problem from the perspective of a mixed methods approach or longitudinally. Thus, this study addresses important gaps in the literature. First, this study focuses on the quality of pupil learning as defined by authentic, critical, and complex engagement with significant and relevant subject matter, central components in democratic education. Second, it attempts to link teacher education to pupil learning by looking at the authentic intellectual quality of teaching and learning inside the classroom and evaluating typical learning opportunities that are presented to pupils on a daily basis. This provides insight into the relationship between assessment, pupil learning, and classroom context. Third, this study uses multiple indicators to examine learning by taking into account teacher candidates'/teachers' perceptions of learning, assessment, and classroom context, while at the same time conducting an independent evaluation of classroom assignments and pupil work. When examined together, these two measures can provide a deeper understanding of the contexts and conditions that influence authentic intellectual work. The use of both qualitative and

quantitative approaches in this study differs from previous studies that tend to examine authentic intellectual work through one method. Finally, this study looks at teacher candidate learning over time through the student teaching period and into the first two years of teaching, a focus that other studies lack.

## Organization of the Dissertation

This dissertation is comprised of six chapters. Chapter 2 introduces the theoretical framework of democratic education as a way to understand the type of teaching and learning necessary to prepare citizens for participation in a democratic society. Democratic education promotes skills such as rational deliberation and critical thinking. More specifically, the discussion turns to Fred Newmann's (1996) framework of authentic intellectual work as a lens to define one way in which teachers can attend to the principles of democratic education. This framework emphasizes the construction of knowledge and in-depth conceptual understanding of significant and relevant material. This chapter presents a review of the empirical research related to authentic intellectual work, including descriptive, cross-sectional effect, and longitudinal studies, to provide insight into using authentic intellectual work to evaluate teaching and learning. Research suggests that when teachers incorporate authentic intellectual work and focus on higher-order thinking, their pupils are able to demonstrate their knowledge in more complex ways and even outperform pupils who have not engaged with authentic intellectual work on national standardized achievement exams. The empirical studies illustrate that teachers, on the whole, are not engaging their pupils in high levels of authentic intellectual work. This literature review illuminates some of the challenges related to

assessing teacher and pupil learning and makes the case for utilizing authentic intellectual work as a framework for assessment in teacher education.

Chapter 3 describes the study's research design. This study, which draws from a larger longitudinal qualitative case studies project, utilized a qualitatively-driven, embedded mixed methods approach (Creswell, Plano Clark, Gutmann, & Hanson, 2003; Morse & Niehaus, 2009) and cross-case study methodology (Stake, 1994) to understand how 11 beginning teachers and their pupils engaged in authentic intellectual work. In this chapter, I present an overview of mixed methods and describe how qualitative and quantitative data were collected concurrently and integrated during the data interpretation stage to provide a nuanced and detailed understanding of how teachers consider aspects of authentic intellectual work. Using quantitative and qualitative data together can illuminate particular circumstances and meanings that teacher candidates/teachers experience, while at the same time, provide a uniform way to compare authentic intellectual work across all 11 cases. I then describe the research design, including the sample, data collection, and data analysis. The chapter concludes with a discussion of the integrity of the study.

Chapters 4 and 5 present the results of the Teacher Assessment/Pupil Learning (TAPL) analysis of these 11 teachers. My analysis draws upon mixed methods to analyze teachers' use of authentic intellectual work through both quantitative and qualitative perspectives and consists of two parts. The first part, which draws upon quantitative methods, is presented in Chapter 4. This chapter examines the authentic intellectual quality of teachers' assessments and their pupils' work as evaluated by researchers

according to a rubric for authentic intellectual work (Research Institute on Secondary Education Reform (RISER), 2001). First, I describe the types of learning opportunities that the beginning teachers in this study provided for their pupils and illustrate how assessments and pupil work were scored by researchers. Statistical analyses were conducted to examine how assessments and pupil work samples were rated on the rubric depending upon factors that have been considered in previous studies such as academic discipline, grade level, and school context. Finally, the chapter concludes by examining the relationship between the authentic intellectual quality of learning opportunities and pupil learning.

Chapter 5 focuses on the second part of the TAPL analysis. In this section I describe how qualitative methods were used to analyze how teachers talked about their goals for assessment and pupil learning and the degree to which their understandings were consistent with concepts underpinning the framework of authentic intellectual work. The beginning teachers in this study were placed on a continuum based on how their understandings aligned with authentic intellectual work. In this chapter I use the qualitative findings to describe and explain the quantitative findings presented in Chapter 4. The analysis of both the qualitative and quantitative data suggests that teachers' practices were influenced by contextual factors such as time, school culture/accountability, student ability and motivation, behavior/classroom management, and the content area in which they taught. The chapter concludes with two case studies to illustrate how teachers' understandings of assessment and pupil learning and contextual factors interacted very differently in two classrooms where one teacher was able to

engage her pupils in high quality authentic intellectual work while another teacher had low expectations of authentic intellectual work. Findings suggest that the ways in which teachers' goals for assessment and pupil learning and their ability to negotiate contextual factors influenced the extent to which these beginning teachers were able to engage pupils in authentic intellectual work.

The final chapter presents a review of the study's findings and explores how this study contributes to our understanding of beginning teachers and authentic intellectual work. The degree to which teachers foster authentic learning opportunities consistent with democratic education is complicated by teachers' beliefs about assessment and pupil learning and particular classroom factors. Teachers face constraints that are not always apparent in single measures of pupil achievement. In this chapter I discuss the complexity of evaluating teacher and pupil learning and the importance of using multiple measures of assessment to account for the quality of pupils' learning. Then I consider several implications for research and teacher education based on this study's findings. My analysis suggests that authentic intellectual work is not wholly supported in schools. School factors such as pressures to prepare pupils for standardized assessment and the time it takes teachers to implement authentic intellectual work influenced teachers' assessment choices and learning opportunities they provided their pupils. This dissertation argues that policymakers and teacher education need to emphasize authentic learning in schools. Teacher education can do this by shaping teacher candidates' understandings of assessment and pupil learning to incorporate authentic intellectual

work and schools need to  consider support structures to help beginning teachers foster this type of learning.

Preparing pupils for participation in democratic society is dependent upon preparing teachers to understand the value of engaging pupils in critical thinking and in-depth exploration of meaningful content. This study contributes to our understanding of how new teachers consider authentic intellectual work and the factors that influence the degree to which they engage their pupils in this type of work. This study raises questions about the authentic intellectual quality of learning that currently occurs in schools and what that means for the future of American democratic society.

**CHAPTER TWO: LITERATURE REVIEW**

This section begins with a discussion of democratic education, the theoretical framework that drives this study's research design. Within the framework of democratic education, this dissertation draws from Amy Gutmann's (1987) theory of democratic education, the relationship between democratic education and critical thinking, and the embodiment of those ideas in Fred Newmann's (1996) framework of "authentic intellectual work." Finally, since this study uses the framework of authentic intellectual work as a way to evaluate the extent to which teacher candidates/teachers address democratic education and pupil learning, I reviewed empirical work related to authentic intellectual work. This literature review provides the background for how authentic intellectual work addresses the principles of democratic education in examining the work of teacher candidates/teachers and their pupils.

**Theoretical Perspective**

This dissertation draws upon the framework of democratic education as a lens to examine the types of learning opportunities that teacher candidates/teachers should provide for their pupils and what pupils should be able to do. Democratic education fosters learning that prepares citizens to deliberate on significant issues, make sound judgments, and solve problems. Democratic teacher education, therefore, seeks practices and assessment that reflect these learning goals.

**Democratic Education**

In the United States, the word *democracy* is frequently used, and there is often an unspoken assumption that all Americans share a common understanding of its definition.

In actuality, however, there is seldom a consensus on what *democracy* is or means. Typically, democracy is understood as representative governance by the people, with an emphasis on concepts such as equality, rights, and law. Key features of a democracy include the rule of law and its role to protect citizens from tyranny, as well as the reliance on an educated, wise citizenry engaged in debates related to the creation of laws, but who, despite disagreement, collectively adhere to the rule of law (Woodruff, 2005). However, democracy has also been defined as an ideal beyond the rule of law. Educational philosopher John Dewey (1916) described democracy as "a mode of associated living, of conjoint communicated experience," pertaining to individuals' interactions with each other and societal interests (p. 87). Disagreements over what democracy is and what democracy means have influenced the ways in which the role of education in American democracy has been conceptualized.

The centrality of public education to American democratic governance goes back to the early years of the republic. Drawing on classical republican governments, the founders believed that democracy depended upon educated citizens. Legal scholar Martin Carcieri (1997) argues that founders, such as Thomas Jefferson and James Madison, "saw education as one of the *necessary* conditions for the success of democracy, for the healthy functioning of republican institutions" (p. 4). In his *Bill for the More General Diffusion of Knowledge*, Jefferson (1778/1984) called for public education to guard against tyranny from an ignorant citizenry. An educated citizenry, Jefferson wrote, would be "able to guard the sacred deposit of the rights and liberties of their fellow citizens" (p. 365). For Jefferson and Madison, democratic education needed to provide students with:

vocational skills that would aid the economy, civic education to further the rights and responsibilities of citizens, judgment to negotiate between individual and public interests, and a liberal education to critically analyze and evaluate political and social issues (Carcieri, 1997). Citizens, therefore, were responsible for knowing how to support the civic and economic structure of the nation.

Democratic education, from this perspective, is necessary to promote the skills necessary to sustain and participate in representative democratic government. Citing Alexis de Toqueville's notion of the "arduous…apprenticeship of liberty," political scientist Benjamin Barber (2001) states that public education is the arena in which citizens work at learning how to protect liberty and become citizens. Education, he writes, is vital to serve and maintain democratic institutions:

> The literacy required to live in civil society, the competence to live in democratic communities, the ability to think critically and act deliberatively in a pluralistic world, the empathy that permits us to hear and thus accommodate others, all involve skills that must be acquired….Empower the merely ignorant and endow the uneducated with a right to make collective decisions and what results is not democracy but, at best, mob rule: the government of private prejudice and the tyranny of opinion – all the perversions that liberty's enemies like to pretend (and its friends fear) constitute democracy. (pp. 4-5)

Here, education provides individuals with the tools to become citizens and make collective decisions in spite of disagreement in order to protect democratic principles.

People who see democratic education as preparing citizens for governance advocate for students to learn about government and citizenry, with an emphasis on civic education. Ravitch (2008) describes democratic education as a liberal, "universal education" where all students learn about American history, civics, and the Constitution, with emphasis on science, math, language, arts, literature, and history (p. 56). Civic

16

education also entails learning about government and law, looking at current events, community service, and participation in democratic processes (Carnegie Corporation of New York & CIRCLE, 2003). In teacher education this type of democratic education occurs in the preparation of social studies teachers to teach pupils how to become citizens in a democracy by focusing on concepts such as civic knowledge, cognitive civic skills, participatory civic skills, and civic dispositions (Patrick & Vontz, 2001).

Democracy in education has also been characterized in terms that go beyond government structures to apply to social issues. John Dewey (1916) describes two aims for democratic education. The first aim applies to government but the second is related to the aforementioned "mode of associated living" (p. 87). For Dewey, according to Mike Boone (2008), "education is essentially a social process, through which individuals are taught to live into a particular democratic ideal" and one where individuals need a "personal stake in its social relationships and the habits and dispositions to accomplish needed social change without instigating chaos" (p. 17). Democratic education, therefore, is an instrument for social change. This thinking supported the social efficiency movement during the early 20th century and the development of a social studies curriculum that addressed what Harold Rugg (1921/1996) described as "principle[s] of social worth" which included examining economic, social, and political relationships; "an understanding of established modes of living;" and problems related to contemporary issues (p. 52). Democratic education, in this case, is centered around examining relationships and current social issues.

Changes in American demographics have also influenced how democratic citizenship and education have been conceptualized. Democratic education has been characterized as education that serves the interests of and incorporates *all* students. Along these lines, James Banks (2007) calls for *multicultural citizenship* education centering on the creation of an "*authentic unum*" where "citizens from diverse racial, ethnic, and cultural communities will participate" (p. xii). Here the role of teachers is to "help students become reflective citizens in pluralistic democratic nation-states" which includes "enabl[ing] students to acquire a delicate balance of cultural, national, and global identifications and to understand the ways in which knowledge is constructed; to become knowledge producers; and to participate in civic action to create a more humane nation and world" (p. 19). The role of teacher education is to challenge teacher candidates to think critically about and confront their beliefs regarding race and culture (Banks, 2001; Dixson & Dingus, 2007). Public schools in a democracy, therefore, are "democratic public spheres" where students learn to be critical of societal power structures (Giroux, 1988) and address diverse perspectives. Sleeter (2008) describes democratic education as an environment where all students receive high quality education that legitimates cultural and linguistic identities; students look at social issues and practice democratic decision-making; and where teachers have students critically engage with real life social issues that involve diverse viewpoints and knowledge. Citizens in a democracy should have a problematic understanding of the world (Cornett & Gaudelli, 2003). Democratic education in this view is critical education.

Some individuals, however, see the emphasis on critical democratic education as harmful to the underlying conceptions of democracy. The tension here focuses on the interpretation of the hallmark principle of American democracy: *e pluribus unum*. For a diverse society charged with coming together to make decisions for the common good, "out of many, one" is central to maintaining an operational government and society. However, interpretations of what this principle means in contemporary society differ, particularly in relation to education. According to Chester Finn (2004):

> Some educators harbor worrisome values: moral relativism, atheism, doubts about the superiority of democracy, undue deference to the "pluribus" at the expense of the "unum," discomfort with patriotism, cynicism toward established cultural conventions and civic institutions. Transmitting those values to children will gradually erode the foundations of a free society. (As quoted in Michelli, 2005, p. 9)

The fear is that embracing diversity focuses on negative aspects of democracy and limits Americans' ability to unite as citizens and further erodes democratic ideals. Focusing on differences detracts from coalescing as a people around the common good (Ravitch, 2001). This critique is used to challenge the views of multiculturalists and critical theorists, such as Banks (2007), who advocate for the interrogation of alternative points of view. The challenge is how to come to consensus among diverse ideas where deliberation is central (Barton & Levstik, 2004).

Debate over the definition of democratic education has intensified in the wake of September 11, 2001. Amid concerns that Americans are uninformed about politics and civics (Macedo et al., 2005), many have called for a re-emphasis on civic education and democracy in schools (Carnegie Corporation of New York & CIRCLE, 2003; Finn, 2002; Westheimer & Kahne, 2003). However, there is disagreement over what that means in

today's political climate. Some educators characterize democratic education as a transmission model that focuses on teaching about citizenship and patriotism (Cheney, 2002; Ravitch, 2001) where the definition of a patriot is someone who "love[s] our country and the ideals for which it stands," offering unquestioning and unwavering support (Finn, 2002, p. 5). Others conceive education in a democracy as a forum for serious debate and disagreement (Apple, 2002; Giroux, 2002; Nash, 2005) where patriotism "combines national loyalty with critical reflection" (Nash, 2005, p. 215) and "social criticism is the ultimate act of patriotism" (Apple, 2002, p. 1768). Here questioning is seen as a vital component of being a patriotic citizen. This example illustrates how democratic education becomes a contested terrain in the struggle to define the foundational principles of American democracy and identity.

**Gutmann's theory of democratic education**. In her theory of democratic education, Amy Gutmann (1987) reinforced the connection between democratic government and democratic education, claiming that the two are interdependent; democratic education is vital to create citizens capable of maintaining democratic government while, at the same time, democratic government is necessary in order to provide democratic educational structures. Democracy, she argues, is built on disagreement among citizens over how to serve the public good and requires deliberation among citizens where citizens are free to consider and discuss issues. In examining who should hold power over decisionmaking in democratic education, Gutmann proposes that democratic education relies on two fundamental principles. The first is nonrepression,

which prohibits citizens from using education to subvert rational deliberation, and the second is nondiscrimination, which is not excluding people from education.

To be prepared to take part in deliberative democracy, Gutmann (1987) argues, students need an education that develops their moral character through "exemplary" models of societal rights and wrongs and an education where they "develop capacities for criticism, rational argument, and decisionmaking by being taught how to think logically, to argue coherently and fairly, and to consider the relevant alternatives before coming to conclusions" (p. 50). Here students learn skills related to understanding and dealing with disagreement. This is important because as citizens, students will be faced with hard choices as they negotiate societal disagreements over what is best for the common good. Therefore, she continued, "Children must learn not just to *behave* in accordance with authority but to *think* critically about authority if they are to live up to the democratic ideal of sharing political sovereignty as citizens" (p. 51). Hence, democratic society depends on teaching students how to think and resolve disagreements within the confines of nonrepression and nondiscrimination.

For Gutmann (1987), teachers and universities, who, for the most part, prepare teachers, are critical to sustaining democracy. Teachers, who act as both citizens and experts in the field of education, need "to uphold the principle of nonrepression by cultivating the capacity for democratic deliberation" (p. 76). They are obligated to act within professional and disciplinary norms while at the same time "shedding critical light on a democratically created culture" (p. 76). Universities act as "sanctuaries of nonrepression" where students can deliberate new ideas without fear of consequences (p.

174). Universities foster ideas that create knowledge that best serves the interest of democracy, even if those ideas go against current societal norms. Although Gutmann does not refer to teacher education specifically, one can extrapolate that teacher education would be seen as a place to further nonrepression in teaching future teachers. Teacher education, therefore, should engage teacher candidates in critical reflection and teach them how to make rational decisions. This is necessary to prepare teachers to engage in similar practices with their pupils.

Other people have drawn from Gutmann's work to look at the role of teacher education and deliberation, particularly focusing on critical thinking. Drawing on John Rawls' (1993) work, Kunzman (2006) argues that "recognizing reasonable disagreement" is the central civic virtue (p. 166). Students need to recognize that reasonable people disagree and people, therefore, need to be able to compromise and accommodate. According to Michelli (2005), teacher education is vital in this endeavor:

> It is through learning knowledge in a way that depends on critical thinking that students, and their teachers, will learn how to make good judgments, how to argue well for their positions, and how to be flexible and willing to alter their positions in the face of evidence. These are the essential ingredients of democratic life, and the things we can expect from a focus on critical thinking. (p. 15)

Critical thinking is one way to strengthen democracy. One goal of teacher education should be to foster critical thinking, especially since preservice teachers often avoid controversial, debatable topics (Rubin & Justice, 2005) and therefore need exposure to argumentation and reasoning (Ten Dam & Volman, 2004).

This study draws heavily from Gutmann's (1987) theory of democratic education. The assumption is that critical thinking and rational decisionmaking are imperative to

serving the interests of democracy and are skills that students need to develop through education. Teachers and universities play a valuable role in supporting this task. This dissertation uses these ideas by arguing that teacher candidates/teachers should be prepared to engage their pupils in critical thinking. In addition, this study also draws on how democratic education has been conceptualized to argue that pupils should be exposed to learning opportunities where they study significant social issues, rigorous content, and material that is related to the world outside of the classroom. This knowledge should be evaluated when assessing pupil learning.

**Critical thinking**. A central component to democratic education is critical thinking, but although the term is ubiquitous, defining the term is less concrete. In their literature review on critical thinking and civics education, Ten Dam and Volman (2004) found three different conceptions of critical thinking. Critical thinking has been defined in philosophical terms as "the norm of good thinking, the rational aspect of human thought, and as the intellectual virtues needed to approach the world in a reasonable, fair-minded way (Gibson, 1995)" (Ten Dam & Volman, 2004, p. 361). It is seen as rational, deliberate, and reflective thought (Ennis, 1989; Williams, 2005). Critical thinkers analyze and evaluate information in a disciplined way. According to Paul and Elder (2006), "Critical thinking is, in short, self-directed, self-disciplined, self-monitored, and self-corrective thinking. It requires rigorous standards of excellence and mindful command of their use. It entails effective communication and problem solving abilities and a commitment to overcome our native egocentrism and sociocentrism" (p. 4). In this perspective, critical thinking is the systematic process in which one thinks.

Psychologists conceptualize critical thinking by concentrating on the processes of higher-order thinking skills. Halpern's (1998), definition of critical thinking adds a purpose to engaging in such thought:

> The term critical thinking refers to the use of those cognitive skills or strategies that increase the probability of a desirable outcome….Critical thinking is purposeful, reasoned, and goal-directed. It is the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions. (pp. 450-451)

In this view there are "clearly identifiable and definable thinking skills" that people can develop and use to become more effective thinkers (Halpern, 1997, p. 17). Lipman (1988) defines critical thinking as "skillful, responsible thinking that facilitates good judgment because it (1) relies upon criteria, (2) is self-correcting, and (3) is sensitive to context" (p. 39). Lipman's six elements of critical thinking include skillful thinking, responsible thinking, judgment, criteria, self-correction, and sensitivity. Facione (1990) considers particular cognitive skills such as interpretation, analysis, evaluation, inference, explanation, and self-regulation as central elements to critical thinking. Critical thinking, in this perspective, is viewed as particular cognitive skills that pupils should possess.

Critical thinking has also been conceptualized as more than just cognitive skills. Critical theorists make up a "second wave" of research in critical thinking and pedagogy (Walters, 1994). The philosophical and psychological definitions of critical thinking operate under "the unwarranted assumption that good thinking is reducible to logical thinking" (Walters, 1994, p. 1). Instead, critical thinking goes beyond objectivity to include other types of thinking such as imagination and intuition (Walters, 1994), and includes "the capacity to recognize and overcome social injustice" (Ten Dam & Volman,

2004, p. 362). Critical thinking from this perspective involves reflective thought (Papastephanou & Angeli, 2007), the ability "to discern certain kinds of inaccuracies, distortions, and even falsehoods" (Burbules & Berk, 1999), and is linked to "local and global projects of social transformation and social justice" (McLaren, 1994, p. xii). Critical thinking is the ability to critique power structures and examine inequities. This fits into the critical democratic stance where pupils challenge ideas.

The majority of the literature on critical thinking instruction considers critical thinking in terms of cognitive skills as opposed to the critical theory perspective (Ten Dam & Volman, 2004). This includes teaching general critical thinking skills that are transferable across disciplines (Ennis, 1989; Halpern, 1998) and domain-specific skills that differ depending upon subject matter (Smith, 2002). Here teaching critical thinking involves "engaging students in dealing with tasks that call for reasoned judgment or assessment, helping them develop intellectual resources for dealing with these tasks, and providing an environment in which critical thinking is valued and students are encouraged and supported in their attempts to think critically and engage in critical discussion" (Bailin, Case, Coombs, & Daniels, 1999, p. 299). Critical thinking often has a negative connotation, with a common assumption that critical means to criticize as opposed to reflecting and evaluating an issue in depth, which teachers and schools often try to avoid (Winn, 2004).

There is very limited conceptual and empirical work examining critical thinking and teacher education (Williams, 2005). Teacher educators have not typically stressed critical thinking in preservice education (Martin, 2005). Suggestions for teacher

education include having courses in critical thinking, embedding it throughout the program, and explicitly teaching critical thinking skills (Martin, 2005; Williams, 2005). Sears and Parsons (1991) argue that teacher educators need to model an Ethic of Critical Thinking for new teachers which includes attitudes that knowledge is not fixed, questions should be asked, awareness of alternative views, tolerance for ambiguity, alternative ways of knowing, skepticism, and understanding of complexity. Michelli (2005) describes teacher education for democracy and critical thinking as programs that have communities of inquiry where teacher candidates can develop ideas, support their arguments, and consider criteria for good judgments, particularly with opposing views. It is not only learning critical thinking skills but "acquiring the competence to *participate* critically in the communities and social practices to which a person belongs" (Ten Dam & Volman, 2004, p. 372). Critical thinking, therefore, is a necessary skill for participating in a democratic community.

In current policy debates, critical thinking has become a buzzword, particularly with the "21st Century Skills" movement (Achieve, 2008; Kay & Honey, 2006; Partnership for 21st Century Skills, 2006; Silva, 2009). Advocates of the 21st century skills movement argue that pupils need new types of knowledge to meet current global, economic, and technological challenges. This knowledge includes subject matter knowledge, global awareness, information and technology literacy, leadership skills, and critical thinking and problem solving skills (Partnership for 21st Century Skills, 2006). In a recent speech, Secretary of Education Arne Duncan (2010) promoted encouraging teachers "to teach the higher-order, critical thinking skills so desperately needed in the

knowledge economy," making the link between higher-order thinking and the economy. Critical thinking in the 21st century skills movement is similar to the type of thinking promoted by democratic education, but serves a very different purpose in this context. The 21st century skills movement is primarily a business model; critical thinking is necessary so that workers can solve problems in the workforce and transfer skills as new technologies require (Achieve, 2008). This is a different purpose than supporting deliberative democracy.

Democratic education and the economic interests underpinning the 21st century skills movement have been described as diametrically opposed to one another. Christine Sleeter (2008), a proponent of democratic education, uses John Perkins' (2004) term "corporatocracy" – defined as the alignment of business, government, and financial institutions with the purpose of consolidating global economic power for the elite – to represent the business interests that are in opposition to democratic interests in education. Democracy and corporatocracy, Sleeter argues, have competing notions of the purposes of education, curriculum, and instruction. First, democratic education values education for the collective public good while corporatocratic education values education as a means to secure market competition and private interests. Second, democratic education embraces a diverse curriculum drawing from a variety of different bodies of knowledge and multiple perspectives and opinions while corporatocratic education emphasizes universal content and skills related to corporate economic interests. Third, democratic education embraces the production of knowledge and critical thought while corporatocratic education advances specific content and skills, oftentimes emphasizing

factual knowledge over the processes of how to produce knowledge. Finally, democratic education values evaluating student learning in multiple ways and incorporates communal and diverse interests while corporatocratic education values specific measures, particularly related to business interests. Sleeter makes a distinction between common and private interests, arguing that democratic education serves all people while corporatocratic education serves only the interests of a few.

Some suggest, however, this dichotomy is not as pronounced as it may seem. Changes in political and economic forces have brought about an "unprecedented convergence of interests" between progressive, democratically-minded educators and business interests (Walser, 2008). There now exists some common ground between proponents of democratic education and 21st century skills regarding what people should know and be able to do in the world. In his work interviewing world business leaders, Tony Wagner (2008) argues that as American workers now compete for jobs with people all over the world, the distinction between "preparing students for the world of work and teaching them about their role as citizens" diminishes greatly (p. xvi). In today's society, technology and the availability and accessibility of information require the same abilities for both work and citizenship. "Thus, work, learning, and citizenship in the twenty-first century," Wagner argues, "demand that we all know how to *think* – to reason, analyze, weigh evidence, problem-solve – and to *communicate effectively*. These are no longer skills that only the elites in a society must master; they are essential skills for all of us" (p. xxiii). According to this perspective, citizenship and work draw from the same body of knowledge and skill set.

To become effective citizens able to participate in American deliberative democracy and a rapidly changing international economy, education needs to pursue knowledge and skills that emphasize critical analysis, knowledge production, communication, and connection to community, national, and global issues. This connects with Gutmann's (1987) framework of democratic education and the emphasis on deliberation because it involves preparing pupils to think about how they make decisions. In order to reach these ideals, teacher education must prepare teachers who are capable of addressing these issues with their pupils and focus on critical thinking. In assessing teacher candidates/teachers teaching and their pupils' learning, this dissertation uses these ideas, as articulated by the framework of "authentic intellectual work" (Newmann & Associates, 1996), as the criteria to evaluate effective teaching and learning.

**Authentic intellectual work**. This study argues that Fred Newmann's framework of "authentic intellectual work" (Newmann & Associates, 1996) is one way that teacher education can evaluate how well teacher candidates/teachers engage their students in work that reflects the goals of democratic education and critical thinking. This section of the literature review begins with a discussion of the conceptual definitions of authentic work, which includes, but is not exclusive to, Newmann's framework. Here I present an overview of the major concepts related to authentic work since the term became a fixture in the literature beginning in the late 1980s. Following this discussion, I present a detailed description of Newmann's framework of authentic intellectual work.

The concept of *authenticity* has a long tradition in education. Plato, Rousseau, and John Dewey, among others, have all written about the significance of learning through

29

real experiences and obtaining knowledge within a situated context (Splitter, 2009). The term *authentic*, however, gained particular prominence in educational research and policy during the 1980s and 1990s. At this time, the use of the term was specifically identified with creating learning experiences and assessments that replicated (or came as close as possible to replicating) the type of learning and work that occurs outside of the classroom that professional adults engage in.

The push for authentic learning emerged during a time where there were a number of different criticisms of the educational system. There were concerns that American schools were falling behind internationally (The National Commission on Excellence in Education, 1983), prompting an emphasis on standardized testing as a way to measure student achievement. However, traditional standardized testing was also under attack at this time. Critics argued that standardized tests were detrimental to student achievement due to the ways in which testing tended to narrow the curriculum (because teachers were only teaching what was on the test) and limited student thinking (because the multiple choice format of standardized tests prohibited students from explaining and expanding upon their answers) (Shepard, 1989). There were also concerns that classroom learning was disconnected from the "real world" and that students were not engaged in learning because there was no relevance to what students were learning in school and what they would encounter outside of school (Resnick, 1987).

In response to these criticisms, school reformers focused on changing the structures of schools in ways that re-centered schooling on students and student achievement (Corbett & Blum, 1996). To do this, many restructuring efforts focused on

changing assessment practices in ways that would better connect classroom learning experiences to assessments (Newmann, 1991; Tye, 1987; Wiggins, 1989b). Instead of traditional standardized multiple choice tests that were isolated from the everyday learning experiences inside the classroom, reformers called for assessment experiences that were more genuine – more *authentic* – to regular classroom practices. In this view, assessments measure "not just any kind of achievement, but rather, valuable or meaningful forms of mastery" (Archbald & Newmann, 1988, p. vi).  It was within this context that "authentic" came to be a term used to describe types of assessment and ways in which pupils thought and engaged in learning. The use of the term became commonplace, but definitions of authentic teaching and learning varied.

One way that "authentic" has been applied to teaching and learning is through the framework of "authentic intellectual work." Fred Newmann and Associates (1996) describe authentic as "something that is real, genuine, or true, rather than artificial, fake, or misleading" (p. 22) where pupils engage in "tasks that are worthwhile, significant, and meaningful" and connected to actual experiences (Archbald & Newmann, 1988, p. 1). The goal of academic work is to engage students in challenging and meaningful intellectual work. Authentic intellectual activities replicate activities that are "undertaken by successful adults: scientists, musicians, business entrepreneurs, politicians, crafts people, attorneys, novelists, physicians, designers, and so on" (Newmann & Associates, pp. 23-24). In authentic intellectual work, pupils construct new knowledge and develop in-depth understandings of subject-matter. To demonstrate this knowledge, pupils produce something that is similar to what adults in the in the field would produce such as

an essay, report, performance, or speech. Unlike standardized tests where students typically have a finite amount of time, choose standardized answers, and work alone, authentic intellectual work involves activities where time is flexible, pupils produce their own answers, and pupils collaborate with one another. The underlying assumption here is that when students learn about material in a way that allows them to examine the subject deeply and connect the material to what is important to their own lives, then students will care about what they are learning and their new knowledge will remain with them longer. This dissertation draws heavily from the concept of "authentic" as developed in Newmann's work; a further discussion of this framework follows in the next section after a description of other conceptions of authentic teaching and learning.

Attending to criticisms of traditional standardized tests, Wiggins (1989a; 1989b) used authentic to describe tests and assessments that went beyond traditional standardized paper-and-pencil tests. Critiquing the drill-like nature of standardized tests that force students to recall information out of context, Wiggins (1989b) argued "a true test of intellectual ability requires *the performance of exemplary* tasks" (p. 703). Similar to Newmann, Wiggins argued that authentic tests are representative of the performance of someone in the field. Authentic tests "involve students in the actual challenges, standards, and habits needed for success in the academic disciplines or in the workplace: conducting original research, arguing critically, and synthesizing divergent viewpoints" (Wiggins, 1989b, p. 706). Authentic tests require students to use judgment and "efficiently and effectively use a repertoire of knowledge and skill to negotiate a complex and multistage task" (Wiggins & McTighe, 2005). Authentic tests are ones in which students know the

criteria upon which they will be assessed and are given time to rehearse, receive

feedback, and refine their work. Unlike traditional standardized tests, authentic tests

allow students to publicly present and defend their work and are often comprised of

multiple tests over time, as in a portfolio. The emphasis is on students mastering the

material. Newmann and Wiggins both describe authentic assessments as having students

do something that a professional adult would do, such as writing a science report or

conducting original historical research. Wiggin's work differs from Newmann's because

it focuses primarily on assessment, whereas Newmann's authentic intellectual work

framework applies to classroom instruction, assessment practices and pupil performance.

Although Newmann and Wiggins were most prolific in the development and use

of authentic learning and assessment, the term authentic has been applied in other ways

as well. Authentic has been used in a general sense to refer to any assessment that is

alternative to "standardized, norm-referenced, multiple choice tests" (Darling-Hammond,

Ancess, & Falk, 1995, p. 4) where students create and construct their own responses

(Arter & Bond, 1996; Fischer & King, 1995; Mabry, 1999). These authentic assessments

include portfolio assessments where students produce a collection of multiple and varied

assessments over time (Fischer & King, 1995), and performance assessments where

students perform a task or produce a product that is then judged for its quality (Blum &

Arter, 1996). Authentic has also been defined as "actual work samples" from the

disciplines where students solve the same problems that people in the discipline solve,

using the same tools and in the same situations (Lesh & Lamon, 1992). For example, in

an authentic math assessment, students solve a problem that mathematicians have solved

using the same resources used by the mathematicians. In teacher education, authentic assessment has been used to refer "to opportunities for developing and examining teachers' thinking and actions in situations that are experience based and problem oriented and that include or simulate actual acts of teaching" (Darling-Hammond & Snyder, 2000, p. 524). Authentic teacher education experiences are ones in which teacher candidates are involved in teaching. For example, in the Performance Assessment for California Teachers (PACT), teacher candidates reflect on their performance in the classroom and examine samples of pupil work to demonstrate their understandings of pupil learning (Chung, 2008; Darling-Hammond, 2006; Pecheone & Chung, 2006); this work is then evaluated by the teacher education program to assess teacher candidates' and pupils' learning.

With so many different definitions of authentic used interchangeably to refer to alternative, performance, and portfolio assessments and intellectual work, it is difficult to differentiate what is meant by authentic teaching, learning, and assessment. Alleman and Brophy (1997) claim that the "literature is fuzzy regarding the distinctions" between authentic assessment and performance assessment (p. 339). Common features of these types of learning opportunities include that they involve the application of knowledge to a situation that normally occurs and the assessments do not take the form of standardized, multiple choice tests. However, there are features that distinguish authentic work and assessment. Wiggins argues:

> But 'authentic' is *not* synonymous with 'performance assessment.' Many so-called performance tasks or constructed-response items do not replicate the conditions under which adults are challenged in context. That's why a potentially

misunderstood term such as 'authentic' must be used. The point is not to measure performance but to *improve* it. (Newmann, Brandt, & Wiggins, 1998, p. 21)

Wiggins emphasized the centrality of reproducing the conditions in which adults operate. Although connecting assessments to the work that professionals do is central to authentic work and assessment, the mere connection and assessment design does not guarantee that students will engage in the high-quality intellectual thought and work that defines authentic intellectual work (Newmann, Brandt et al., 1998). Newmann argues that there needs to be criteria to define the type of intellectual thought and processes that comprise authentic thinking that goes beyond assessment.

*Newmann's authentic intellectual work*. Although authentic work and assessment have been defined in a number of ways, this dissertation draws primarily from Fred Newmann's work (Newmann & Associates, 1996). His work is based on the idea that successful academic engagement comes from looking at what is necessary for significant intellectual accomplishments in the real world and recreating experiences that challenge pupils in similar ways. As noted above, Newmann and his colleagues identify authentic work as the type of thinking that pupils engage in, which includes, but is not limited to, the types of assessments and activities that pupils undertake. Here the emphasis is on implementing "standards of intellectual quality rather than teaching techniques or processes as the central target of innovation" (Newmann et al., 1996, p. 282). Authentic intellectual work is dependent upon the characteristics of learning opportunities that lead to authentic intellectual achievement. This is not solely dependent upon the type of teaching strategies and assessment. That is to say, the particular teaching strategies or assessments can vary and do not on their own guarantee authentic

intellectual achievement for students. Any type of learning activity can be authentic intellectual work as long as it has the particular characteristics of authentic intellectual work.

According to Newmann, authentic intellectual work is comprised of three criteria: 1) *construction of knowledge*, 2) *disciplined inquiry*, and 3) *value of learning beyond school*. The first criterion, construction of knowledge, is based upon the idea that adults produce knowledge through communication, creation, and performance. Therefore, "students should hone their skills and knowledge through guided practice in producing original conversation and writing, repairing and building of physical objects, or performing artistically" (Newmann & Associates, 1996, p. 24). This type of construction of knowledge "involves organizing, interpreting, evaluating, or synthesizing prior knowledge to solve new problems" (Newmann, King, & Carmichael, 2007, p. 4). Students engage in higher-order thinking as they build new ideas and critically evaluate information.

This emphasis on higher-order thinking illustrates how authentic intellectual work reflects critical thinking. Newmann and Wehlage (1993) propose this definition of higher-order thinking:

> Higher-order thinking (HOT) requires students to manipulate information and ideas in ways that transform their meaning and implications, such as when students combine facts and ideas in order to synthesize, generalize, explain, hypothesize, or arrive at some conclusion or interpretation. Manipulating information and ideas through these processes allows students to solve problems and discover new (for them) meanings and understandings. When students engage in HOT, an element of uncertainty is introduced, and instructional outcomes are not always predictable. (p. 9)

This is in contrast to their definition of lower-order thinking:

Lower-order thinking (LOT) occurs when students are asked to receive or recite factual information or to employ rules and algorithms through repetitive routines. As information-receivers, students are given pre-specified knowledge ranging from simple facts and information to more complex concepts. Students are in this role when they recite previously acquired knowledge by responding to questions that require recall of pre-specified knowledge. (p. 9)

The definitions of lower- and higher-order thinking reflect the way that psychologists have conceptualized critical thinking with an emphasis on particular cognitive skills such as synthesis, generalization, and interpretation. At the same time, it draws from a critical perspective in that students should construct new knowledge. The emphasis on manipulating meanings and grappling with uncertainty is consistent with democratic education.

The second criterion, disciplined inquiry, guides the construction of knowledge, as students examine subject matter in particular ways that adults employ in the field, in order to insure that the knowledge is valid and consistent with the field. In particular, disciplined inquiry requires that students "(1) use a prior knowledge base, (2) strive for in-depth understanding rather than superficial awareness, and (3) express their ideas and findings through elaborated communication" (Newmann et al., 2007, p. 4). Here students draw upon a discipline's prior knowledge base such as facts, concepts, theories and vocabularies that comprise a particular field and develop a deep understanding of the discipline so that students can solve specific problems in the field through theorizing and examining relationships among significant disciplinary concepts and ideas. Finally, students present their understandings in a form of communication that adults in the field would employ and in a way that allows students to demonstrate nuanced and complex ideas. Although written essays are the most common form of elaborated communication,

other formats where students demonstrate and articulate complex ideas, such as a debate, artistic rendering, mathematical proof, or scientific research report, also fit into this category. Developing the ability to communicate is central in democratic deliberation.

The third criterion for authentic intellectual achievement is value beyond school. Meaningful intellectual achievements have "aesthetic, utilitarian, or personal value" (Newmann & Associates, 1996, p. 26). That is to say, when adults partake in activities, there is a purpose for the activity that goes beyond simply demonstrating competency, as is often the case with school work. Adults strive instead to do things such as influence an audience, create a product, or convey ideas (Newmann, Marks, & Gamoran, 1995). Often, tests in school only serve the purpose of students demonstrating that they are competent in the subject matter and lack an application. Authentic intellectual work in classrooms occurs within an authentic context. However, value beyond school means more than just "relevant," "student-centered," or "hands-on" activities, as some have suggested, because it is more than just activities that students find interesting "but those involving particular intellectual challenges that when successfully met would have meaning to students beyond complying with teachers' requirements" (Newmann et al., 2007, p. 5). Value beyond school reflects democratic education because it applies to any intellectual challenge that could occur in any aspect of life including as a citizen or as a worker.

Newmann's framework of authentic intellectual work requires that all three criteria – construction of knowledge, disciplined, inquiry, and value of learning beyond school – be present in the activity. Learning opportunities that only include one of these

criteria are not authentic. For example, if students conducted scientific research but did not present their findings in an elaborated way, then the experience would be inauthentic. Students would need to write up and/or present their findings to make the activity authentic. Newmann and Associates (1996) do recognize that not all classroom learning need be authentic. There are instances where students need to be engaged in information retrieval, memorization, and drills in order to build their prior knowledge base or prepare for mandated standardized tests. However, they argue that teachers should strive to create as many authentic learning experiences as possible.

Newmann (1991) argues that authentic intellectual work will motivate students because it is relevant to their lives outside of the classroom. If students are learning about something that has relevance to their lives and the world outside of the classroom then they might be more inclined to care about the problem they are investigating (Newmann, Bryk, & Nagaoka, 2001). The higher-order thinking and problem-solving skills that students gain are also transferable beyond school which adds to the real world relevance. As students integrate skills and facts with their prior knowledge through disciplined inquiry, students internalize and own their new constructed knowledge. These are skills that all citizens need for participating in democratic society but they are also applicable to emerging workplace challenges that require "intellectual competence" to solve new problems and communicate in new ways (Newmann, Lopez, & Bryk, 1998, p. 15). Newmann also draws upon research that positively correlated authentic intellectual work, pupil engagement, and pupil performance. Student surveys from Chicago public schools demonstrated that students were more engaged in classrooms with more authentic work

(Marks, 1995). In turn, the students who were more engaged, had higher academic achievement than students who were less engaged and had lower levels of authentic intellectual work.

To guide teachers in developing authentic instruction and evaluating authentic pupil achievement, Newmann and Associates (1996) developed standards and rubrics to explicitly detail how the three criteria for authentic intellectual work (construction of knowledge, disciplined inquiry, and value beyond school) might look in practice. There are standards that address those criteria for instruction, assessment, and pupil achievement and those standards differ depending upon the type of practice evaluated. Standards for authentic classroom instruction include teaching activities that emphasize higher-order thinking, substantive conversation, deep knowledge, and connectedness to the world. Authentic assessments involve assessment tasks that involve students in constructing knowledge, elaborated communication, and connections to students' lives. Pupils engage in authentic achievement when they analyze disciplinary concepts and communicate their understandings.

Over the years, the categories and criteria comprising the standards have been altered and refined (Newmann et al., 2007; Newmann, Secada, & Wehlage, 1995; Newmann & Wehlage, 1993). In Table 2.1, I show how the standards have been presented in different iterations. The left hand column displays the three criteria for authentic intellectual work (construction of knowledge, disciplined inquiry, and value beyond school) and the columns to the right of the first column list the standards for the criteria according to instruction, assessment, and achievement. There have been two

different sets of standards for instruction. The first iteration, "Authentic Instruction," had

five standards (higher-order thinking, depth of knowledge, substantive conversation,

social support for student achievement, and connectedness to the world). A second

iteration, "Authentic Pedagogy: Classroom Instruction" removed the standard of "social

support for student achievement." The most noticeable difference in the standards applies

to the standards for assessments. In the first iteration, "Authentic Pedagogy: Assessment

Tasks" had seven standards (organization of information, consideration of alternatives,

disciplinary content, disciplinary process, elaborated written communication, problem

connected to the world, and audience beyond school). In a more recent iteration, the

name was changed to "Authentic Assignments" and the seven standards were streamlined

to three (construction of knowledge, elaborated written communication, and connections

to students' lives). The standards for pupil performance, "Authentic Achievement,"

remained unchanged. Changes to the rubrics were made to help in the evaluation of the

standards (Newmann et al., 2007), but the modifications are important because they make

it difficult to compare research that has used the framework but operated under a different

set of scoring standards. The standards articulate a definition of authentic intellectual

work that can be uniformly applied to different assessments and pupil work examples.

Table 2.1

*Iterations of Authentic Intellectual Work Standards*

| CRITERIA | STANDARDS | | | | |
|---|---|---|---|---|---|
| | Authentic Instruction[1] | Authentic Pedagogy: Classroom Instruction[2] | Authentic Pedagogy: Assessment Tasks[2] | Authentic Assignments[3] | Authentic Achievement[2] |
| **Construction of Knowledge** | Higher-Order Thinking | Higher-Order Thinking | Organization of Information<br><br>Consideration of Alternatives | Construction of Knowledge | Analysis |
| **Disciplined Inquiry** | Depth of Knowledge<br><br>Substantive Conversation<br><br>Social Support for Student Achievement | Deep Knowledge<br><br>Substantive Conversation | Disciplinary Content<br><br>Disciplinary Process<br><br>Elaborated Written Communication | Elaborated Written Communication | Disciplinary Concepts<br><br>Elaborated Written Communication |
| **Value Beyond School** | Connectedness to the World | Connections to the World Beyond the Classroom | Problem Connected to the World<br><br>Audience Beyond the School | Connections to Students' Lives | |

*Note.* Categories reproduced from: [1](Newmann & Wehlage, 1993); [2](Newmann, Secada et al., 1995); [3](Newmann et al., 2007).

This study used the framework of authentic intellectual work to analyze assessments tasks that teacher candidates/teachers utilized in their classroom and samples of pupil work on those assessments. However, this study did not use the standards created to evaluate authentic pedagogy. This dissertation used the more recent iterations of the standards for Authentic Assignments and Authentic Achievement and not the standards related to authentic instruction. A further discussion of these standards follows in Chapters 3 and 4.

***Critiques of authentic intellectual work.*** The concept of authentic intellectual

work has been criticized on a number of grounds. For one thing, the use of the word

*authentic* has proved problematic. Terwilliger (1997) argues that using authentic to

describe certain types of assessments implies that those assessments are *real* and *genuine*.

Therefore, authentic assessments privilege themselves over other types of assessments by

claiming they are more real than other assessments. Authentic, in a philosophical sense, is

a misnomer because assessments that do not fall within the definition of Newmann's

authentic intellectual work still exist and are *authentic* in that they are real (Ladwig,

1998). In an attempt to resolve this tension and distance their work from the notion that

there is a "true" or "real" performance, pedagogy, or assessment, researchers using

Newmann's authentic intellectual work framework in Queensland, Australia, refer to

authentic intellectual work as "productive pedagogies" (Hayes, Mills, Christie, &

Lingard, 2006). This dissertation acknowledges that the term authentic may be

problematic and uses the characteristics as described by Newmann to define the term.

The subjective nature of the term authentic has also sparked criticism; what is

authentic to one person may be inauthentic to another (Splitter, 2009). Certainly, as the

discussion in the previous section suggests, there are different definitions for authentic

work and assessment. This raises the question of whether or not it is possible to have

students in school engage in true authentic work. There are also judgments involved in

determining what is representative of a particular field (Terwilliger, 1997).  In his

analysis of Newmann's authentic intellectual work, Splitter (2009) challenges the notion

that students construct new knowledge in authentic work. Since disciplined inquiry is

also a criterion for authentic intellectual work, then students are operating within a given set of facts, concepts, and theories that professionals in the field have deemed accurate for the field. Therefore, students are not producing new knowledge that is authentic to their own understandings if they are re operating within someone else's framework. However, there are boundaries to what is acceptable in a field as accurate information and students need to fit their work within those confines because they will face those constraints as a professional in the discipline (Newmann et al., 2007).

Critics have also questioned whether it is possible to recreate authentic "real world" contexts in a classroom. In their work evaluating the authenticity of a standardized Document-Based Question (DBQ) history exam, Grant, Gradwell, and Cimbricz (2004) argue that even an exam that attempts to replicate the work of historians by asking students to analyze primary sources is still confined by parameters that historians in the profession do not face. The sources and questions have been chosen for and given to the students, which differs from the experience of an actual historian who selects her own research question and sources. Splitter (2009) also contends that disciplines construct their own sets of operating guidelines that confine student thinking within the confines of the discipline. He proposes that students should not try to replicate what happens in the profession but construct "what *ought to be*" in the field (p. 143). However, this critique is weakened when one considers that real world relevance is not the only criteria for authentic intellectual work (Newmann & Associates, 1996). Authentic experiences have value beyond school but also include construction of

knowledge and disciplined inquiry. It is the combination of the three that results in authentic intellectual achievement.

Another criticism is that authentic intellectual work neglects basic skills and specific content knowledge (Terwilliger, 1997). Concerns have been raised that the emphasis on higher-order thinking neglects the teaching of basic skills and that the omission of specific content matter suggests that students will not gain content knowledge. However, this concern ignores the fact that the framework's criterion, disciplined inquiry, requires that students acquire in-depth, conceptual understanding of the discipline, which implicitly includes content knowledge. Complex knowledge does not preclude students from knowing basic skills and content. In fact, as Ladwig (1998) points out, studies have shown that students in authentic classrooms performed "as well or better" than students in inauthentic classrooms on standardized tests of basic skills (Newmann & Associates, 1996, p. 43). In addition, the omission of particular content is purposeful because content is continually being revised and debated. Authentic intellectual work is not dependent upon content or teaching strategies and can be applied across disciplines and grade level.

Other criticisms of the framework of authentic intellectual work have led to new versions of the framework. Researchers in Queensland, Australia added criteria connected to critical pedagogy to Newmann's work when they created their framework of *productive pedagogies* (Queensland School Reform Longitudinal Study (QSRLS), 2001) while researchers in Singapore incorporated elements of curriculum theory and discourse analysis into Newmann's framework in their work on pupil intellectual work (Luke,

Freebody, Shun, & Gopinathan, 2005). A more detailed analysis of these critiques and changes follows in the next section.

   ***Authentic intellectual work and democratic education***. The framework of authentic intellectual work is an appropriate lens to assess the knowledge needed for democratic education because it accounts for the cognitively complex activities that pupils encounter, abilities that are relevant to society, both in terms of citizenship and employment. Newmann and colleagues (1996) are explicit in the relationship between authentic intellectual work and society at large:

> [T]eaching students to master authentic academic challenges should benefit both individuals and the society. The complexities of contemporary society demand that citizens be problem solvers and lifelong learners capable of adapting to changing economic and social conditions. Whether trying to make a living, manage personal affairs, or participate in civic life, citizens are increasingly called on to exercise the kinds of intellectual capacities reflected in authentic achievement. Schools that fail to help students face these challenges deny them opportunities for security, productivity, and fulfillment. (pp. 27-28)

The three criteria for authentic intellectual achievement develop skills necessary for deliberation and encourage knowledge and skills transfer that are required in the contemporary economy. Newmann's framework, however, differs from some aspects of the 21st century skills movement. Authentic intellectual work does not focus on discrete thinking skills; instead, the focus is on thinking in complex ways (Newmann et al., 2007). Authentic intellectual work is more aligned with democratic education. In fact, democracy is explicitly stated as part of the rationale for authentic intellectual work:

> … the argument for democracy assumes that citizens are capable not only of basic literacy, but also of exercising principled and reasoned judgment about public affairs. Arriving at defensible positions on controversial public issues… all require interpretation, evaluation, in-depth understanding, and elaborated

communication that extends well beyond traditional tests of knowledge. (Newmann et al., 2007, pp. 11-12)

The emphasis on interpretation, evaluation, in-depth understanding, and elaborated communication are consistent with what Amy Gutmann (1987) describes as fundamental to deliberative democracy. Citizens need to know how to evaluate differing interests and become effective communicators as they deliberate on issues for the public good.

This dissertation utilizes the framework of authentic intellectual work as a measure of teacher and pupil outcomes in a way that is consistent with democratic education because of the ways in which construction of knowledge, disciplined inquiry, and value beyond school engage pupils in critically examining issues in-depth and developing effective communication skills to engage in dialogue about issues. These skills are fundamental to deliberative democracy. A more detailed discussion of the selection of this framework follows in the research design section.

### Empirical Research on Authentic Intellectual Work

This dissertation examines teacher candidate/teacher and pupil learning through the lens of "authentic intellectual work" (Newmann & Associates, 1996). As the previous section details, this framework was developed during the late 1980s and early 1990s and was closely connected to the school restructuring movement at the time. Since then, researchers have used this framework to measure the intellectual quality of teaching and learning in regards to inservice and pre-service teachers and their pupils. Much of the empirical research utilizing this framework has been embedded within large-scale research projects investigating the necessary school support structures for high quality teaching and learning and centers on in-service teachers involved in these professional

development and school restructuring initiatives. Many of these studies include an examination of assessments and pupil work samples and investigate how factors such as socio-economic-status and prior achievement influence pupils' exposure to authentic intellectual work and their achievement. A few of the studies utilize the framework of authentic intellectual work to examine teacher education and pre-service learning, although there is very little research in this area.

This review includes all of the major research studies using this framework since Newmann and colleagues articulated the framework in the early 1990s. However, certain criteria were used to select the studies reviewed here. First, this review includes rigorous and methodologically sound studies as evidenced by their inclusion in peer reviewed journals. Second, since much of the research has been conducted by various research centers and teams, such as the Center on Organization and Restructuring of Schools, the Queensland School Reform Longitudinal Study, and the Research Institute on Secondary Education Reform for Youth with Disabilities, reports and other publications emanating from these groups have been included as well. Third, some studies have modified the framework of authentic intellectual work, oftentimes utilizing different terminology and criteria for authentic intellectual work. Although the frameworks have significant differences from the framework used in this study, select studies of this type of research are included here because they were derived from and draw heavily upon Newmann's concepts. Finally, this review includes empirical research that provides relevant insight into using authentic intellectual work to evaluate teaching and learning, and more specifically, teacher candidate/teacher and pupil achievement.

This literature review is divided into four parts. First, since a number of studies related to authentic intellectual work were embedded within larger, longitudinal research projects investigating a variety of factors related to teaching, learning, and schools, I provide a brief overview of three research projects that have used the framework at different times and in different ways to examine these issues. The individual studies pertaining to authentic intellectual work related to these large-scale projects are integrated with other empirical work in the literature discussion that follows to better illustrate the overall themes that researchers have addressed and the approaches with which they have done so. Second, I present empirical work that has examined the authentic intellectual work of inservice teachers and their pupils in terms of instruction, assessment, and achievement. This work is sub-divided into three parts based on the ways the studies approached and reported the concept of authentic intellectual work: descriptive studies; cross-sectional effect studies, and longitudinal studies. Third, since this dissertation examines the work of teacher candidates during the preservice period, I include work on authentic intellectual work and teacher education. I conclude with a discussion of the challenges related to this type of research. Each section of the review includes qualitative and quantitative research from diverse methodological perspectives. For the most part, studies are presented chronologically to illustrate the ways in which researchers have drawn upon their own work and the work of others to further investigate these issues.

**Overview of Large-Scale Studies of Authentic Intellectual Work**

A majority of the research on authentic intellectual work was conducted within larger studies on restructuring and improving schools and educational systems. Three research teams worked extensively to develop and investigate this framework: 1) Fred Newmann and colleagues at the Center on Organization and Restructuring of Schools (CORS) and the Consortium on Chicago School Research; 2) James Ladwig, Bob Lingard, and Jenny Gore who worked with the Queensland and New South Wales schools; and 3) Alan Luke and colleagues' work in Singapore at the Centre for Research in Pedagogy and Practice. The work of these three research groups is interconnected and, in some cases, has involved the same researchers. A general overview of these research projects is presented first; detailed findings from these studies follow.

**CORS and Chicago Annenberg Challenge.** Fred Newmann and colleagues at the Center on Organization and Restructuring of Schools (CORS) and the Consortium on Chicago School Research have conducted the majority of the empirical research on authentic intellectual work during the mid- to late 1990s. From 1990 to 1995, CORS engaged in a federally-funded, large-scale research study examining issues related to school restructuring (Newmann & Wehlage, 1995). As part of this initiative, eighteen different studies were conducted, resulting in dozens of reports (Newmann & Associates, 1996). The main study, the School Restructuring Study (SRS), was designed to explore restructuring efforts by studying schools that had adopted policies advocated by school reformers and included 24 public schools in 16 states across the country. Researchers visited each school on two different occasions during the school year for a week at a time

where visits included classroom observations; interviews with teachers, staff, parents, and administrators; teacher surveys; and a collection of teacher assessments and pupil work. To evaluate teaching and learning, Newmann and colleagues developed the criteria and standards for "authentic intellectual work," outlined in the previous section. One of the research questions that SRS addressed, and the one most pertinent to this dissertation, was "How can school restructuring nurture authentic forms of student achievement?" (Newmann & Associates, 1996, p. 8). This study used the framework of authentic intellectual work to measure teaching and learning and evaluate the impact of school restructuring efforts.

Newmann continued his work on authentic intellectual work school restructuring as part of the Consortium on Chicago School Research funded by the Chicago Annenberg Challenge grant (Newmann, Lopez et al., 1998). From 1996 to 2001, researchers studied 400 classrooms in 19 different Chicago schools. The ultimate purpose of this study was to improve student achievement by deepening school and community relationships and reforming teaching and learning. As with the SRS study, teaching and learning were evaluated using the criteria of authentic intellectual work. The empirical research that came out of this work, as well as the CORS work, examined the relationship between school restructuring efforts and pupil learning as measured by the framework of authentic intellectual work.

**Queensland School Reform Longitudinal Study and New South Wales.**
Researchers in Queensland, Australia, built on Newmann's authentic intellectual work framework in their work on the Queensland School Reform Longitudinal Study (QSRLS)

(2001). Under the co-direction of James Ladwig and Bob Lingard, this government-commissioned research initiative acted as the foundation for the restructuring of Queensland schools where the main focus of the research was to explore what institutional and classroom practices will lead to better academic and social performance for all pupils (Lingard, Hayes, & Mills, 2003). Researchers involved in the QSRLS were interested in the foundation of authentic intellectual work, but concluded that Newmann's framework did not fully consider issues of critical pedagogy. As a result, QSRLS researchers expanded upon Newmann's work to incorporate dimensions that "recognize that intellectual outcomes are also characterized by social factors that render some knowledge problematic" and was informed by literature in "sociology of school knowledge, school effectiveness and ethnographies of classroom discourse" (QSRLS, 2001, pp. 4, 6). Their new framework, "productive pedagogies," added elements that addressed issues of problematic knowledge, cultural practices, and issues related to the context of Australian schools, while retaining the fundamental elements of authentic intellectual work (Hayes et al., 2006).

There are four underlying dimensions of productive pedagogies: 1) *intellectual quality*, 2) *connectedness*, 3) *socially supportive classroom environment*, and 4) *recognition of difference* (QSRLS, 2001). The first two dimensions reflect concepts from the three authentic intellectual work criteria construction of knowledge, disciplined inquiry, and value beyond school. The third dimension draws upon the concept of social support for learning, a concept that Newmann had included in early iterations of authentic work. The fourth dimension incorporates issues of critical pedagogy and social outcomes,

where teaching and learning address issues of citizenship, cultural knowledges and group identities in learning communities. Although some of these concepts, such as citizenship and social support for learning, underpin the theory of authentic intellectual work, they did not become explicit criteria. These concepts, however, are still relevant to authentic intellectual work. Productive pedagogies emphasize transformative teaching. Similar to the authentic intellectual work framework, there are different items that address these four dimensions with regard to instruction, assessment, and achievement, although the terms have been changed to "productive pedagogies," "productive assessment tasks," and "productive performance," respectively.

Table 2.2, reproduced from Hayes, Mills, Christie, and Lingard (2006, pp. 22-23), lists the productive pedagogies dimensions and items for instruction (productive pedagogies), assignments (productive assessment tasks), and pupil achievement (productive performance). As can be seen by the individual items, the main concepts of Newmann's authentic intellectual work (marked with an asterisk on the table) are included in the item dimensions, although they have been disaggregated. For example, Newmann's criterion, construction of knowledge, was separated out to have separate items for construction of knowledge (problematic knowledge), consideration of alternatives (problematic knowledge), and higher-order thinking, all under the productive pedagogy dimension *intellectual quality*. By disaggregating the category, researchers are able to look at cognitive processes in a more nuanced way than Newmann's work. The item names for productive pedagogies also reflect the influence of critical pedagogy; here "problematic" knowledge is based on articulating the assumption that knowledge is

constructed, not fixed (Hayes et al., 2006). The items also include standards from an earlier iteration of authentic pedagogy (Newmann, Secada et al., 1995). For example, the framework for authentic pedagogy had separate standards for "disciplinary content" and "disciplinary processes," which were later streamlined into one standard "disciplinary concepts" (Newmann et al., 2007). In productive pedagogies, "Productive Assessment Tasks" includes items for "Depth of Knowledge: Disciplinary Content" and "Depth of Knowledge: Disciplinary Processes." Productive pedagogies draws heavily from Newmann's framework.

Table 2.2

*Dimensions and Items of Productive Pedagogies*

| Dimensions | Items | | |
| --- | --- | --- | --- |
| | Productive Pedagogies | Productive Assessment Tasks | Productive Performance |
| Intellectual Quality | Problematic Knowledge* | Problematic knowledge: construction of knowledge* | Problematic knowledge* |
| | | Problematic knowledge: consideration of alternatives* | |
| | Higher-order thinking* | Higher-order thinking* | Higher-order thinking* |
| | Depth of knowledge* | Depth of knowledge: disciplinary content* | Depth of understanding* |
| | Depth of students' understanding* | Depth of knowledge: disciplinary processes* | |
| | Substantive conversation* | Elaborated communication* | Elaborated communication* |
| | Metalanguage | Metalanguage | |
| Connectedness | Connectedness to the world beyond the classroom* | Problem connected to the world beyond the classroom | Connectedness to the world beyond school* |
| | Knowledge integration | Knowledge integration | |
| | Background knowledge | Link to background knowledge | |
| | Problem-based curriculum | Problem-based curriculum | |
| | | Audience beyond school* | |
| Supportive Classroom Environment | Students' direction | Students' direction | |
| | Explicit quality performance criteria | Explicit quality performance criteria | |
| | Social support* | | |
| | Academic engagement | | |
| | Student self-regulation | | |

| Dimensions | Items | | |
| --- | --- | --- | --- |
| | Productive Pedagogies | Productive Assessment Tasks | Productive Performance |
| Working With and Valuing Difference | Cultural knowledges | Cultural knowledges | Cultural knowledges |
| | Active citizenship | Active citizenship | Responsible citizenship |
| | | | Transformative citizenship |
| | Narrative | Narrative | |
| | Group identities in learning communities | Group identities in learning communities | |
| | Representation | | |

*Note*. Adapted from (Hayes et al., 2006, pp. 22-23).

The QSRLS (2001b) research design replicated, in many ways, the CORS research. Researchers used backward mapping to account for factors that would influence pupil achievement consisting of pedagogy, school organizational capacity, and external systemic supports. From 1998 to 2000, data were collected in 24 schools (primary and secondary grades), which were involved in restructuring efforts and involved classroom observation (*n*=975), teachers' questionnaires, staff interviews, and collection of assessment tasks and pupil work samples. At the completion of this study, a second research project, "The New Basics Trial," sought to implement productive pedagogies into 38 Queensland schools (The State of Queensland Department of Education and the Arts, 2004).

Researchers involved in the work in Queensland extended their work on authentic intellectual work in New South Wales (NSW) and the longitudinal study *Systemic Implications of Pedagogy and Achievement for NSW Public Schools* (SIPA) (Ladwig & King, 2003). There, researchers modified "productive pedagogies" and renamed the new framework "quality teaching." The quality teaching model has three dimensions: 1)

*intellectual quality*, 2) *quality learning environment*, and 3) *significance*. Although the

terminology changed, the concepts of construction of knowledge, disciplined inquiry, and

value beyond the classroom remained in quality teaching. The SIPA study examined "the

relationships between teacher learning, the quality of pedagogy, and the quality of

outcomes for students, to explore implications of these findings for the structure and

substance of teacher education programs (with particular interest here in preservice

teacher education)" (Gore, Ladwig, Griffiths, & Amosa, 2007, p. 1). Between the years

2004 and 2007, this study followed 3 cohorts of students (*n*=3000) through classroom

observations and analysis of assessment tasks and pupil work in terms of quality

teaching.

**Centre for Research in Pedagogy and Practice Core Research Project.**

Newmann's framework of authentic intellectual work and the research in Queensland

also informed researchers in Singapore at the Centre for Research in Pedagogy and

Practice (Luke et al., 2005). Under the direction of Alan Luke, a scholar involved with

the Queensland research, the Centre's Core Research Project created their own

framework to measure the authentic intellectual quality of teaching and learning in their

large-scale study on the "factors that contribute to educational success and outcomes in

Singapore schools" (p. 14). The Core Research Project involved six different studies and

included classroom observations, interviews, surveys, and the collection and analysis of

assessment tasks and pupil work. As Table 2.3, adapted from Koh et al. (2005),

illustrates, the framework includes elements from authentic intellectual work and

productive pedagogies such as depth of knowledge, knowledge criticism, sustained

writing, and connection to the real world, but includes elements drawn from curriculum

theory and discourse analysis. Here, attention is paid to "the representation and

scaffolding of knowledge, using a range of scales to examine epistemological sources of

knowledge, disciplinary framing, depth of disciplinary concepts and discourse,

knowledge reproduction and construction, and levels of critique" (Luke et al., 2005, p.

19). These additions allow researchers to account for support structures in teacher

practices.

Table 2.3

*Assessment Tasks and Pupil Work Standards for the Centre for Research in Pedagogy and Practice*

| **Standard** | |
|---|---|
| **Teacher's Assessment Tasks** | **Student Work** |
| Depth of Knowledge | Depth of Knowledge and Understanding |
| Factual Knowledge | Factual Knowledge |
| Procedural Knowledge | Procedural Knowledge |
| Advanced Concepts | Advanced Concepts |
| Knowledge Criticism | Knowledge Criticism |
| Presentation of Knowledge as Given | Presentation of Knowledge as Given |
| Compare and Contrast Knowledge | Compare and Contrast Knowledge |
| Critique of Knowledge | Critique of Knowledge |
| Knowledge Manipulation | Knowledge Manipulation |
| Reproduction | Reproduction |
| Organization, Interpretation, or Evaluation of Information | Organization, Interpretation, or Evaluation of Information |
| Application/Problem-Solving | Application/Problem-Solving |
| Generation/Construction of Knowledge New to Students | Generation/Construction of Knowledge New to Students |
| Sustained Writing | Sustained Writing |
| Connections to the Real World Beyond the Classroom | Connections to the Real World Beyond the Classroom |
| Supportive Task Framing | Quality of Student Writing/Answers |
| Structure of the Task | |
| Content Scaffolding | |
| Procedural Scaffolding | |
| Strategy Scaffolding | |
| Clarity and Organization | |
| Learner Support | |
| Student Control | |
| Explicit Performance | |
| Standards/Marking Criteria | |

*Note*: Adapted from (Koh et al., 2005).

The framework of authentic intellectual work has evolved, and although the work

in Australia and Singapore uses variations on Newmann's framework, this literature

review includes these iterations together very intentionally. The later work is based heavily on the concepts of authentic intellectual work as articulated by Newmann and colleagues. The rubrics are different and the terminology is different, but the fundamental concepts such as construction of knowledge, disciplined inquiry, and value beyond school remain. The drawback, however, is that the different rubrics and terminology make it difficult to compare the levels of authentic intellectual work across the research projects because the rubrics and the subsequent ratings of the quality of work are inconsistent with each other. Second, many of the researchers involved in this work were involved with research that preceded it. Third, these research studies address similar research questions and share similar research design methodology. These studies all used the qualities of authentic intellectual work to examine and evaluate instruction, assessment, and pupil achievement. The studies address different aspects of teaching and learning as well. For example, Newmann's framework has primarily been used in restructuring efforts with inservice teachers whereas the work in Queensland and New South Wales applied their framework to teacher education. Finally, each study observed classroom practice and analyzed teacher assessment tasks and pupil work samples, straying away from using conventional, standardized multiple choice tests to measure outcomes. Since these studies approach the evaluation of pupil learning in ways similar to authentic intellectual work, this dissertation draws from some of the methodological approaches these studies have used and uses these findings to provide insight into this dissertation's findings.

**Research on Authentic Intellectual Work and Inservice Teachers**

The aforementioned large-scale research projects, as well as a number of smaller studies, form the basis for this literature review on the ways that researchers have used the framework of authentic intellectual work to examine teaching and learning. In their review of the literature on this topic, Ladwig, Smith, Gore, Amosa, and Griffiths (2007) categorize analyses of this work into three types: 1) descriptive studies of the intellectual quality of instruction, assessment, and achievement; 2) cross-sectional effect studies analyzing the impact of authentic intellectual work on pupil outcomes; and 3) longitudinal studies examining the effect of authentic intellectual work on pupil gains on national achievement tests. Here I use Ladwig and colleagues' categories to categorize the literature on inservice teachers and their pupils.

**Descriptive studies**. One way in which researchers have utilized the framework of authentic intellectual work is to describe the quality of authentic intellectual work found in classrooms. These descriptive studies typically involve the use of a scoring rubric based on the framework to quantify the levels of authentic intellectual work found in instruction (through classroom observation), learning opportunities (assessment tasks), and pupil achievement (pupil work). In these descriptive studies, researchers have reported the levels of intellectual work in terms of subject matter, gender, grade, and scores on the different criteria. There are also descriptive studies on pupils' views of authentic work.

As part of the CORS study, researchers sought to determine the authenticity of teachers' practices by asking, "To what extent do teachers in restructured schools offer

authentic pedagogy according to these standards and how much variation is there

between teachers, schools, grade levels, and subjects?" (Newmann et al., 1996, p. 292).

They examined this question by looking at math and social studies classes for grades 4-5

and 7-10 from the 23 public schools involved in the CORS study throughout the country.

Three math and three social studies classes from each grade were selected for this study

and at least one teacher from each subject area was involved in a reform initiative at their

school. Each teacher submitted two assessment tasks that were used in their classes

($n$=234), one used in the fall and one used in the spring, and completed a questionnaire

about the context in which the assessment task was administered. Teachers also

submitted a class set of pupil work on the assessment that they submitted ($n$=3128) and

pupils ($n$=2128) completed a questionnaire about their thoughts on the task and their

work. Researchers used the rubric for authentic pedagogy and authentic student

performance to score the assessment tasks and pupil work samples. The mean scores for

both pedagogy and performance were lower than the rubric's midpoint where "even in

restructured schools, pedagogy was rarely rated at the higher levels" of the scale (p. 296).

Researchers concluded that teachers and pupils were not engaging in high levels of

authentic intellectual work. They also used descriptive statistics to describe differences in

the quality of authentic intellectual work between grade level and subject matter. Pupils

engaged in higher levels of authentic performance in social studies than in math, although

there were no statistical differences between social studies and math in the authentic

pedagogy scores. There was also some variation in pedagogy and achievement scores in

social studies across grade levels, with middle and high school pupils achieving higher

than elementary pupils, but math scores remained stable across grade level. Authentic intellectual work was not a common practice in schools.

Similarly, researchers conducted a descriptive study of schools involved in the Chicago Annenberg grant, although the research design differed slightly from the CORS research. Newmann, Lopez, and Bryk (1998) collected data from 12 Chicago public schools to evaluate the quality of authentic intellectual work. These schools were involved in reform measures although they were representative of Chicago public schools as a whole. Two teachers in grades 3,6, and 8 submitted two assessments that represented "typical" assessments used in their class and two assessments that they considered to be "challenging" in language arts and math. The teachers in this study were aware of the Annenberg Challenge but did not know the standards for authentic intellectual work. In total, researchers collected 349 assessments from 74 teachers and collected 1864 samples of pupil work in writing and 1436 samples of pupil work in math. Researchers scored the assessment tasks and a sample of the pupil work ($n$=965 writing, $n$=727 math) according to the standards for authentic pedagogy and achievement. The scores were scaled, using Rasch analysis, so that the samples fell into four categories, from lowest to highest quality: *no challenge/no authentic intellectual work*; *minimal challenge*; *moderate challenge*; and *extensive challenge/extensive authentic intellectual work*. Examples of assessments and pupil work were presented in the report to describe what authentic intellectual work looked like and illustrate how samples were scored. Researchers found that the majority of the writing and math assignments in all three grades were scored as "no challenge" or "minimal challenge," the lowest two out of four categories. They also

found that the assessments categorized as "challenging" by the teachers scored higher than the "typical" assessments and found that assessments and pupil work in writing scored higher than assessments and pupil work in math.

In addition, this study (Newmann, Lopez et al., 1998) also used the framework of authentic intellectual work to analyze the curriculum frameworks for Chicago public schools to see the level of intellectual work that teachers were expected to engage in. In this analysis they found that the curriculum frameworks had more instances of construction of knowledge and disciplined inquiry than connections to things outside of school. As a result, they recommended that teachers relate more assignments to explore significant personal or social topics in ways that demand construction of knowledge and disciplined inquiry. They also suggested that school reform at all levels support teachers to become more authentic in their teaching.

The National Writing Project (The Academy for Educational Development, 2002) used the standards for authentic intellectual work to evaluate assessments and pupil work as part of their longitudinal study of 36 third and fourth grade teachers and their pupils ($n$=1914) in five different states on the impact of a professional development program on pupil achievement in writing. Researchers collected two writing assessments from each teacher ($n$=154) and pupil responses to those assessments ($n$=1992) over a three year period. Assessment tasks were scored on a 3-point scale for three criteria: construction of knowledge, disciplined inquiry, and value beyond school. Pupil work was scored on a 3-point scale according to three criteria: construction of knowledge; organization and coherence; and usage, mechanics, and spelling. The report for the National Writing

Project claimed that a "majority" of assessments provided opportunity for authentic intellectual work and that the "majority" of pupils produced "at least some" level of authentic intellectual work. However, researchers in the report conflated the top two scores on the 3-point scale and combined the number of students who received the two highest ratings for each criterion to categorize those pupils as the "majority." Although the National Writing Project used a different scale than the CORS and Chicago Annenberg Challenge studies and it is difficult to directly compare the studies, based on the small percentage of assessment tasks and pupil work samples that scored the highest rating, it is reasonable to suggest that teachers and pupils involved in the National Writing Project professional development program produced moderate levels of authentic intellectual work.

One purpose of the Core Research Project conducted by the Centre for Research in Pedagogy and Practice in Singapore was to examine assessment tasks and pupil work (Luke et al., 2005). Researchers collected assessment tasks and pupil work from 36 schools (18 primary and 18 secondary) in seven subjects (English, social studies, math, science, and Chinese, Malay and Tamil languages) (Koh et al., 2005). From this sample, Koh et al. analyzed assessments and pupil work in English, social studies, math, and science from grade levels primary 5 and secondary 3. Teachers submitted assessments ($n$=170) that were embedded within their classroom instruction and those lessons were observed by researchers. Each teacher submitted what they considered to be four high quality, four medium quality, and four low quality pupil work samples ($n$=1749).

Assessment tasks and pupil work samples were scored on a 4-point scale according to the intellectual work standards described above for the Core Research Project.

This study found that assessments were low in intellectual quality and consisted primarily of multiple choice and short answer workbook and worksheet pages, with some opportunity for extended writing. When describing the range of scores on the individual rubric items, researchers found that there were more opportunities for students to reproduce factual knowledge than engage in authentic criteria such as "organization, interpretation, or evaluation of information;" "application/problem-solving;" and "generation/construction of knowledge new to students." The intellectual quality of pupil work was also relatively low, with the exception of primary social studies, which had higher scores when compared to the other subjects in the criteria advanced concepts, critique of knowledge, interpretation, problem-solving, construction of new knowledge, and connections to the real world. Researchers concluded that Singapore teachers were more inclined to use traditional assessments as opposed to authentic assessments.

There are two main findings from these descriptive studies. First, overall, for the most part, these descriptive studies found that the authentic intellectual quality of pedagogy and achievement was relatively low. The majority of teachers in these studies did not design assessments that required high levels of authentic intellectual work and pupils did not engage in high levels of this type of work. Authentic intellectual work was not the norm. This is supported by additional studies described in the following sections that also found low levels of authentic intellectual work (Bryk et al., 2000; Gore et al., 2004; Ladwig et al., 2007).

Second, certain subjects may represent authentic intellectual work more than others. In Newmann, Marks, and Gamoran (1996), social studies had higher levels of authentic work than math and in Newmann, Lopez, and Bryk (1998), assessments and pupil work in writing were more likely to be authentic than math. One possible explanation is that since one aspect of authentic intellectual work is elaborated written communication, there are fewer opportunities to express ideas in that matter in traditional math assignments. However, that may not be the case. For one thing, the criterion, elaborated written communication, in math assessment tasks include "tasks that ask students to generate prose (e.g., write a paragraph), graphs, tables, equations, diagrams, or sketches" and diagrams or graphic representations for pupil work (RISER, 2001, p. 11). Therefore it is possible for math assessments and work to score high. There are also indicators that it is possible to construct authentic assessments and pupil work in subjects that do not generally emphasize standard writing. In a preliminary study for the New South Wales SIPA study, researchers used the Quality Teaching framework to analyze 73 assessment tasks for secondary Physical Development, Health and Physical Education classes to examine the authentic intellectual quality of physical education classes (Gore, Ladwig, Amosa, & Griffiths, 2008). Researchers found that the physical education assessments tasks scored higher than assessments in English, math, science, and Human Society and Its Environment. However, this is further complicated by the fact that the study also found that classroom instruction in physical education ($n$=42) had lower quality teaching scores for observed classroom instruction than the other subjects, although researchers accounted for this difference based on the challenge physical

education teachers have implementing theory into practice. This suggests that there may be conditions other than subject matter that account for differences in the quality of authentic intellectual work. This dissertation uses interview data to further explore conditions and contexts that may contribute to teachers' ability to engage their pupils in authentic intellectual work.

The descriptive studies depicted above involve inservice teachers in schools where there were structures or professional development opportunities promoting practices that were consistent with authentic intellectual work. These teachers were not prepared in initial teacher education preparation to engage in this type of work. Therefore, one question that this raises is whether or not teachers in a teacher education program that focuses on pupil learning in ways that are consistent with authentic intellectual work would produce higher levels of authentic intellectual work than these teachers. This study examines preservice teachers to see if this is the case.

**Cross-sectional effect studies**. A second way that researchers have explored the topic of authentic intellectual work has been through cross-sectional effect studies on the relationship between authentic intellectual work and a variety of topics including pupil engagement, teacher efficacy, pupil achievement, assessment tasks and pupil performance, SES/race/prior achievement, the achievement of pupils with disabilities, and professional development/school reform. These analyses go beyond describing the authentic intellectual quality of teaching and learning and utilize quantitative approaches to account for differences among groups and to measure the effect of authentic intellectual work on different factors.

***Authentic intellectual work and pupil engagement***. Researchers have examined

the relationship between authentic intellectual work and pupil engagement as measured

by pupil effort and interest in learning. The CORS study paid specific attention to how

authentic intellectual work influenced pupil engagement with subject matter (Marks,

1995). Marks surveyed 144 math and social studies teachers in grades 5, 8, and 10 from

the 24 schools across the country that had organizational structures and practices

associated with school restructuring. Teachers were asked questions about their

instructional goals, teaching practice, curriculum, and background to determine the

quality of authentic intellectual work in their practice. Their pupils ($n$=3660) were also

given a survey about their attitudes, behaviors, and experiences in the classroom. Marks

found that authentic practices had more of a positive effect on middle and high school

pupils than on elementary pupils. Middle and high school pupils who were exposed to

higher quality intellectual work were more likely to report factors that indicated they

were more engaged in school, whereas for elementary pupils, there was no significant

relationship between authentic intellectual work and engagement. This research has been

used to suggest that pupils who are engaged in authentic intellectual work are more

interested in their work and schooling. There is also evidence that suggests teachers took

more responsibility and efficacy for pupil learning when they had higher authentic

intellectual pedagogy score (Gore, Griffiths, & Ladwig, 2002).

This research is supported by one study of how pupils perceive authentic

intellectual work and learning. Van't Hooft (2005) surveyed 99 seventh grade science

students about their experiences in a technology-rich science class. Specifically, van't

Hooft asked, "What is the effect of participation in technology-infused projects such as the Ohio Schools Going Solar (OSGS) project on student perceptions of learning science, specifically focusing on perceptions related to disciplined inquiry, construction of meaning, and application beyond the classroom?" (p. 225). In this classroom, pupils took part in lessons where they engaged in activities such as conducting research and using the research to build a device powered by an alternative energy source and teaching fifth graders about electrical circuits. Lessons engaged pupils in the underlying criteria for authentic intellectual work and class assessment tasks were scored "extensively challenging" according to the framework's standards. For two different units, pupils were given a survey prior to the start of each unit to register a baseline for their ideas about authentic intellectual work and science and were then given a second survey after each unit about the same concepts. Van't Hooft concluded that pupil engagement in authentic learning in science had "a positive effect on students with regards to their perceptions of learning science, especially when it comes to disciplined inquiry and construction of meaning" (p. 232). Pupils in this case perceived authentic intellectual work as more engaging.

*Authentic intellectual work and pupil achievement*. One of the underlying assumptions in the framework of authentic intellectual work is that when pupils engage in this kind of work, their achievement is higher. Therefore, there is great interest in the ways that classroom practices, primarily instruction and assessment, influence pupil outcomes. Researchers have approached this by examining the relationship of authentic

pedagogy, authentic pupil achievement, and pupil outcomes on other standardized measures.

Researchers in the CORS study (Newmann et al., 1996) measured the relationship between authentic pedagogy and authentic achievement, as defined by the standards for authentic intellectual work. Using a three-level Hierarchical Linear Model (HLM) analysis to control for pupil social background and prior achievement, Newmann et al. measured the effect of authentic pedagogy scores on pupil authentic achievement scores based on the scoring of 234 teacher assessment tasks and 3128 samples of pupil work. In this analysis, authentic pedagogy accounted for "about 35% of the variance among classrooms in student performance" (p. 299). That is to say, only about one-third of student performance could be attributed pedagogy. One weakness of this study is that the "design did not allow the strongest possible test of the link between pedagogy and performance" (p. 300) because pupils did not have a common assessment task or curriculum across grades and classes. Pupil performance was also constrained by the limits of the assessment task; that is to say, if pupils were given a low quality assessment task, they had fewer opportunities to produce high quality work. It is very difficult for pupils to demonstrate high levels of authentic work if they are not asked to do so (Newmann, Lopez et al., 1998). However, what this indicates is that teaching practices and assignments matter in pupil performance.

Newmann, Lopez, and Bryk (1998) supported the preceding finding with their work on the Chicago Annenberg Research project. They investigated the statistical correlation between authentic assessment scores for 349 math and language arts

71

assessments and the authentic performance score on pupils' responses to those

assessments (*n*=1692). They found that a strong, statistically-significant, positive

correlation existed between the two, suggesting that higher quality assessment tasks were

related to higher quality pupil performance. Similarly, correlations between assessment

task scores and pupil achievement scores in two other studies supported this finding

(Clare & Aschbacher, 2001; King, Schroeder, & Chawszczewski, 2001; The Academy

for Educational Development, 2002). These studies make the case that assessments

matter in pupil outcomes.

Supporting Newmann and colleagues' findings related to the CORS and

Annenberg research, Ladwig, Smith, Gore, Amosa, and Griffiths (2007) replicated the

research design with data from the New South Wales SIPA study for an international,

cross-cultural comparison. Although they altered the terminology slightly, researchers

used the same measures of authentic intellectual work used by Newmann, Marks, and

Gamoran (1996). Ladwig et al. utilized a two-level MLWin (v2.02) to control for the

effect of gender, language background, race, SES, and prior achievement and measured

the effect of authentic assessment tasks (*n*=78) on pupil performance on those assessment

tasks (*n*=2236). Teachers in this study had been involved in professional development

work related to authentic intellectual work. As with the CORS research, this study found

that assessment tasks had a positive and statistically significant effect on authentic

achievement, although the effect in the SIPA study was smaller than the effect found in

the CORS study. One difference between the two studies was that the CORS study used a

combined authentic pedagogy score that included a classroom observation and

assessment task score, whereas the SIPA study only included an assessment task score; Ladwig et al. suggest that may be one reason the effect sizes were different. However, the findings do suggest that even in contexts outside of the United States, authentic pedagogy influences the level of pupil performance and that authentic intellectual classroom practices and assessments should be promoted.

The above studies used classroom-specific measures of authentic pedagogy and pupil performance, meaning teachers and pupils in different classrooms in the study were not evaluated on the same assessment. Although this construct strengthens classroom validity because the evaluated assessment was part of the normal classroom routine, the use of different assessment tasks across classes weakens the ability to determine a relationship between assessment task and pupil performance. In an attempt to provide evidence for a stronger correlation, Avery (1999) examined the relationship between authentic instruction and pupil performance by using a common assessment task with five different teachers and their 116 pupils (predominantly 11th graders). Avery found a strong, positive correlation ($r=.686$) between authentic instruction (via classroom observation) and authentic pupil performance. Using a statistical model to control for demographics, pupil engagement, and instruction, Avery found that authentic instruction accounted for 40% of the variation on pupil performance, and argued that authentic instruction is a better predictor of pupil performance than demographics and engagement. This further supports, and strengthens, previous assertions that authentic pedagogy is critical to higher authentic pupil performance.

Since many of these studies have used assessment tasks as a measure of teacher quality, researchers have also looked at whether or not this is a valid measure. Researchers at the National Center for Research on Evaluation, Standards and Student Testing (CRESST) were involved in a 4-year Annenberg research grant to explore the use of assessments and pupil work as indicators of classroom practice (Clare, 2000; Clare & Aschbacher, 2001; Matsumura, 2003). This 4-year research study focused on four elementary and four middle schools in Los Angeles that were part of an Annenberg Challenge school restructuring initiative. Researchers collected language arts assessments from third and seventh grade teachers and samples of pupil work on those assessments; teachers submitted a cover sheet for each assessment where they detailed their learning goals and evaluation criteria. Researchers also observed classroom lessons and conducted 15 minute interviews with the teachers about the observed lessons. During the fourth year of the study CRESST piloted an assignment rubric in the Los Angeles Unified School District. They scored classroom instruction, assessment tasks, and pupil work according to a rubric that included indicators of authentic intellectual work and conducted statistical analyses on the scores. Although CRESST modified the authentic intellectual work rubric, this work is pertinent because it looks at the validity of using assessment tasks and pupil work as measures of classroom practice and pupil achievement. Although they investigated a number of different research questions, one research focus pertinent to this dissertation is their analysis of the number of assignments needed to estimate the quality of classroom practice. Clare and Ashbacher (2001) conducted a generalizability study and concluded that at least two assessments were required to estimate teacher quality. They

also suggest that "typical" assessments were a better indicator to measure pupil performance because they were more likely to correspond with "typical" instruction inside the classroom. This implies that examining assessments is a valid way to assess teacher practices.

Critics of authentic intellectual work are skeptical about the use of assessment tasks and pupil work as measures of pupil performance in favor of more standardized evaluative measures (Tanner, 2001; Terwilliger, 1997), but research on the relationship between authentic intellectual work and pupil performance as measured by standardized tests suggests otherwise. Newmann, Bryk, and Nagoaka (2001), as part of the Chicago Annenberg Challenge study, compared authentic assessment and achievement scores to pupil test scores on the Iowa Test of Basic Skills (ITBS) and the Illinois Goals Assessment Program (IGAP). Over a three year period they collected and scored assessment tasks (*n*=2017) according to a rubric for authentic intellectual work. Using a Many-Facets Rasch Analysis, researchers controlled for raters, grade, and subject. They then examined the effect of authentic intellectual quality scores to pupil tests scores on the ITBS and IGAP in reading and math, through an HLM analysis that controlled for race, gender, and SES. Newmann et al. found "a consistent positive relationship between student exposure to high-quality intellectual assignments and students' learning gains on the ITBS" per year (p. 22). Since the IGAP assessment was not administered for each grade level, researchers presented a "value-added" effect where pupils in highly authentic classrooms were more likely to outperform other pupils on the IGAP reading, writing, and math tests with effect sizes that ranged from .43 and .64, which are quite substantial.

Pupils did well on basic skills tests even though they were not focusing on basic skills. The higher level thinking that authentic work produced did not detract from, and most likely only helped, pupils' knowledge of basic skills. This refutes critics who claim that authentic intellectual work ignores basic skills and suggests that pupils exposed to authentic intellectual work do better on tests of basic skills than their peers.

These studies illustrate that authentic instruction has a positive impact on authentic performance and that using classroom instruction, assessment tasks, and pupil work samples are strong indicators of classroom practice. When students are exposed to highly authentic intellectual work, they are more likely to produce highly authentic work. The skills and knowledge that pupils gain in this type of work also transfer to gains on standardized tests. In one regard the strong relationship between assessment and achievement scores is rather intuitive considering that the rubric for authentic instruction/pedagogy and achievement share a conceptual framework. That is to say, the criteria for authentic achievement are dependent upon the criteria for authentic pedagogy. However, "merely asking for authentic intellectual work offers no guarantee that all students will succeed in producing it" and requires instructional support (Newmann, Lopez et al., 1998, p. 36). There are additional conditions and contexts that support this type of work as well. The emphasis on instruction requires an examination of why and how teachers make instructional choices. This dissertation looked at classroom context and teachers' goals for pupil learning to see how they might influence how teachers design learning opportunities that reflect authentic intellectual work.

***Authentic intellectual work and SES, race, and prior achievement***. In addition

to the effect of authentic instruction on authentic pupil performance, researchers have

also examined the relationship between authentic intellectual work and factors such as

socio-economic status (SES), race and prior achievement. Many of these studies have

been previously mentioned, but their findings in regards to these topics are included here

to illustrate how studies have examined these factors.

For the most part, the CORS and Chicago Annenberg Research projects found no

differences in authentic intellectual work with regard to SES and race. Newmann, Marks

and Gamoran (1995) found that students from different SES backgrounds had "equal

access to authentic pedagogy" (p. 8) while the Chicago Annenberg Research also found

instances of high quality authentic pedagogy distributed equitably throughout even the

most economically disadvantaged schools (Newmann et al., 2001). Overall, there were

very low levels of intellectual work throughout all of the schools, and classroom

variations in authentic intellectual quality were not due to socio-economic status or the

racial composition of the classes. That was not the case, however, in some Australian

schools. The New South Wales SIPA study found that pupils from low SES background

and Indigenous descent were more likely to receive lower quality productive pedagogy

than their white, advantaged peers (Amosa, Ladwig, Griffiths, & Gore, 2007; Gore et al.,

2007; Griffiths, Amosa, Ladwig, & Gore, 2007). At the school level, there were no

differences between the quality of tasks at high and low SES schools but there were

differences at the classroom level of the quality of assessment tasks given to low SES and

minority pupils when compared to the quality of assessment tasks given to high SES and

majority pupils (Griffiths et al., 2007).  Amosa, Ladwig, Griffiths, and Gore (2007) did find, however, that when minority pupils were given higher quality intellectual work they performed at a higher level, thereby reducing the achievement gap. The gap was reduced further when pupils were supported by authentic instruction in the classroom and had teachers who emphasized high expectations.

On the surface, the fact that SES and prior achievement had no effect on authentic intellectual work seems surprising and refutes research that suggests that pedagogy in disadvantaged schools is lower-order and watered-down. One possible explanation for the CORS findings is that the disadvantaged schools involved in the study had been working to bring high quality teaching to disadvantaged schools (Ladwig et al., 2007). Ladwig et al. suggest that this implies that high quality authentic intellectual work is possible in disadvantaged schools as long as the schools provide it. Their findings of differences in particular populations in Australian schools suggest that there may be more to understand about the conditions. Another possible explanation, especially since the levels of authentic intellectual work were low, is that schools in general are not providing high quality learning opportunities to their pupils and so in a sense there is a level playing field because very few students are engaged in high quality work.

The factor that appears to be more of a predictor of which students will be exposed to high quality authentic intellectual work is prior achievement. Both the CORS and Chicago Annenberg Challenge research projects found that pupils who had higher levels of prior academic achievement (as measured by NAEP pre-test scores for the CORS study and ITBS pre-test scores for the Chicago study) were more likely to receive

higher levels of authentic pedagogy in the classroom (Newmann et al., 2001; Newmann et al., 1996). This achievement gap shrank, however, when low-achieving pupils received authentic pedagogy and these pupils were more likely to improve at a greater rate than already high-achieving pupils. This is significant because it suggests that good teaching, as evidenced by high quality assessments, has an impact on achievement. The SIPA study found similar conditions (Ladwig et al., 2007); higher-achieving pupils in New South Wales were also more likely to receive higher intellectual quality assessment tasks and productive pedagogy than their low-achieving peers. This work suggests that all pupils, particularly low achieving and minority pupils, can produce high levels of authentic intellectual work when they are exposed to that type of work.

*Authentic intellectual work and students with disabilities*. The framework of authentic intellectual work has also been used as an integral component of a large-scale study on school reform and students with disabilities. As part of their five year national study, the Research Institute on Secondary Education Reform for Youth with Disabilities (RISER) used the framework of authentic intellectual work to examine research questions related to authentic teaching and learning for students with disabilities, the impact of authentic practices on school outcomes, and the support structures necessary to promote authentic learning and performance for all pupils (Braden, Schroeder, & Buckley, 2001; Collet-Klingenberg, Hanley-Maxwell, & Stuart, 2000; Hanley-Maxwell, Phelps, Braden, & Warren, 1999, 2000; King et al., 2001; Mooney & Phelps, 2001; Mooney, Phelps, & Anctil, 2002). In particular, RISER examined the authenticity of assessments in secondary inclusion classrooms and the performance on those assessments of pupils with

and without disabilities (King et al., 2001). King et al. scored 51 assessment tasks and 314 samples of pupil work from inclusion classrooms. In one part of the study, teachers submitted whole class sets of work, while the second data set consisted of samples of work from a matched set of pupils (with and without disability) on the same assessment.

Researchers concluded that pupils, both with and without disabilities, who were exposed to higher quality authentic assessments produced higher quality authentic intellectual work than their peers who were given lower quality assignments. In addition, pupils with disabilities who were given higher authentic tasks outperformed pupils, both with and without disabilities, who were given less challenging tasks. This work, like the effect studies on low-achieving pupils, suggests that pupils with disabilities can also perform at high levels when given authentic tasks. The presence of authentic intellectual learning opportunities is beneficial for all learners.

*Authentic intellectual work and professional development/school reform*. The framework of authentic intellectual work originated within the larger context of school restructuring efforts. Therefore, research in this field has been conducted to determine the effect of particular school reform initiatives or professional development programs on the quality of authentic intellectual work in the classroom and pupil performance. This review details four such studies.

Two of the studies were primarily quantitative and large-scale. As part of the Chicago Annenberg Challenge research, Bryk, Nagoaka, and Newmann (2000) analyzed the authentic intellectual quality of assessment tasks at the beginning of the study in 1997 and compared the scores to assessment tasks collected at the end of the study in 1999 to

assess whether or not programmatic changes that were a part of the reform initiative contributed to improved assessment quality. Based on previous research, this analysis assumes that assessment tasks that are higher in authentic intellectual quality will result in higher authentic intellectual pupil achievement; no samples of pupil work were analyzed. Although the overall intellectual quality of all tasks remained low, there was an improvement in the quality of assessment tasks that teachers labeled as "challenging." There was more variation in the quality of the assessment tasks labeled as "typical."

Researchers involved in the SIPA study also examined the impact of reform initiatives on the authentic intellectual work scores, in this case Quality Teaching scores, of classroom instruction, assessment tasks, and pupil work. Gore, Williams, and Ladwig (2006) drew on data from the larger study to explore the effect of a professional development program that specifically addressed, and provided support on, Quality Teaching on the work of 206 early career teachers (0-3 years experience). Quality Teaching scores for these novice teachers were compared to the Quality Teaching scores of teachers in these programs who had more than four years of teaching experience ($n$=775). There were no statistical differences in Quality Teaching scores between first year teachers, teachers with 1-3 years experience, and teachers with more than 4 years experience. This suggests that early career teachers were performing at the same level as experienced teachers, although Gore et al. also contend that it suggests that novice teachers were not receiving any more support than experienced teachers through the initiative.

One study utilized a pre- and post-test design to examine the effect of a professional development program. In this small study related to the larger QRSLS study, Gore, Griffiths, and Ladwig (2002) examined the impact of a series of professional development sessions on Productive Pedagogies (18 hours in total). Twenty-six teachers at two different schools were observed on two to three occasions prior to and following the professional development program. Classroom instruction was scored according to the Productive Pedagogies scoring rubric. Gore et al. found significant statistical differences between teachers' Productive Pedagogies scores pre- and post-intervention. They suggest that the workshop had a positive effect on teacher quality.

Finally, one small qualitative study explored the experiences of teachers involved in an authentic intellectual work professional development program. Avery, Kouneski, and Odendahl (2001) provide a program description of a monthly "Authentic Pedagogy in the Social Studies" seminar for 16 teachers at two Minnesota high and middle schools. In this seminar, teachers spent one day of class each month to discuss Newmann's authentic intellectual work framework. Teachers developed authentic assessments collaboratively and provided each other with peer feedback. They also videotaped and critiqued their own instruction and scored pupil work based on the standards for authentic pedagogy and pupil achievement. Through their analysis of interview data, Avery et al. conclude that the program provided teachers with a common language – that of authentic intellectual work – to talk about teaching and learning which contributed to a renewed sense of collegiality and energy. In interviews, teachers reported that their teaching

changed "drastically" (p. 100) as they began to implement aspects of authentic intellectual work into their instruction.

These four studies suggest that when teachers are exposed to the ideas of authentic intellectual work, they are more likely to engage in authentic intellectual work in their classroom. Again, the assumption here is that as teachers engage in more authentic practices, their pupils will perform at higher authentic intellectual levels. Since these studies looked at inservice teachers who may not have been exposed to concepts related to authentic intellectual work, it raises the question of how teachers who are prepared to consider these concepts might think about authentic intellectual work as they create assessments and think about pupil learning. This dissertation looks at preservice teachers in a preparation program that emphasized pupil learning.

**Longitudinal studies**. The cross-sectional effect studies suggest that the use of authentic intellectual work has an impact on pupil learning, but they do not get at how the gains measure over time. Researchers have attempted to address that by linking authentic intellectual work to longitudinal data on pupil performance.

As part of the CORS study, researchers used data from the National Educational Longitudinal Study (NELS) to measure pupil learning gains (Lee & Smith, 1994; Lee, Smith, & Croninger, 1995, 1997). This longitudinal study used NELS survey data to determine which schools nationwide had features of restructured schools, one of the traits being authentic intellectual work, based on how teachers and principals self-reported information on the survey about their teaching practices and school structure. This survey looked at 11,000 students in 820 schools and they labeled schools as traditional,

moderate, and restructuring. They looked at math, social studies, science, and reading scores for 8th graders in 1988 through 10th graders in 1990 on the test questions on the NELS survey, which were taken from the National Assessment of Educational Progress, to measure pupil learning gains (Lee & Smith, 1994) and used a Hierarchical Linear Model to control for differences in SES, race, and prior achievement. There were increased gains in pupil performance in schools that had more restructuring characteristics. They followed those same students in their last two years of high school and their achievement in math and science (Lee et al., 1995) and found that the learning gains remained in the later years of high school. In schools that had consistently higher levels of authentic intellectual work, pupils learned more and there was a smaller achievement gap among high and low SES pupils.

This work suggests that exposure to authentic intellectual work influences gains in pupil achievement from one year to the next and that these gains can be sustained throughout high school if high levels of authentic intellectual work are present. In addition, this work suggests that authentic intellectual work may have an impact on reducing the achievement gap between pupils of different race, SES and prior achievement.

**Authentic Intellectual Work and Teacher Education**

The majority of research on authentic intellectual work has focused on inservice teachers, many of whom were involved in school restructuring studies. For the most part, these teachers were experienced teachers, although oftentimes they were involved in professional development programs where they learned new teaching strategies and

approaches. That work is valuable in gaining insight into the relationship between teaching, learning, and authentic intellectual work. However, they do raise the question as to whether or not these relationships are the same for teacher candidates and teachers in their early career. Since this dissertation focuses on teacher education and novice teachers, this section of the review examines literature where the framework of authentic intellectual work has been used in teacher education. This work includes instances where teacher education has incorporated the framework into their program and/or instances where teacher candidates have been evaluated according to these concepts. The research is limited and suggests that teacher education has not explicitly addressed this framework in relation to assessing teacher candidates. There are only two studies in this review which utilize authentic intellectual work specifically in the teacher education context. Both studies are small and qualitative and were conducted by the same group of researchers, although they involve different groups of teacher candidates. A third study involving authentic intellectual work and university professors is also included here because it has implications for teacher education.

The two studies on authentic intellectual work and teacher education come out of the research in Queensland and New South Wales. In their work, Gore, Griffiths, and Ladwig (2002) conclude that in order for authentic intellectual work, in this case productive pedagogies, to become embedded in teacher practice, the concepts need to be integrated within the teacher education program from the beginning and the central focus for how teachers think about teaching and learning. Gore (2001) calls for teacher education to incorporate dimensions of productive pedagogies throughout preparation

programs and that teacher educators need to attend to and model those practices as well for their teacher candidates so that teacher candidates also engage in high quality intellectual work. Teacher education, she argues, has not done enough to foster authentic intellectual work practices and needs to place classroom pedagogy and practice at its center to create teachers who teach for pupil learning.

The two research studies presented here explored the impact of coursework in authentic intellectual work and how teacher candidates/teachers utilized these concepts in their teaching. Gore, Griffiths, and Ladwig (2004) undertook a study to examine if productive pedagogies "provide a framework for bringing greater coherence and a firmer, more confident knowledge base to the work of teachers and teacher educators" to see if it can "help produce better teachers" (p. 376). In this study, 30 teacher candidates took a course in productive pedagogies during the last year of their teacher education program. Throughout the course, teacher candidates learned about the dimensions of the framework, the individual items that comprised the dimensions, and the coding practices that were developed to evaluate productive pedagogies. Researchers then observed 10 of these teacher candidates during student teaching and coded the observed lessons in accordance with the productive pedagogies scoring manual. For these observed lessons, teacher candidates were told to simply create a lesson as they normally would and not to specifically add elements of productive pedagogies. Two weeks after the observation, participants were interviewed about applying productive pedagogies to their teaching.

Gore, Griffiths, and Ladwig (2004) found that teacher candidates' productive pedagogies scores were similar to the productive pedagogy scores of experienced

teachers involved in the QSRLS, which suggests that teacher candidates who took a course in the framework were as comfortable with productive pedagogies as experienced teachers who were involved in professional development reform initiatives related to productive pedagogies. However, the relatively low level of productive pedagogies that these teacher candidates – and experienced teachers – engaged in suggest that the teacher education program did not foster productive pedagogies enough. Teacher candidates were successful, however, in one dimension of productive pedagogies. They scored higher on the dimension "supportive classroom environment" than they scored on "intellectual quality," "relevance," or "recognition of difference." Teacher candidates appeared comfortable with that dimension whereas they tended to simplify "intellectual quality" as an student-centered practice instead of viewing intellectual quality as a cognitive feature where students critically engage in knowledge construction. Based on the interview data, they also found that teacher candidates felt that productive pedagogies were not integrated in their teaching and that the course came too late in the program to have much of an impact. As a result of this study, these researchers suggest that teacher educators should rethink their overemphasis on classroom environment and refocus attention to teaching experiences that are grounded in the quality of learning, identification of central concepts, and depth of understanding. In addition, they also recommend that productive pedagogies be integrated throughout the teacher education program.

The second study in this line of research suggests that even when teacher education supports and nurtures authentic intellectual practices, novice teachers find little support for this type of work in the early years of teaching. Gore, Williams, and Ladwig

(2006) later conducted a case study of seven new teachers who had taken a course on Quality Teaching (the New South Wales version of productive pedagogies) during their university preparation program. Through lesson observations and interviews with the teacher education graduates, this study examined how these teachers used Quality Teaching during their first year of teaching. Researchers concluded that these novice teachers demonstrated an understanding of Quality Teaching following their teacher education, and in fact had Quality Teaching scores that were similar to experienced teachers' scores in other studies, but that these novice teachers found very little support for Quality Teaching where they were teaching. Gore, Williams, and Ladwig suggest that the teacher education course may have helped these teachers sustain good pedagogical practices in situations where those practices were not mirrored by other teachers.

These two studies indicate two important implications for teacher education. First, teacher candidates involved in courses in authentic intellectual work have demonstrated that they are able to engage in authentic intellectual work practices on par with experienced teachers. In further support of this, the SIPA study found no correlation between the quality of teaching and the years of teaching experience (Gore et al., 2007). Preparation and exposure to authentic intellectual work may have more of an impact on quality teaching than experience, which makes the inclusion of authentic intellectual work practices in teacher education critical to authentic pupil achievement. This is relevant because of the importance to teacher education. Second, these two studies, along with the research on inservice teachers, suggest that teachers are not being prepared to create assessment tasks and pedagogical practices of high intellectual quality. Teacher

education has an opportunity to address this gap, Gore, Ladwig, Griffiths, and Amosa (2007) argue, by developing teacher candidates who have a strong conceptual understanding of pedagogy, an extensive knowledge base, instruction and assessment design, and engage in powerful learning experiences that foster commitment and responsibility to pupil learning.

The final study reviewed here is related to university teaching in general, but is relevant to the presence of authentic intellectual work in teacher education programs. Gore, Ladwig, Amosa, Griffiths, Parkes, and Ellis (2008) used the Quality Teaching framework to study university teaching. In this preliminary report of an ongoing study, researchers examined assessment tasks ($n$=22) and student performance ($n$=233) for courses in five different disciplines (Education, Law, Humanities and Social Sciences, Medicine and Public Health, and Nursing and Midwifery) at one university. Here, university assessment tasks scored higher on the Quality Teaching scale than assessments in secondary schools in studies using the same rubric. In addition, unlike secondary schools, university pupils were still able to produce high levels of authentic achievement when they were given low scoring assessment tasks. Researchers found no correlation between the quality of the task and achievement. The implication for teacher education is that teacher candidates are able to engage in high quality intellectual work and may be familiar with this type of work through their own higher educational experiences.

The limited research involving teacher education suggests that there is a need to examine how preservice teachers engage in authentic intellectual work with their pupils. These studies suggest that preservice teachers are capable of the same levels of authentic

intellectual work as experienced teachers, particularly when the concepts are taught in the teacher preparation program and there is support for this type of work in the school where they teach. This dissertation provides additional insight into the relationship between teacher education programs and engaging pupils in authentic intellectual work.

**Literature Review Discussion and Conclusion**

This literature review described the theory of democratic education and explained how this theoretical framework provides a foundation for the type of knowledge needed to participate in contemporary American democracy. The emphasis on rational deliberation and critical thinking is central to the framework of authentic intellectual work, which can be used to measure teacher candidates', teachers', and pupils' learning.

The empirical research on authentic intellectual work suggests that there are a number of challenges in using this framework to evaluate teacher and pupil authentic intellectual work. First, this type of research is time- and labor-intensive and difficult to maintain high participant compliance rates. The majority of the studies that utilize this framework were large-scale and involved a team of researchers. Research was time consuming for the participants as well which contributed to low compliance rates. Researchers cited instances where teachers and schools did not always submit the required assessment task and pupil work samples (Ladwig et al., 2007; Newmann et al., 1996). Teachers complained that it took a significant portion of their time to collect and photocopy pupil work and provide relevant information when they were asked to do things such as fill out a required cover sheet explaining the context of the assessments

and pupil work (Matsumura, 2003). This illustrates the costs of evaluating learning in a way that accounts for complexity and is connected to the classroom.

In addition, many studies experienced difficulties tracking pupil learning outcomes for specific pupils. Researchers were not always able to collect multiple samples of pupil work from the same pupils. Factors that contributed to this problem included pupil absences, changes in course schedule, and instances where secondary teachers provided pupil work samples from different classes that they taught. Newmann, Marks, and Gamoran (1996) were only able to collect multiple samples of pupil work from the same pupils over time for 45% of their sample.

Finally, this research suggests it is challenging to evaluate the authentic intellectual work of teachers. In some cases, teachers did not always submit assessment tasks that they created themselves (Gore et al., 2007). Since oftentimes the research design of these studies evaluate the quality of instruction based on the quality of assessment, it is not an accurate reflection of the teacher's ability when she submits assessment tasks that were created by someone else. Teacher candidates/teachers in this dissertation did submit assessments they did not create. Although I discuss this in further detail in Chapter 5, my analysis used qualitative data to better understand teachers' thoughts about assessment and pupil learning.

This dissertation addresses gaps in the literature by using the framework of authentic intellectual work to examine teacher candidates' work during the preservice period and first two years of teaching. First, previous studies tend to focus on inservice teachers. Therefore, this study contributes to the literature by examining the relationship

between learning to teach and teacher preparation to authentic intellectual work. Second, whereas previous longitudinal studies such as the CORS and the Chicago School Research work did not always follow the same teachers and pupils, this dissertation follows the same 11 participants over the course of three years. This provides insight into whether or not teachers' engagement of authentic intellectual work changed over time. In addition, these longitudinal studies are primarily quantitative. This dissertation draws upon significant qualitative data to better understand how teachers think about creating assessments and evaluating pupil learning in ways that are consistent with authentic intellectual work.

# CHAPTER THREE: RESEARCH METHODS

The purpose of this study is to examine the extent to which teacher candidates/teachers, during their preservice period and first two years of teaching, think about authentic intellectual work, as a means of promoting the type of learning consistent with the theoretical frame of democratic education, when they create assessments and when they talk about pupil learning, and the extent to which their pupils engage in authentic intellectual work. This dissertation, which draws from a larger longitudinal qualitative case studies project, utilized a qualitatively-driven, concurrent embedded mixed methods approach (Creswell et al., 2003; Morse & Niehaus, 2009) and cross-case study methodology (Stake, 1994) to understand how 11 teacher candidates/teachers did this.

In particular, as a means to examine teacher candidates' use of authentic intellectual work, this study used the Teacher Assessment/Pupil Learning (TAPL) protocol, a research instrument where participants took part in an interview about assessment and pupil learning in relation to specific samples of assessment tasks/assignments and pupils' work that they used in their classroom. This protocol also features an outside examination of these samples of assessments and pupil work. The TAPL protocol, described in detail in a later section, was used with participants in this study at five different points in time during their preservice period and first two years of teaching. This study used an interpretive qualitative approach to explore TAPL interview data to discern how participants thought about concepts related to authentic intellectual work when they talked about assessment and pupil learning and a quantitative rubric to

evaluate the quality of authentic intellectual work in the teacher candidates'/teachers' assessment tasks and pupil work samples. Qualitative and quantitative analyses were integrated to construct a nuanced description and understanding of the extent to which teacher candidates/teachers and their pupils engaged in rigorous authentic intellectual work and how context and conditions influenced this work as beginning teachers developed in the early phase of their career.

This section presents an overview of mixed methods research and describes particular elements of mixed methods research that are related to the dissertation, including qualitatively-driven, concurrent, and embedded designs, followed by an overview of case study methodology. Next, drawing on the framework of mixed methods research, I describe the research design for this study detailing the larger qualitative case studies project from which this study derives, study participants, and data collection and analysis strategies. I conclude with a discussion of the integrity of the study, considering the issues of rigor, reflexivity, validity, and limitations.

## Overview of Mixed Methods Research

To explore how teacher candidates/teachers reflect the ideas of authentic intellectual work in their creation of assessments and understandings of pupil learning, this study utilized a mixed methods approach. This study drew upon data collected during a larger, longitudinal qualitative case studies project and used interpretive qualitative techniques to analyze qualitative interview data from that study. At the same time, this study used a quantitative approach to analyze artifacts collected from the larger study and

employed statistical analyses to examine relationships between variables and variation over time.

Mixed method research has been called the "third methodological movement" (Tashakkori & Teddlie, 2003, p. ix) in social science research and is a methodology that draws upon and combines the two general paradigms that preceded the movement: quantitative and qualitative methods. Quantitative and qualitative researchers operate with differing and competing assumptions concerning research ontology, epistemology, and methodology (Guba & Lincoln, 1994), and mixed methods research has been seen, by some, as a way to bridge the two paradigms in a way that takes advantage of the strengths of both traditions to answer new types of research questions and create new understandings.

**Qualitative and Quantitative Methods**

It has been argued that identifying research by the terms "quantitative" and "qualitative" is a simplistic, and not a very useful, distinction (Howe, 1988). However, in a very general sense, these terms have common and particular meanings. Quantitative research comes from a historical tradition linking science with quantification (Guba & Lincoln, 1994), where social science researchers worked predominantly within the positivist traditions that drew upon numerical analyses and "objective" measurements (Teddlie & Tashakkori, 2003). Scientific objectivity and the construction of replicable research designs are set as ideals for many quantitative researchers (Guba & Lincoln, 1994). Quantitative research is typically seen as using a deductive approach to test the relationship between variables and confirm previous theory (Creswell, 2009; Tashakkori

& Teddlie, 2003). Methods used in quantitative research often include experimental design, statistical analyses and interpretation, and instruments, such as surveys, featuring closed-response items or likert scales (Creswell, 2009).

Qualitative research, on the other hand, operates under the assumption that research is not objective and is dependent upon more than just numerical analyses. With this tradition, qualitative researchers see research as value-laden and "stress the socially constructed nature of reality, the intimate relationship between the researcher and what is studied, and the situational constraints that shape inquiry" (Denzin & Lincoln, 1994, p. 4). Also referred to as interpretivist research, qualitative researchers operate under the assumption that it is important to understand the meanings behind people's actions and beliefs and "that to understand this world of meaning one must interpret it" (Schwandt, 1994, p. 118). In qualitative research the researcher is seen as the instrument, constantly negotiating a relationship between researcher and participant, and research takes place in a natural setting (Denzin, 1994). There are many different forms of qualitative research methods ranging from grounded theory (Charmaz, 2004; Glaser, 1992; Glaser & Strauss, 1967), to ethnography (Eisenhart, 2001; Wolcott, 1999), to phenomenology (Benner, 1994). Qualitative data includes *experiencing*, *enquiring*, and *examining* (Wolcott, 1992) which often take the form of observations, interviews, and material culture. Qualitative research is sometimes seen as exploratory (Teddlie & Tashakkori, 2003) where themes emerge from the data as researchers organize data through an iterative process to construct inferences and generalizations (Creswell, 2009).

**Mixed Methods**

Although quantitative and qualitative approaches are sometimes considered to be dichotomous, mixed methods research attempts to combine the two. Due to the inherent tensions between quantitative and qualitative perspectives, researchers disagree over the degree to which mixed methods can, or even should, be used as a sound research methodology. Teddlie and Tashakkori (2003) identify six stances towards mixed methods research: 1) methods and paradigms are separate from one another, which makes mixed methods possible because they are unrelated; 2) qualitative and quantitative paradigms are incompatible which makes mixed methods impossible; 3) methods must be kept separate and remain true to their paradigms, but they can work together to answer a research question; 4) one single method can serve as the foundation for mixed methods; 5) different paradigms should engage with one another to examine the tensions between the two; and 6) multiple paradigms can be used, depending on the type of study. Although some researchers continue to believe in the incompatibility stance, that view "has now been largely discredited, partially because scholars demonstrated that they had successfully employed mixed methods in their research" (Teddlie & Tashakkori, 2003, p. 19). At the other extreme, some researchers advocate for the dissolution of the identification of research in quantitative or qualitative terms since "all research is interpretive" (Schwandt, 2000, p. 210).

Researchers tend to agree that the most important factor related to choosing a research methodology is the research question (Guba & Lincoln, 1994; Teddlie & Tashakkori, 2003) and that therefore, research methods should be chosen based on the

most appropriate way to answer a research question. From this perspective, mixed

methods research is often associated with the philosophy of pragmatism as a practical and

applied stance to tackle research questions that either quantitative or qualitative methods

alone are unable to address (Howe, 1988; Johnson, Onwuegbuzie, & Turner, 2007;

Tashakkori & Teddlie, 1998). Pragmatists suggest that qualitative and quantitative

methods are compatible (Howe, 1988). Howe (1988) argues there is a "two-way

relationship between methods and paradigms" which makes it possible for methods to

inform paradigms and thereby makes the two methods compatible (p. 10).

Although there have been different terms used to describe this combining of

methods, such as blended research, multimethod, triangulated studies, multiple method,

and mixed research (Creswell, Plano Clark, Gutmann, & Hanson, 2002; Johnson et al.,

2007; Schwandt, 2000; Teddlie & Tashakkori, 2003), the term "mixed methods" has

become the most commonly used. However, a common definition of mixed methods has

been more difficult to come by. In its most basic form, mixed methods research

"intentionally combines different methods" (Greene & Caracelli, 1997, p. 7) and includes

combining the "qualitative and quantitative viewpoints, data collection, analysis, [and]

inference techniques" (Johnson et al., 2007, p. 123). Mixed methods research "also

involves the use of both approaches in tandem so that the overall strength of a study is

greater than either qualitative or quantitative research" (Creswell, 2009, p. 4). Building

on an analysis of the various definitions that researchers have used, Johnson,

Onwuegbuzie, and Turner (2007) define mixed methods as such:

> Mixed methods research is an intellectual and practical synthesis based on
> qualitative and quantitative research…Mixed methods research is the research

paradigm that (a) partners with the philosophy of pragmatism…(b) follows the logic of mixed methods research (including the logic of the fundamental principle and any other useful logics imported from qualitative or quantitative research that are helpful for producing defensible and usable research findings); (c) relies on qualitative and quantitative viewpoints, data collection, analysis, and inference techniques combined according to the logic of mixed methods research to address one's research question(s); and (d) is cognizant, appreciative, and inclusive of local and broader sociopolitical realities, resources, and needs. Furthermore, the mixed methods research paradigm offers an important approach for *generating* important research questions *and* providing warranted answers to those questions. (p. 129)

Therefore, mixing methods provides an opportunity to address research questions in a more complete and powerful way where researchers can both answer questions and generate theory. Greene, Caracelli, and Graham (1989) also point to other purposes of mixed methods research including *triangulation* to see how different methods answer the same question; *complementarity* to add to one method with another; *development* to use one method to inform another; *initiation* to explore questions; and *expansion* to add to the field by looking at the topic in a new way. In further defining the field, mixed methods researchers have created nomenclature to refer to the different methods, adopting "qual" for qualitative methods and "quan" for quantitative methods (Tashakkori & Teddlie, 2003).

There are different types of mixed methods research designs, depending primarily upon how the quantitative and qualitative portions work in tandem, particularly in terms of data collection and data analysis. Data design in mixed methods does not have to follow one particular paradigm (Onwuegbuzie & Teddlie, 2003) and the mixing of methods can occur in a single study (Creswell et al., 2002) or across related studies (Johnson et al., 2007). In sequential mixed methods, one method is used to collect and

analyze data, the results of which then inform the construction of a new study with the

second method (Morse, 1991; Tashakkori & Teddlie, 1998). An example of sequential

design is a study where a quantitative survey is given and analyzed and then followed by

interviews with a few participants about their responses to further clarify issues related to

the survey. In mixed methods nomenclature this is represented by an arrow (i.e., quan $\rightarrow$

qual). Another example of a sequential design is a study that conducts qualitative

interviews with a focus group and uses that work to inform the construction of a survey;

this type of design would be designated qual $\rightarrow$ quan. It should be noted, however, that

Morse and Niehaus (2009) portray qual $\rightarrow$ quan designs as more advanced than simply

using an interview to inform a survey since that practice is an accepted and common

component to instrument design; instead, they argue, a qual $\rightarrow$ quan design should

involve either collecting new data for the quantitative piece or transforming qualitative

data into a form that can then be quantified. In concurrent mixed methods studies

(Creswell et al., 2003), also referred to as parallel or simultaneous mixed methods

(Tashakkori & Teddlie, 1998), qualitative and quantitative data are collected and/or

analyzed at the same time. For example, a survey might include close- and open-ended

responses which are then analyzed with quantitative and qualitative methods respectively.

This is represented by a "+" (i.e., quan + qual).

As a developing field, mixed methods research designs have taken a variety of

approaches to integrating the qualitative and quantitative processes. Creswell, Plano

Clark, Gutmann, and Hanson (2003) argue that researchers have integrated qualitative

and quantitative approaches at various points in time including "[d]uring the phases of

problem/question specification, data collection, data analysis, and interpretation" (p.

220). Typically in mixed methods research, integration occurs at the analysis and

interpretation phases. In fact, Morse and Niehaus (2009) argue that mixing methods can

only occur during either of these two times. Qualitative and quantitative data, they claim,

should be kept separate and dealt with in an appropriate and rigorous manner according

to their respective paradigms until the *point of interface* which can be at the analysis or

the results discussion phase. That is to say, qualitative and quantitative data can be

transformed prior to the analytic stage so that both data are used together in the analysis

phase or that the two methods can be analyzed separately and then integrated in the

discussion section where the two different analyses are compared to create a combined

interpretation of the research phenomenon. In this study, integration of qualitative and

quantitative data occurred in the question phase (individual questions addressed particular

methods), data analysis (qualitative data were translated into quantitative data), and

interpretation (results of the various analyses were compared to understand the larger

research problem).

Data analysis in mixed methods research can take a variety of forms but involves

general procedures. According to Onwuegbuzie and Teddlie (2003), there are seven

stages of mixed methods data analysis: (1) data reduction, (2) data display, (3) data

transformation, (4) data correlation, (5) data consolidation, (6) data comparison, and (7)

data integration. Not all of the stages need to be present and they may occur in a different

order. In stage 1, data are reduced through such processes as the computation of

descriptive statistics for quantitative data or creating themes from qualitative data. In

stage 2, data are displayed visually in tables, graphs, matrices, or charts. Stage 3 involves transforming qualitative and/or quantitative data into forms that can be used with the other method. The next three stages involve correlating, consolidating, and/or comparing the quantitative and qualitative data, depending upon the research design. Finally, in stage 7, quantitative and qualitative data that have either already been combined or have been kept separate through the analysis process are brought together to interpret as a whole. The data analysis in a later section describes how data in this dissertation follows these stages in the analysis of data.

**Embedded Mixed Methods**

In both sequential and concurrent designs, one method is often more dominant than the other. Although some researchers see mixed methods as a way to incorporate inductive and deductive theoretical perspectives and assert that the ideal is equality among approaches (Johnson et al., 2007), others argue that it is important to know whether the study has an inductive or deductive theoretical drive (Caracelli & Greene, 1997; Morse & Niehaus, 2009). Morse (1991) argues that "a project must be either theoretically driven by the qualitative methods incorporating a complementary quantitative component, or theoretically driven by the quantitative method, incorporating a complementary qualitative component" (p. 121). Caracelli and Greene (1997) describe designs where one methodology is located within a larger methodology as "embedded or nested designs" that "feature one methodology located within another, interlocking contrasting inquiry characteristics in a framework of creative tension" (p. 24). From this perspective, the predominant method guides the study in the construction of research

questions, data collection, and data analysis. In mixed methods nomenclature, the dominant method is capitalized while the supplemental method is written in lower case (i.e. QUAN → qual) (Morse, 1991). At this point in time, however, there are very few descriptions in the literature about how to conduct an embedded design (Creswell, 2009).

Johnson, Onwuegbuzie, and Turner (2007) define qualitative dominant mixed methods research as "the type of mixed research in which one relies on a qualitative, constructivist-poststructuralist-critical view of the research process, while concurrently recognizing that the addition of quantitative data and approaches are likely to benefit most research projects" (p. 124). The study here is an example of a concurrent embedded design with an inductive, qualitative theoretical drive (QUAL + quan). That is, the larger study from which this dissertation is derived is a qualitative case studies project where observations, interviews, and a collection of artifacts, assessment assignments, and pupil work samples were collected simultaneously at different time-points. As part of the larger qualitative case study, the Teacher Assessment/Pupil Learning (TAPL) protocol was developed to analyze pupil learning using both a qualitative and quantitative component. This dissertation is an analysis of the TAPL data. Both the quantitative and qualitative TAPL data were collected concurrently. The overall drive of the case studies project is inductive qualitative. The study here retains that inductive theoretical drive to see what is happening. The addition of the quantitative component is used to "enhance QUAL studies with measurement" in a way that can enable comparison, enhance description, illustrate, and triangulate (Morse & Niehaus, 2009, p. 99) and can also be used "to enrich the description of the sample participants" (Creswell et al., 2003, p. 230).

Figure 3.1 presents a graphic representation of this study's research design to illustrate how this is a mixed methods study. The outside rectangle represents the larger longitudinal Qualitative Case Studies Project (QCS). The data sources for this larger study included classroom observations, interviews, artifacts (teacher candidates' teacher preparation coursework), and samples of assessments and pupil work used in their classrooms. Part of the QCS research design included the TAPL protocol which drew from qualitative interviews (Internal TAPL) and a quantitative analysis of assessments and pupil work (External TAPL). This dissertation, which is an analysis of the TAPL protocol, is represented by the center box. A more detailed description follows in the research design section.
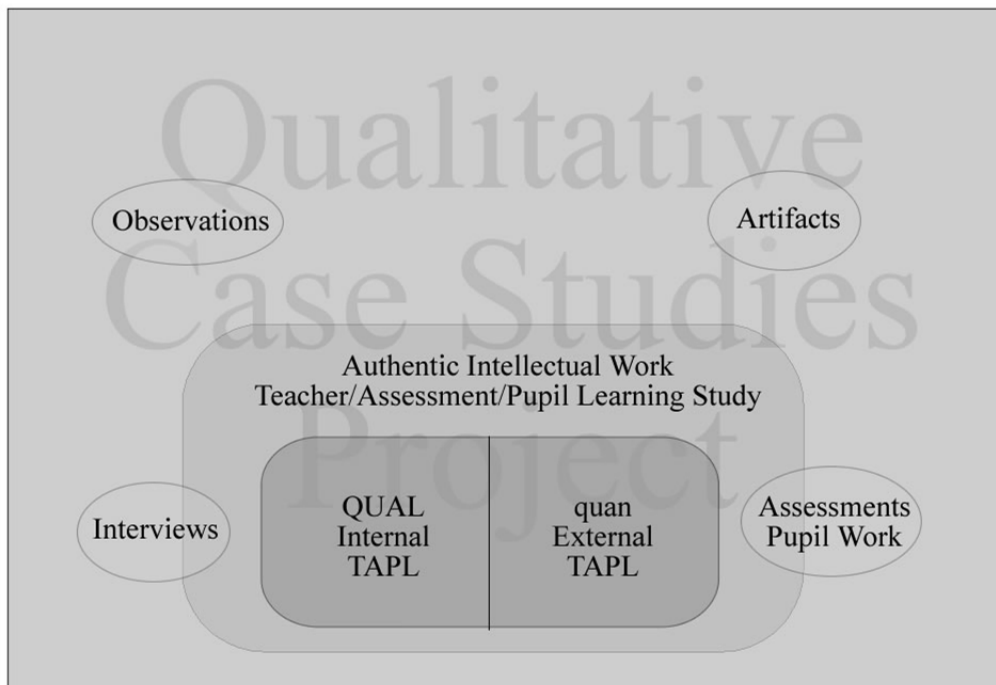


*Figure 3.1*. Research design.

**Quantitizing Qualitative Data**

In a qualitatively-driven mixed methods research design, one way to integrate qualitative and quantitative data during the data reduction stage of data analysis is to quantify qualitative data, meaning to represent qualitative data in numerical form. Incorporating numerical analyses in qualitative data is not new and normally occurs, to some extent, in qualitative methodology and in qualitative methods such as ethnography, because, as Dobbert and Kurth-Schai (1992, p. 137) point out, "the ability of mathematical techniques to clarify patterns has proven valuable" (p. 137). Even before the recent paradigmatic shift legitimizing mixed methods, Mitchell (1979) described how quantitative data could be used in qualitative work and how anthropologists have used quantitative "analytical procedures" as "aids to description" of their fieldwork:

> [T]he more detailed knowledge which quantitative methods allow and the correlation between phenomena which statistical reasoning can educe should be the essential foundation on which anthropologists start to erect their generalizations about the social behavior of the people they study. Quantitative methods are essentially aids to description. They help to bring out in detail the regularities in the data the fieldworker has collected. Means, ratios, and percentages are ways of summarizing the features and relationship in data. Statistical measures based on the theory of probability go beyond the mere quantitative data and use devices to bring out the association between the various social facts the observer has collected. These are essentially analytical procedures and, as Fortes puts it, "are nothing more than a refinement of the crude methods of comparison and induction commonly used." (p. 20)

From this perspective, quantitative data was used as another way to describe qualitative data. Providing an example of what this might look like, Miles and Huberman (1994) describe the process of "quantizing" qualitative data "where qualitative information can be either counted directly (say, the number of times a doctor interrupts a patient during an interview) or converted into ranks or scales (in this teacher's classroom 'moderate'

mastery of a specific innovation has occurred)" (p. 42). Tashakkori and Teddlie (1998),

use the term "quantitizing techniques," a term that is more commonly used and is used

throughout the remainder of this study, to characterize this as "converting qualitative

information into numerical codes that can be statistically analyzed" (pp. 125-126). More

specifically,

> Quantitizing might include a simple frequency count of certain themes, responses, behaviors, or events. On the other hand, it may consist of a more complex rating of the strength or intensity of these events, behaviors, or expressions. Depending on the type of transformation, different QUAN techniques might be used for their analysis. For example, descriptive statistics might be used to summarize/organize the frequency counts, or more complex procedures such as factor, correlation, or regression analysis might be performed on the ratings. (Tashakkori & Teddlie, 1998, pp. 128-129)

When qualitative data are quantitized, an additional level of statistical methods can be

used to better understand or triangulate qualitative information and confirm qualitative

findings (Hesse-Biber, 2010; Sandelowski, 2001), providing one way that qualitative and

quantitative data can be integrated (Creswell, 2009).

Researchers have pointed to a number of ways that quantitized qualitative data

can contribute to the research process particularly to identify patterns, maintain analytical

integrity, and verify hypotheses (Miles & Huberman, 1994; Sandelowski, 2001).

Quantitizing data opens up a new range of statistical techniques available to qualitative

researchers. When narrative data are reduced to a variable, for example, they can be

correlated to other variables with statistical techniques (Hesse-Biber, 2010; Sandelowski,

2001), and matrices of quantified data can aid in constructing effect sizes and correlations

(Onwuegbuzie & Teddlie, 2003). Sandelowski (2001) argues, "Counting is integral to the

analysis process, especially to the recognition of patterns in data and deviations from

those patterns, and to making analytic or idiographic generalizations from data" (p. 231). Displaying information numerically can illuminate patterns, generate new questions, and clarify meaning.

At the same time, there are drawbacks to using quantitized data. By converting qualitative data into numerical codes with a singular meaning, researchers lose the multiple meanings that qualitative data often represent (Sandelowski, 2001). This may result in "an oversimplification of emergent themes that does not capture the complexity of the meaning conveyed by the unit" (Onwuegbuzie & Teddlie, 2003, p. 370). This reflects the inherent tension between quantitative and qualitative views. According to Sandelowski (2001), "Qualitative researchers tend to prefer figures of speech over figures, and tableaux of experience over tables of numbers. Numbers illuminate but also obscure and thereby complicate the necessary tension between science and art in qualitative research" (pp. 235-236). Counting, like any research technique, is not always appropriate (Onwuegbuzie & Teddlie, 2003), can be misleading, and lead to misleading findings if not used in the proper context (Sandelowski, 2001). It is worth noting here that quantitative data can also be transformed into qualitative data in a process referred to as qualitizing, and although a few researchers argue that both quantitizing and qualitizing must occur in the same study to truly integrate qualitative and quantitative data (Sandelowski, Voils, & Knafl, 2009), the two processes do not always have to be used together (Onwuegbuzie & Teddlie, 2003), as is the case in this dissertation.

This study used quantitizing techniques to analyze qualitative data. Participant artifacts (assessments and pupil work samples) received a numerical score based on a

rubric for authentic intellectual work. Descriptive and inferential statistical analyses were used to describe the extent to which participants demonstrated authentic intellectual work and how and why that may have changed over time. These quantitative data provided a measure to compare authentic intellectual work among and between participants and helped describe other qualitative data. A more detailed explanation follows in the section on research design.

**Case Studies**

This dissertation also draws upon case studies methods. Case study methods are appropriate for mixed methods research because they do not apply to any one particular research paradigm (Stake, 1994). Rather, case study has been described as "a transparadigmatic and transdisciplinary heuristic that involves the careful delineation of the phenomena for which evidence is being collected" (VanWynsberge & Khan, 2007, p. 90). That is to say, case study methodology can be used in any discipline or paradigm and is an approach to focus the analysis of a particular phenomenon. A case is what is being studied, and can take the form of an individual, event, process, or activity (Creswell, 2009). There are boundaries to the case and the "boundedness and the behavior patterns of the system are key factors in understanding the case" (Stake, 1994, p. 237). According to Stake, "Case researchers seek out both what is common and what is particular about the case, but the end result regularly presents something unique" (p. 238). When studying a case, researchers decide how to define the boundaries for the case and what themes to focus on, looking for patterns in the data related to those themes, triangulating observations and ideas, looking at alternative interpretations, and making claims or

generalizations concerning the case. Case studies focus in depth on a small sample, provide detailed information on context, occur in natural settings, have particular boundaries, generate hypotheses and new analysis, utilize multiple data sources, and extend the reader's experience (VanWynsberge & Khan, 2007). A case can be one that a researcher finds interesting, that is instrumental to a phenomenon, or a collective group of cases (Stake, 1994).

The collective case study, also referred to as multiple- or cross-case study, is comprised of a number of cases that can be compared and contrasted to see how a particular phenomenon (here authentic intellectual work) is influenced by contextual factors. Looking across cases extends validity and increases generalizability to test ideas in different circumstances. The difficulty is "reconciling the particular and the universal: reconciling an individual case's uniqueness with the need to understand generic processes at work across cases" (Huberman & Miles, 1994, p. 435). To analyze cases, a researcher could repeat the same analysis with all of the cases to see if patterns match, look at themes that cut across cases, examine multiple exemplars, or form "types" to see if there are patterns (Huberman & Miles, 1994). Researchers have to balance how they look within and across cases to refine their inferences based on how the phenomenon is portrayed in different cases, paying particular attention to reintegrate generated themes back into the cases (Ayres, Kavanaugh, & Knafl, 2003). Case study research allows researchers to use the same sequence of data collection to compare the same sequence of data across cases over time.

Case studies are typically suited to qualitative research because they concentrate in-depth on particular cases (Onwuegbuzie & Teddlie, 2003). This study follows 11 teacher candidates/teachers, each of whom constitutes a "case" in a larger longitudinal qualitative case studies project. Within this larger study there are extensive and rich qualitative data structured similarly for each participant, making it possible to engage in cross-case analysis. Results from this study draw on and feed back into the larger QCS study to provide an additional level of analysis of these cases for a more detailed understanding of how teacher candidates use authentic intellectual work in their classrooms.

## Appropriateness of Design

Previous studies of authentic intellectual work have generally worked within a quantitative framework to analyze the authentic intellectual quality of teacher and pupil work (Amosa et al., 2007; King et al., 2001; Newmann et al., 2001). These studies analyzed the effect of particular, pre-determined variables, such as socio-economic status and prior academic performance, on authentic intellectual work, without necessarily providing explanation as to how and why teachers and pupils experienced authentic intellectual work. Teachers' perspectives were frequently neglected in these studies. This dissertation attempts to develop a more contextual and complex understanding of how teacher candidates/teachers engage pupils in authentic intellectual work in their beginning years as a teacher.

With an inductive qualitative theoretical drive, this research study is designed to get at the contexts and conditions that influence teachers' provision of authentic

intellectual work as well as how those change over time. In addition, the rich qualitative

data provide insight into the ways that teacher candidates/teachers conceptualize

authentic intellectual work, which, as evidenced in the literature review, is missing in

other studies. A mixed methods approach that uses an instrument to quantitize qualitative

artifacts but also invites participants to explain their understanding and actions can shed

light onto particular circumstances and meanings that teacher candidates/teachers

experience. At the same time, it provides a uniform way to compare authentic intellectual

work across all 11 cases to see how the factors influencing authentic intellectual work

occur in different contexts.

## Research Design

Drawing upon data gathered during a longitudinal case studies project on the

experiences of teacher candidates learning to teach in a university-based teacher

education program, this study utilized a cross-case, mixed methods approach to explore

two general questions: In a teacher education program that emphasizes pupil learning and

social justice, to what extent do teacher candidates/teachers engage pupils in democratic

education through authentic intellectual work during their preservice period and first two

years of teaching? How do pupils respond to these opportunities? To address these

general questions more specifically, this study addresses the following questions:

1. What is the level of authentic intellectual work that teacher candidates/teachers create for their pupils?
2. What is the level of authentic intellectual work of pupils' responses to these learning opportunities?
3. What is the relationship between the quality of learning opportunities and the quality of pupil learning?
4. How does the way that teachers talk about their goals and understandings of assessment and pupil learning reflect the concepts of authentic intellectual work?

111

5.  How does the way that teachers talk about assessment and pupil learning compare to the quality of authentic intellectual work on their assessment tasks and pupil work?
6.  What conditions and contexts influence the level of authentic intellectual work teachers provide over time?

These research questions were answered by examining the cases of 11 teacher candidates/teachers involved in a longitudinal qualitative case studies project on their experiences of learning to teach during their preservice period and first two years of teaching. At five different time-points throughout the preservice year and the first two years of teaching, participants took part in the Teacher Assessment/Pupil Learning Protocol where they selected samples of assessment tasks/assignments and pupils' work used in their classrooms and discussed the creation of the assessments, contextual factors related to teaching these assignments, and their assessment of the quality of pupils' learning. In this QUAL $\rightarrow$ quan study, interpretive qualitative methods were used to analyze interview data to examine how teacher candidates/teachers think about assessment and pupil learning in relation to authentic intellectual work. At the same time, the assessment tasks and pupil work samples that participants selected were analyzed using a rubric designed to rate the quality of authentic intellectual work. Here the qualitative artifacts were quantitized to enable statistical analyses of the data in order to facilitate cross-case analysis.  The scores from the assessment tasks and pupil work samples were compared to the analysis of the interview data to develop a description of the quality of authentic intellectual work and possible supports and constraints that influence teacher candidates'/teachers' engagement with authentic intellectual work.

**Overview of Qualitative Case Studies (QCS) Project**

  This study draws upon data collected for the larger longitudinal Qualitative Case Studies (QCS) project which studied the experiences of teacher candidates/teachers at one university during their teacher education program and beginning years of teaching. This teacher preparation program has an explicit emphasis on teaching for social justice and pupil learning. A team of researchers followed two successive cohorts of teacher candidates/teachers over a 3 or 4 year period of time (from entry into the Master's level initial teacher preparation program and throughout their first 3 or 4 years of teaching) to examine relationships between teacher candidates'/teachers' experiences, knowledge, beliefs, and practices related to teaching and pupil learning over time. The 22 original participants in the two cohorts of the QCS project volunteered to participate in the study at the start of their teacher education program and were selected by the research team based on these criteria: 1) participants had no prior teaching experience, 2) planned to remain in the university's geographic area during their first year of teaching, and 3) represented roughly the demographics of students enrolled in the university's teacher education program. Data collected during the QCS project involved a series of structured participant interviews; interviews with teacher education faculty, clinical supervisors, and cooperating teachers; structured classroom observations; collection of artifacts related to the teacher education program (i.e. course assignments, syllabi, etc.); and collection of assessments/assignments and pupils' work from participants' k-12 classrooms over the course of a 5-year period.

In terms of research design and analysis, the Qualitative Case Studies project was broadly informed by sociocultural theory as a way to examine social behaviors through the lens of culture and value. Schools and schooling have been described as having a cultural structure consisting of the relationships and norms of the interactions among students, teachers, and administrators (Sarason, 1971). Anthropologist Clifford Geertz (1973) describes culture as public "socially established structures of meaning" (p. 12) related to the values and beliefs that a society creates to govern relationships. From this view, culture is "an historically transmitted pattern of meanings embodied in symbols, a system of inherited conceptions expressed in symbolic forms by means of which men communicate, perpetuate, and develop their knowledge about and attitudes toward life" (p. 89). Culture influences what people do and how they act in their daily lives. Margaret Eisenhart (2001), an educational anthropologist, suggests that people negotiate their surrounding contexts as they engage in the world and cultures, interpreting different meanings from different environments based on their own experiences.

The QCS project drew upon these general theorists in constructing the research design and data analysis, drawing particular connections to Eisenhart's work on how people are continually negotiating multiple cultural contexts. For the QCS project, this means that understanding how people learn to teach involves examining how values and beliefs develop over time and how they shape and are shaped by the teacher education program, the schools in which they teach, and people's prior experiences and characteristics. Teaching and learning, therefore, are negotiated, based on social and cultural beliefs, and value-laden (Gee, 1996). In addition, learning to teach is also

"socially negotiated" (Britzman, 1991, p. 8), in the sense that beginning teachers draw upon past and present experiences, including educational biography and their own experiences and beliefs related to schooling, as they develop their teaching practices and beliefs. The QCS research design and analysis rests upon the assumptions that school is a socially negotiated culture, teaching and learning are not value-neutral, and that learning to teach is a process that involves individual beliefs and values that develop over time as teachers interact with the realities of schools. This notion of teaching and learning also informed this dissertation.

**Research Participants and Sampling**

The participants for this study are a subset of the larger QCS participants. Since this study was interested in how teacher candidates/teachers engaged pupils in authentic intellectual work over time (i.e., during their preservice and first two years of teaching), only QCS participants who completed the teacher preparation program, taught for at least 2 years following their preparation program, and continued to follow all aspects of the QCS research design protocol were selected ($n$=11). Due to attrition, not all of the 22 QCS participants met these requirements; some participants did not complete the teacher education program ($n$=1), delayed program completion ($n$=3), moved out of the geographic area of the study ($n$=2), did not continue with the third year of the study ($n$=1), did not teach within two years after the teacher education program ($n$=2), or did not teach after their first year of teaching ($n$=2). The 11 participants in this study include female and male elementary and secondary teachers who span subject areas and school contexts. Table 3.1 describes the participants in terms of these characteristics. It is

important to note here that throughout the QCS study, all participants were identified by pseudonyms only. The identities of the 11 participants in this study were unknown to this researcher.

Table 3.1

*Study Participants*

| Participant | Gender | Teacher Ed Program of Study | Teaching Grade/Subject | School Context* |
|---|---|---|---|---|
| Amanda | F | M.Ed. Secondary English/TELL certificate | Secondary English/Humanities | Title 1 |
| Craig | M | M.Ed. Secondary Science | High/Middle School Science | Non-Title 1 |
| Elizabeth | F | M.Ed. Secondary English | Secondary English | Title 1 |
| Lola | F | M.Ed. Elementary | Elementary/Middle School Science | Charter |
| Mara | F | M.Ed. Secondary History | Secondary History | Non-Title 1 |
| Mark | M | M.Ed. Secondary History | Secondary Science/Secondary History | Title 1 |
| Matt | M | M. Ed., Secondary Math | Secondary Math | Title 1 |
| Rachel | F | M.Ed. Elementary | Elementary | Title 1 |
| Riley | F | M.Ed. Elementary | Elementary | Title 1 |
| Sonia | F | M.Ed. Elementary | Elementary (bilingual class) | Title 1 |
| Sylvie | F | M.Ed. Elementary | Elementary/ESL pull-out | Title 1 |

* Refers to whether the school is in a district that receives federal "Title 1: Improving the Academic Achievement of the Disadvantaged" funding for schools that may be high-poverty or low-achieving.

**Data Collection**

This study draws upon data collected in the QCS study to explore teacher candidates'/teachers' and their pupils' engagement with authentic intellectual work. Participants in the QCS study took part in 12 interviews over a 3-year period throughout the one year of the teacher education program and during the first two years of teaching.

These semi-structured interviews were designed to get at issues related to teacher preparation and learning to teach in a program that emphasized social justice and pupil learning. Embedded within five of these 12 interviews was a series of questions related to assessment and pupil learning, forming the Teacher Assessment/Pupil Learning (TAPL) protocol. The TAPL protocol became a standard set of questions and procedures that participants were asked during interviews that also touched upon other topics throughout the course of the QCS study. Participants took part in five TAPL interviews at five different points in time: 1) at the end of student teaching (Interview 5[2] of the QCS research design); 2) in the fall/winter of the first year of teaching (Interview 8); 3) in the spring of the first year of teaching (Interview 9); 4) in the fall of the second year of teaching (Interview 10); and 5) in the spring of the second year of teaching (Interview 11). Table 3.2 provides a description of the 12 QCS interviews and illustrates how the TAPL protocol fit within the larger QCS study. (See Appendix A for all 12 interview protocols in their entirety.) From this point forward, the TAPL interviews will be referred to by time-point: Student Teaching, Year 1 Fall, Year 1 Spring, Year 2 Fall, and Year 2 Spring.

---

[2] The title of the interview (i.e. "Interview 5") refers to the Qualitative Case Studies (QCS) research design protocol where each interview was assigned a number. The interview titles used here come from the QCS design in an explicit attempt to maintain this study's embedded nature within the larger QCS study.

Table 3.2

*QCS Interviews*

| | Interview | Interview Topics |
|---|---|---|
| **Teacher Education** | **1** – Personal History and Education Experience<br><br>(Beginning of program) | Background<br><br>• Educational Experiences<br><br>• Beliefs<br><br>• Knowledge<br><br>• Practice (Expectations for Teaching) |
| | **2** – Pre-Practicum Experience<br><br>(Fall semester) | Pre-practicum Experience<br><br>• School and Classroom Community<br><br>• Cooperating Teacher<br><br>• University Pre-practicum Supervisor<br><br>• Students<br><br>• Social Justice<br><br>• Theory and Practice |
| | **3** – University Coursework<br><br>(End of fall semester) | Course Experiences<br><br>• Foundations<br><br>• Methods |
| | **4** – Full Practicum Experience<br><br>(Beginning of student teaching) | Full Practicum<br><br>• School and Classroom Community<br><br>• Role in the Classroom<br><br>• Cooperating Teacher<br><br>• Clinical Faculty Supervisor<br><br>• Students<br><br>• Teaching for Social Justice |
| | **5** – Pupil Learning<br><br>(End of student teaching) | • TAPL 1<br><br>• General Thoughts on Pupil Learning<br><br>• Inquiry |
| | **6** – End of Teacher education<br><br>(End of Teacher Education Program) | • Beliefs about Learning<br><br>• Impact of Teacher Education Program<br><br>• Teaching for Social Justice<br><br>• Future Plans<br><br>• Program Feedback |
| **First Year of Teaching** | **7** – First Year Teaching 1<br><br>(November of first year) | • School, Classroom, and Students<br><br>• Teaching for Social Justice<br><br>• Pupil Learning<br><br>• Colleagues |

| Interview | Interview Topics |
|---|---|
| | • Induction/mentor |
| | • Influence of Teacher Preparation Program |
| **8** – First Year Teaching 2 (February/March of first year) | • TAPL 2 |
| | • General Pupil Learning Ideas |
| | • Experiences in Classroom/School |
| **9** – First Year Teaching 3 (End of First Year) | • TAPL 3 |
| | • Ownership and Distancing |
| | • Social Justice |
| | • School Context/Teacher Roles |
| | • Inquiry |
| | • Future Plans |
| **Second Year of Teaching**   **10** – Second Year Teaching 1 (November/December of second year) | • TAPL 4 |
| | • Pupil Learning |
| | • Workload and Planning |
| | • Relationships with Colleagues |
| | • Influence of Teacher Preparation Program |
| | • Teaching for Social Justice |
| | • Inquiry |
| | • Personal Quality of Life |
| **11** – Second Year Teaching 2 (End of second year) | • Beliefs on Teaching |
| | • Induction |
| | • School Context |
| | • Teaching for Social Justice |
| | • Student Goals |
| | • Future Plans |
| | • TAPL 5 |
| | • Teacher Development |

The TAPL protocol, represented in Figure 3.2, consists of two main components. For each TAPL interview, the interview data and work samples were collected at the same time reflecting a concurrent mixed methods design. First, in each of the five TAPL interviews, participants were asked to bring to the interview and submit two lead-up

assessments and one culminating assignment/assessment of their choice and a full class

set of the pupil work produced in response to those assessments. Assessments could be

any assignment, lesson, activity, or project that teacher candidates/teachers used to assess

pupil learning. During the interviews, participants were asked questions about the

creation and implementation of the assessment; their learning goals for pupils and

evaluation of pupil learning; how well the assessment worked and might be modified in

the future; and what they considered "high," "medium," and "low" examples of pupil

performance. The interview and collection of assessments and pupil work is the first part

of the TAPL protocol. It is referred to in this study as the TAPL "internal" evaluation

because the interview allows teacher candidates/teachers to talk about and explain their

assessment practices and understandings of pupil learning. (See Appendix B for the

TAPL interview protocols.)



| **Evaluation of Learning Opportunities and Pupil Performance** | | |
|---|---|---|
| **Internal Evaluation** | *Samples of Assessments and Pupil Work* | **External Evaluation** |
| *Interview Question Categories:*<br><br>**1. Description of Assessments and Classroom/School Context**<br>Questions related to the creation, implementation, sequencing, and rationale for use of the assessment.<br><br>**2. Evaluation**<br>Questions related to pupil learning goals on the assessment and teacher understandings of how pupils met goals and demonstrated proficiency.<br><br>**3. Change**<br>Questions related to how well the assessment worked and how it might be altered for future use.<br><br>**4. High, Medium and Low**<br>Questions related to the selection of high, medium, and low examples of pupil work; the context related to these examples and pupils; and comparisons between these pupils' performance and the performance of the whole class. | Lead up and culminating assessments<br><br>↓<br><br>Class set of pupil work on culminating assessment<br><br>↓<br><br>Teacher selected examples of high, medium and low | *Authentic Intellectual Work Criteria*<br>(Newmann and Associates, 1996)<br><br>**Standard 1. Construction of Knowledge**<br>The assessment requires that pupils analyze, interpret, and/or evaluate relevant content information rather than simply reproduce it.<br><br>**Standard 2. Disciplined Inquiry**<br>The assessment asks pupils to elaborate on understandings, draw conclusions, construct arguments relevant to content area and incorporate extended written communication based on suitability for content.<br><br>**Standard 3. Value Beyond School: Connection to Students' Lives**<br>The assessment demands pupils to engage in problems or topics they are likely to encounter their daily lives and asks them to make connections with the content to their life outside the classroom. |

*Figure 3.2.* Teacher Assessment/Pupil Learning (TAPL) Protocol.

The second part of the TAPL protocol, which is referred to here as the "external"

TAPL evaluation, involves an independent evaluation by researchers of both the teachers'

assessments/assignments and the pupil work submitted by the teacher

candidates/teachers. This evaluation analyzes the authentic intellectual quality of learning

opportunities and also the pupil work produced in response. This external evaluation uses

an established rubric (RISER, 2001) based on Newmann and Associates' (1996)

framework of authentic intellectual work to score the assessment tasks and pupil work

samples. As noted in the literature review, the framework of authentic intellectual work is

based on three standards: 1) construction of knowledge, 2) disciplined inquiry, and 3)

value beyond school. The standards are further divided into different criteria for

assessment tasks and pupil work, based on subject area. The rubric is applicable for both

elementary and secondary grades and includes scoring criteria for assessment tasks and

pupils' work for four subjects (English Language Arts/Writing, Math, science, and social

studies). For example, the three criteria for scoring writing assessments tasks are 1)

*construction of knowledge*, 2) *elaborated written communication*, and 3) *connection to*

*students' lives*. The three criteria for scoring pupil work in math are 1) *mathematical*

*analysis*, 2) *disciplinary concepts*, and 3) *elaborated written communication*. The rubric

accounts for and differentiates aspects of authentic intellectual work as they pertain to the

creation of assessment tasks, pupil work outcomes, and subject matter. Scoring for each

criterion is based on a 3- or 4-point scale, depending upon the criterion/standard. Scores

from the three standards are added together for one total score. Thus, assessment scores

can range from 3.0 to 10.0 while pupils' work scores can range from 3.0 to 12.0. A detailed description of the rubric follows in Chapter 4. Members of the QCS team were trained to score assessment tasks and pupil work on the rubric in an intensive two-day training session led by one of the developers of the rubric, followed by a number of sessions among team members to establish inter-rater reliability.

Artifacts were scored by QCS researchers who worked in pairs according to subject matter expertise. There were 10 pairs of judges. Periodic checks were scheduled to insure the same reliability throughout the scoring process. Fifty percent of the assessment tasks were double coded and 30% of the samples of student work were double coded. Scores that were double coded were averaged together. Interrater reliability was calculated with a consensus estimate of percent agreement. A consensus estimate demonstrates the extent to which raters agreed on their ratings (Stemler & Tsai, 2008). An average of the different pairs of judges' percent agreement was used. The percent exact agreement (where raters had the exact same score for each item they scored) for this study was 59.95%. The percent adjacent agreement (where raters' scores were within one-point of one another for each item they scored) was 95.82%. Raters were consistently within one point for each sample scored. Raters were also consistent in their scorings. An intraclass correlation estimates the consensus and consistency at which raters differ and "represents the ratio of within-subject variance to between-subject variance on a rating scale" (Stemler & Tsai, 2008, p. 39). Intraclass correlation coefficients that are higher than .70 are generally acceptable and close to 1 suggests almost complete agreement. The

intraclass correlation for the 10 pairs of judges ranged from .909 to 1. Table 3.3 lists the

scores for each pair of judges. This indicates that scorers were consistent in their ratings.

Table 3.3

*Interrater Reliability Intraclass Correlation*

| | Pairs of Judges | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Intraclass Coefficient | .922 | .938 | .977 | .942 | .927 | .971 | .966 | .946 | .909 | 1 |

After investigating a number of other instruments to evaluate teacher

candidate/teacher and pupil learning, including the SOLO taxonomy (Biggs & Collis,

1982) and the framework of Productive Pedagogies (QSRLS, 2001), QCS researchers

selected Newmann's framework of authentic intellectual work to use in the TAPL

external evaluation for a number of reasons. As described previously in the literature

review, the concept of authentic intellectual work places an emphasis on higher-order,

deep, meaningful, and relevant learning experiences embedded within the classroom

which QCS researchers identified as consistent with the study's sociocultural theoretical

framework and the teacher preparation program's themes of constructivism and social

justice. In addition, QCS researchers wanted to use a pre-existing instrument that had

been used in previous studies in order to facilitate possible comparisons related to the

quality of authentic intellectual work between the study's participants and teachers in

other studies. Also, due to the nature of the case studies project, researchers wanted an

instrument that could be used to evaluate all participants and their pupils, which included

elementary and secondary pupils in all four major subject areas. Finally, although the QCS team considered utilizing the Productive Pedagogies framework because it is more reflective of the teacher education program's commitment to social justice, the team was dissuaded by the framework's 20-item rubric. Cognizant of limited resources, the small research team chose authentic intellectual work as a more manageable framework to use for the study.

As mentioned above, the main reason QCS researchers chose the framework of authentic intellectual work was that the framework aligned with the teacher education program. Gore, Griffiths, and Ladwig (2002) argue that in order for authentic intellectual work to take hold in teacher practices, the framework should be introduced and emphasized within the teacher education program. Although most of the teachers in this study were not explicitly introduced to Newmann's framework during their course of study, the fundamental concepts of authentic intellectual work were central tenets of the mission of the university preparation program. At the core of the program were the two themes of *promoting social justice* and *constructing knowledge* (Boston College Lynch School of Education, 2010). Through the first theme, "promoting social justice," the program conceptualized "teaching as an activity with political dimensions, and … see[s] all educators as responsible for challenging inequities in the social order and working with others to establish a more just society." These ideas support promoting democratic education by focusing on the role of teaching and society. The second theme of the teacher education program, "constructing knowledge," is also one of the three criteria for authentic intellectual work. In this case, the two are directly aligned. All of the courses in

this program incorporated these core values, promoting the importance of critically engaging with knowledge. In some courses, such as the elementary and secondary social studies methods courses, teacher candidates were directly introduced to the theory of authentic intellectual work through course readings by Fred Newmann and colleagues. Through this preparation program, teachers were prepared to consider one of the fundamental characteristics of authentic intellectual work.

This study analyzed data collected from five TAPL interviews for 11 participants over a 3 year period. This included the text of five structured interviews for each participant and the culminating assessment tasks and pupil work samples collected from each participant during the five TAPL interviews. Therefore, the study included an analysis of five assessment tasks for each participant. In a study on the validity of using assessment tasks as a measure of classroom practice, Clare and Aschbacher (2001) found that a sample of two assessment tasks from an individual teacher was a valid indicator of typical classroom practice and the type of learning opportunities that teachers presented. The use of five assessment tasks from each participant therefore provides a valid measure of typical classroom practice. A random sample of 10 pieces of pupil work from each assessment was used, and when there were fewer than 10 samples from a class, all work collected from that assessment was examined.[3] This sample selection is consistent with, and validated in, previous studies evaluating the quality of authentic intellectual work (King et al., 2001; Newmann, Lopez et al., 1998). 481 samples of pupil work were scored. All identifying pupil information was removed from the samples and replaced

---

[3] In a few instances, participants were unable to collect and submit work from all of their pupils. The number of pupils in each class varied among teacher candidates/teachers and school context.

with researcher-generated codes so that pupils could not be identified during the scoring process.

**Data Analysis**

As a mixed methods study, this study draws from a number of different data analysis procedures with the goal of incorporating the data together to construct a deeper understanding of the extent to which teacher candidates'/teachers' practices and their pupils' learning reflect authentic intellectual work. The integration of qualitative and quantitative data occurred to various degrees in the question, analysis, and interpretation phases. Single or multiple methods were used depending upon the research question. Therefore, this data analysis section is divided by research question and presents an overview of the methods that were used to answer each question. The qualitative and quantitative methods used to analyze each question were determined by the nature of the question. Only a brief overview of the methods of data analysis is presented here in this chapter. Because each research question draws from a different approach and the analysis was intricately connected to the method, a more detailed description of how the specific methods were used for each part of the analysis is incorporated within the discussion of each research question in Chapters 4 and 5.

Research questions (1) and (2) addressed the quality of authentic intellectual work that teacher candidates/teachers created for their pupils and the quality of authentic intellectual work of pupils' responses to these learning opportunities. These two questions involved the external TAPL. They were answered by scoring the assessment tasks and pupil work samples that participants submitted during the TAPL interviews.

Scoring the assessment tasks provided insight into the types of learning opportunities that teacher candidates/teachers offered their pupils to see the extent to which pupils had opportunities to engage in authentic intellectual work. An evaluation of pupil work samples indicated the extent to which pupils produced authentic intellectual work. These artifacts were scored according to the rubric for authentic intellectual work as described in the data collection section. Further explanation of the rubric is also provided in Chapter 4.

Since the assessment tasks and pupil work samples differed within and across participants, using this rubric to analyze the data served two purposes. First, it provided a framework to measure the quality of authentic intellectual work. Descriptive statistics were used to describe the range and variation of the authentic intellectual quality of learning opportunities and pupil outcomes based on the scores. Second, by applying this rubric to the various samples of work, it provided a means of quantitizing qualitative artifacts so that the diverse artifacts could be described in a standard manner. These scores were used in research questions (5) and (6) for within and across case comparisons.

Question (3) addressed the relationship between the quality of learning opportunities and the quality of pupil learning. Previous studies found that pupils of experienced teachers were more likely to produce higher levels of authentic intellectual work when they are given highly authentic learning opportunities (King et al., 2001; Newmann, Lopez et al., 1998). That is to say, an assessment task that scores high on the scale of authentic intellectual work is more likely to produce higher authentic intellectual

pupil work scores than low scoring assessment tasks. Conversely, if an assessment offers few opportunities for authentic intellectual work and scores low, pupils will have little chance to demonstrate authentic learning. Question (3) tests that relationship on beginning teachers and their pupils. A statistical correlation, "a statistical summary of the degree and direction of relationship or association between two variables" (Glass & Hopkins, 1996, p. 103), was run to determine the relationship between scores on the assessment tasks and pupil work scores. The hypothesis for this study is that there will be a positive correlation between TAPL assessment task scores and the TAPL pupil work scores.

To answer Question (4), which addressed the extent to which teacher candidates/teachers thought about the concepts of authentic intellectual work when they talked about learning opportunities and pupil learning over time, I used the internal TAPL interview as a way to examine teachers' goals for assessment and pupil learning. The internal TAPL allowed participants to explain why they created or used the assessments in the manner they did and describe their own understandings of pupils' learning on those assessments. This added an additional layer of data to explain these work samples. Typically, when researchers examine textual artifacts, such as these assessment tasks and pupil work samples, the researcher has to infer the creator of the artifact's intentions (Hodder, 1994). Without additional dialogue and "spoken critical comment from participants" (Hodder, 1994, p. 401), hallmarks to interpretive approaches that rely on participant input, researchers are challenged to account for the contexts of production and interpretation. This study incorporated participants back into the written texts by

providing an opportunity for participants to discuss their thoughts and explain their actions as to why they constructed and assessed the assessment tasks and pupil work.

In this study, I used interpretive qualitative research techniques to analyze the interview data. Although these techniques differ in specific structures (Charmaz, 2004; Glaser, 1992; Glaser & Strauss, 1967; Miles & Huberman, 1994; Strauss & Corbin, 1994), they are iterative processes that use "constant comparative analysis" with the data and draw on such systems as coding data, emerging themes, constructing conceptual categories, testing relationships between categories, analytic memos to refine categories, and generating inferences as to what is happening. TAPL interviews were cleaned and transcribed during the larger QCS project. Data analysis started by open-coding the TAPL interview data, followed by reviewing the data to construct categories related to authentic intellectual work. I continued to interrogate the data to determine possible relationships between categories to pose possible hypotheses related to the ways in which teacher candidates/teachers conceptualize authentic intellectual work. Specific examples of the codes that were used in the analysis are described in Chapter 5. The TAPL interviews were read in three different ways. I read the interviews by time-point (i.e. all of the interviews from the preservice period were read, followed by all of the interviews from the fall of the first year, etc.), by case (i.e. all five interviews for each case), and by external TAPL scores (i.e. all interviews related to assessments that had high scores on the external TAPL analysis, followed by all of the interviews related to assessments that had moderate scores on the external TAPL analysis, etc.) to further refine the categories. I then read the QCS case narratives for each case, which were written by other

researchers on the QCS team and summarized all of the QCS data for each participant, to triangulate my own interpretations with those of the other members of the QCS research team.

Questions (5) and (6) involved comparing and integrating the external and internal TAPL and using quantitative and qualitative analyses to construct a better understanding of how conditions or contexts influenced the level of authentic intellectual work over time. This helped determine whether teacher candidates engaged pupils differently in authentic intellectual work at any point in time from the preservice period to the second year of teaching. These questions involved three parts, although the analysis is not dependent upon the order. First, in the stage of data display, the external TAPL scores were placed on a matrix for each participant (Huberman & Miles, 1994). Displaying information numerically on a chart helped to illuminate patterns of authentic intellectual work within and across participants (Sandelowski, 2001).

Second, to help describe the patterns of the matrix, a repeated measures analysis of variance (ANOVA) was conducted on the external TAPL scores to determine if there were any statistical differences between candidates' TAPL scores over time. Two repeated measures ANOVA tests were conducted: one looked at participants and their assessment task scores over time and one looked at participants and their pupils' work scores over time. When a subject is measured on the same variable over time, it is referred to as repeated measures or a randomized block design (Girden, 1992), where the subject is compared to itself. In this case, the participants were the random factor and the TAPL scores from the five different interview time-points were the fixed repeated

130

measures factor (Glass & Hopkins, 1996). Taking into account independence of observations and sphericity, an F statistic was calculated to determine whether there were differences in assessment task scores or pupil work scores during the five points in time (Leech, Barrett, & Morgan, 2008). An *a priori* power analysis was conducted. Onwuegbuzie and Leech (2004) suggest that "a medium effect size is appropriate for use in most studies because of its combination of being practically meaningful and realistic" (p. 207). Using Cohen's (1992) suggested medium effect size for one-way analysis of variance (.25) and level of significance at .05, a sample size of more than 21 is required to achieve power of .80. This study's small sample ($n$=11) was considered in the interpretation of these results in Chapter 4. In addition, I analyzed the qualitative data from the internal TAPL interviews to provide possible explanations as to the variations in the quality of authentic intellectual work from the external TAPL scores across the five time-points.

Finally, I drew from the analysis of the six individual research questions to consider the larger research question of how teacher candidates/teachers reflect democratic education through authentic intellectual work as they learn to teach. Here I integrated the qualitative and quantitative methods during the interpretation stage to see if there were similarities or contradictions among the different data analyses. For example, a participant may speak highly of an assessment she used but it might have scored low in terms of authentic intellectual quality. An analysis of the interview data provided insight into the participant's conceptions about teaching and learning. In this case, the external and internal TAPL data were used to triangulate data. To illustrate how internal and

external TAPL data added to the understanding of beginning teachers' experiences with authentic intellectual work, two case studies are presented that show how these factors interacted in a classroom where the teacher was able to implement authentic intellectual work and a classroom where the teacher did not implement authentic intellectual work.

## Integrity of the Study

One of the challenges of mixed methods research is that qualitative and quantitative paradigms are grounded in differing assumptions regarding the fundamental principles of what makes research valid. Some researchers suggest that using the usual nomenclature to discuss the integrity of a study from the different approaches is inappropriate for mixed methods and that there should be new terms that transcend and connect both paradigms (Teddlie & Tashakkori, 2003). On the other hand, others suggest that each aspect of a mixed method study should hold fast to its own paradigmatic guidelines (Morse, 2003; Morse & Niehaus, 2009), suggesting that qualitative and quantitative analyses should be able to stand alone as sound and rigorous scholarship according to the standards for their paradigms when taken out of the mixed methods context. Qualitative researchers have taken issue with applying quantitative measures of integrity such as validity and generalizability to qualitative studies because they do not fit within poststructuralist and critical discourse (Denzin & Lincoln, 2005). Although quantitative and qualitative researchers disagree about these ideas, one commonality is the understanding that sound research is dependent upon the quality and accuracy of the researcher's interpretations and analysis of the data and the transferability of inferences. In this case inference refers to "a final outcome of a study" which "may consist of a

conclusion about, an understanding of, or an explanation for an event, [a] behavior, [a] relationship, or a case" (Teddlie & Tashakkori, 2003, p. 35). Researchers need to be able to support the ways in which interpretations and inferences were made. The study here is transparent in my analyses and the multiple components provide an element of triangulation. I was able to compare participants' understandings of teaching and learning through their interviews over time, compare their thoughts to the artifacts they created and the external evaluation of those assessments and pupil work samples, and compare participants to one another. This provided multiple ways to locate consistencies and inconsistencies within the data to refine my interpretations.

**Rigor**

The "rigor" of a study is one example where the two paradigms diverge. In its most basic definition, rigor refers to the ways in which a study "adheres to commonly accepted best practices" (Teddlie & Tashakkori, 2003, p. 37). With quantitative approaches, this usually refers to whether the study can be replicated, whereas qualitative researchers, believing that experiences are socially constructed, negotiated, and contextual, suggest that replication is impossible and not even an ideal. Instead, rigor is determined by the quality of research methods used, a connection to a theoretical framework, the articulation of the connection between interpretations to the data, explicit reasons in support of these inferences, and the likelihood that another researcher with the same set of data would agree with the interpretations (Rossman & Rallis, 2003). Transparency of methodological choices and analysis is key. In the case of mixed methods, Teddlie and Tashakkori (2003) suggest that rigor be defined by how well a

study adheres to "best practices" and that researchers "describe a process whereby the accuracy, or authenticity, of our conclusions/interpretations is assessed" (p. 37). Such considerations include the extent to which the research design is consistent with the research questions; inferences are consistent conceptually and theoretically with each other and other scholarship; inferences offer distinct conceptions; and the possibility of other plausible explanations to account for the phenomena (Teddlie & Tashakkori, 2003).

The various pieces of this mixed methods study attempted to address those considerations. Separate research methodologies were chosen to analyze different research questions because those methods offered the best possible techniques to answer the questions. The multiple pieces also provided an opportunity to triangulate inferences within the study and the use of qualitative and quantitative techniques enabled comparisons with this study to other similar studies that used only one method. Continual analysis within and across cases also shed light on distinct instances and generated ideas related to other explanations.

**Reflexivity**

With qualitative research, researchers interact and make it explicit that their perspectives and experiences influence the interpretation of the data. My own experiences as a researcher, teacher, and pupil influenced the way in which I interpreted the interview data. Although I did not conduct the TAPL interviews and had no interaction with these participants or their pupils, I was a member of the QCS research team and have been involved in data analysis on these case studies throughout the study. Each bimonthly meeting of the QCS research team generated a collaborative data analysis process where

team members, including myself, continually reflected upon our interpretations of the data. Throughout this research study, I collaborated with other members of the research team to see if my interpretations of the data were consistent with their views.

**Validity**

The validity of a study has very different connotations in quantitative and qualitative research. In a general sense, validity refers to how well research conclusions reflect what is actually happening (Teddlie & Tashakkori, 2003) and "the trustworthiness of inferences drawn from data" (Eisenhart & Howe, 1992, p. 644). In quantitative research, internal validity is "the *sine qua non* of good experimental design" (Eisenhart & Howe, 1992, p. 644) and typically refers to the ability to link causal relationships between variables, ruling out other explanations (Teddlie & Tashakkori, 2003). External validity refers to the generalizability of results. Onwuegbuzie (2003) found that researchers have conceptualized more than 50 threats to internal and external validity. In qualitative research, "credibility" is used to refer to ways of making qualitative interpretations more authentic, and "transferability" is used to see how findings occur in different contexts (Teddlie & Tashakkori, 2003). Qualitative researchers stress that findings should be representative of the data and suggest examining confirming and disconfirming cases and using participant-checks to enhance credibility (Miles & Huberman, 1994). For mixed methods approaches, instead of using the term validity, Teddlie and Tashakkori (2003) recommend using the term *inference quality* to relate to the quantitative concept of internal validity and the qualitative concept of credibility and the term *inference transferability* to refer to external validity and transferability.

Onwuegbuzie and Johnson (2006) build on Teddlie and Tashakkori's notion of inference transferability to suggest the use of the term *legitimation* in mixed methods work. They list nine types of legitimation: sample integration, inside-outside, weakness minimization, sequential, conversion, paradigmatic mixing, commensurability, multiple validities, and political. A few pertain specifically to this study. One is *sample integration legitimation* which addresses difficulties in mixed methods related to studies where there is a small sample size due to the qualitative portion that may not be appropriate for quantitative analyses; researchers need to consider "how to combine legitimately different sets of people for use in making quality meta-inferences" (p. 58). Small sample size is one limitation of this study and I was transparent in the way in which I conducted and reported statistical analyses. In *conversion legitimation*, Onwuegbuzie and Johnson caution researchers about the need to scrutinize data that has been converted, such as quantitized qualitative data, to make sure that the use of such data is not misleading. One of the strengths of this study is the interplay between the internal and external TAPL analyses. In the internal TAPL interview, participants were given the opportunity to talk about the samples of assessment tasks and pupil work. Their analyses were triangulated with the external analysis as a way to check the representativeness of the TAPL external scores. Finally, Onwuegbuzie and Johnson use the term *multiple validities legitimation* to make sure that mixed methods researchers maintain high levels of validity, as measured by the particular method employed, for all portions of their research. In this study I adhered to quantitative standards of inter-rater reliability in the scoring of assessment tasks and pupil work samples. In addition, in regards to the qualitative interview data, I

reviewed my interpretations with team members from the larger QCS project, a study that draws on Hill, Thompson, and Williams' (1997) framework of the "consensual" approach to data collection and analysis where a team of researchers work together to come to "consensus judgments" (p. 521) regarding the data.

**Limitations**

Like any study, there were a number of limitations to this study. One challenge of mixed methods research is maintaining paradigmatic methodological integrity, especially with regard to sample size. This study had an inductive, qualitative theoretical drive. The sample purposefully came from a small qualitative case studies project where the intent was to examine in-depth experiences of a few. Although this study attempts to reveal a more general understanding of how beginning teachers enact authentic intellectual work by looking across 11 cases during a 3-year period, this is a small sample. Therefore, the quantitative analyses, in particular the correlation and repeated measures ANOVA, have a small $n$ and lack power. Morse (1991) argues that quantitative measures should not be used unless they have a randomized, representative sample or have a norm reference point outside of the study. This study does not attempt to rectify the non-randomized, non-representative sample but is transparent and cautious in reporting findings. In addition, however, the use of the rubric to score authentic intellectual work does provide a general point of comparison with other studies that have used the rubric (Bryk et al., 2000; King et al., 2001; Newmann et al., 2001) and analyses of the external TAPL scores could be compared to those other studies.

In conclusion, this inductive, qualitatively-driven concurrent embedded mixed methods study explores the extent to which teacher candidates/teachers engage pupils in authentic intellectual work during their preservice period and first two years of teaching and examine their pupils' responses to these learning opportunities by following 11 teacher candidates during their teacher preparation program and beginning years of teaching. This study addresses this by quantitizing assessment tasks and pupil work samples used by teacher candidates/teachers in their classroom through a rubric designed to measure the quality of authentic intellectual work. In addition, this study used a qualitative analysis of participant interviews where participants discuss the creation of their assessments and their understandings of pupil learning to further explore this topic. Both quantitative and qualitative data were used to construct an in-depth understanding of how beginning teachers engage pupils in authentic intellectual work and the extent to which their discussions of assessment and pupil learning reflect the principles of authentic intellectual work.

# CHAPTER FOUR: AUTHENTIC INTELLECTUAL QUALITY OF TEACHERS' ASSESSMENT TASKS AND PUPIL WORK

Amy Gutmann's (1987) theory of democratic education posits that American democracy rests upon citizens' abilities to deliberate on matters related to the collective good by making informed, rational decisions. She argues that schools are vital in maintaining this system. The responsibility of schools and teachers, therefore, is to develop citizens capable of making these decisions by providing opportunities for their students to develop skills such as critical analysis and higher-order thinking. Teacher education plays an important role in preparing teachers to be able to do this type of work (Michelli, 2005). This study looked at the extent to which teacher candidates/teachers, from a teacher education program whose mission was consistent with these goals, engaged their pupils in the skills that Gutmann advocates during the preservice period and the first two years of teaching. To get at the complexities of how beginning teachers prepare their students for democratic participation, this study draws on quantitative and qualitative analyses of classroom artifacts, including teachers' assessments and pupils' work, and interviews.

This study used Fred Newmann's (Newmann & Associates, 1996) framework of authentic intellectual work to assess the skills associated with Gutmann's theory of democratic education. As described in Chapter 2, the framework of authentic intellectual work is an appropriate lens to assess the knowledge needed for democratic participation because it accounts for the cognitively complex activities that pupils encounter, skills that are relevant to participation in a democratic society. Newmann and colleagues explicitly

identify the underlying principles of authentic intellectual work as the abilities necessary to "participate in civic life" (p. 27). Newmann, King, and Carmichael (2007) expand on this connection between democratic education and authentic intellectual work:

> [T]he argument for democracy assumes that citizens are capable not only of basic literacy, but also of exercising principled and reasoned judgment about public affairs. Arriving at defensible positions on controversial public issues…all require interpretation, evaluation, in-depth understanding, and elaborated communication that extends well beyond traditional tests of knowledge. (Newmann et al., 2007, pp. 11-12)

This emphasis on interpretation, evaluation, and communication in the framework of authentic intellectual work is consistent with the theory of democratic education posed by Gutman and others. In the next two chapters I use the framework of authentic intellectual work to examine teachers' assessments and pupils' work, as well as their understandings of assessment and pupil learning, in terms of the knowledge and skills necessary for democratic participation.

This study presents the results of the Teacher Assessment/Pupil Learning (TAPL) analysis of 11 beginning teachers with the intent to examine the extent to which these teachers engaged their pupils in authentic intellectual work and how their pupils responded to those learning opportunities. This analysis consists of two parts: 1) the external TAPL analysis, where teachers' assessments and their pupils' work on those assessments were evaluated by researchers according to a rubric for authentic intellectual work, and 2) the internal TAPL analysis, a qualitative analysis of how and what these teachers talked about in interviews about their assessments and pupils' work on those assessments.

My analyses in this chapter and the next show that the degree to which the beginning teachers in this study and their pupils engaged in authentic intellectual work varied widely. Generally speaking, the teachers and their pupils were able to engage in moderate levels of authentic intellectual work. As I show, the overall quality of authentic intellectual work was influenced by teachers' beliefs, classroom contexts, and variables such as the content area that teachers taught and who created the assessments that teachers used. Beginning teachers who valued authentic intellectual work were more likely to implement it in their classroom and have students who were able to produce authentic intellectual work. The next two chapters describe the ways in which this occurred.

First, I present the results of the external TAPL analysis. Here researchers evaluated teachers' assessment tasks and pupil work samples using a rubric for authentic intellectual work. The evaluations were then used to compare and contrast the quality of authentic intellectual work across teachers. This chapter first looks at teachers' assessment tasks, followed by the pupil work samples. These sections include a description of the tasks and pupil work as well as the results of statistical analyses of how assessment tasks and pupil work samples were scored in terms of the level of authentic intellectual work depending upon factors that have been considered in previous studies such as academic discipline, grade level, and school context. Finally, the chapter concludes by examining the relationship between the authentic intellectual quality of assessment tasks and the quality of pupil work.

The following chapter, Chapter 5, presents the results of the internal TAPL analysis. Here, teachers' interviews were analyzed to examine how teachers talked about teaching and learning in terms of the concepts underpinning the framework of authentic intellectual work and the focus of the analysis was on the relationship between teachers' values and beliefs and their teaching practices. The external TAPL analysis was considered with the internal TAPL data to construct a more complete understanding of how teachers, and their pupils, engaged in authentic intellectual work. This chapter examines how factors such as time and students' academic abilities influenced teachers' capacities to implement authentic intellectual work. It concludes with an extended example of how beliefs, practices, and contextual factors interacted in a case where a teacher was able to implement high levels of authentic intellectual work and a case where a teacher had very few expectations for authentic intellectual work.

**Learning Opportunities and the Quality of Authentic Intellectual Work**

Teaching for democratic participation requires that teachers engage pupils in activities that develop critical and higher-order thinking skills. To examine how beginning teachers think about and implement higher-order thinking in their teaching practice, this study focused on the learning opportunities that teachers provided for their pupils. In this case, the assignments that teachers used with their pupils, referred to here as *assessment tasks*, were used as evidence of practice, particularly the type of learning that beginning teachers expect or desire from their pupils.

The 11 teachers in this study were asked to submit one culminating assessment task that they used in their classroom during each of the five TAPL interviews in which

they participated. Teachers submitted one assessment task from their student teaching period, two assessment tasks from their first year of teaching (Year 1 Fall, Year 1 Spring), and two assessment tasks from their second year of teaching (Year 2 Fall, Year 2 Spring). Nine of the teachers submitted all five assessment tasks, while two of the teachers only submitted four assessment tasks.[4] A total of 53 assessment tasks were included in this study.

Table 4.1 lists the assessments that each participant submitted at each of the five TAPL interviews. As the description of the assessment tasks in Table 4.1 illustrates, the assessment tasks in this study spanned different grade levels and addressed all four major subject areas. There were 25 secondary assessments in writing, math, science, and social studies; eight middle school science assessments; and 15 elementary assessments in writing and math. Since this study was embedded within the larger QCS study (as described in the previous chapter), the sample was limited by the subjects and grade levels that the QCS participants taught during a 3-year period. Therefore, this sample does not include assessments in every subject area for every grade. However, as Table 4.1 demonstrates, the variety of subjects and grades included here provides a fairly representative selection of elementary and secondary teachers and their assessments.

---

[4] One of the participants who submitted four assessment tasks did not hold a teaching position for part of his second year of teaching and, therefore, did not participate in one of the TAPL interviews. The second teacher who submitted four assessments did participate in all five TAPL interviews but one of her assessment tasks was not available for this study.

Table 4.1

*Assessment Task Descriptions*

| Participant | Assessment Task Description | Subject | Grade |
|---|---|---|---|
| Amanda | | | |
| Student Teaching | Essay on the novel *Death of a Salesman* | Humanities | 12 |
| Year 1 Fall | State test preparation essay on the novel *Of Mice and Men* | Writing | 10 |
| Year 1 Spring | Prose poems and vignettes | Writing | 10 |
| Year 2 Fall | Essay comparing Enlightenment philosophers and the novel *A Tale of Two Cities* | Humanities | 12 |
| Year 2 Spring | Final exam | Humanities | 12 |
| Craig | | | |
| Student Teaching | Worksheet on heredity and genetics | Biology | 10 |
| Year 1 Fall | Solar system playing cards | General science | 6 |
| Year 1 Spring | Food chains and food webs poster | General science | 6 |
| Year 2 Fall | Layers of the Earth test | General science | 6 |
| Year 2 Spring | Ecosystems quiz | General science | 6 |
| Elizabeth | | | |
| Student Teaching | Essay on the novel *A Lesson Before Dying* | English | 11 |
| Year 1 Fall | Essay on the novel *And Their Eyes Were Watching God* | English | 11 |
| Year 1 Spring | Persuasive essay | English | 11 |
| Year 2 Fall | Advance Placement (AP) essay | English | AP |
| Year 2 Spring | Essay on the novel *The Things They Carried* | English | 12 |
| Lola | | | |
| Student Teaching | Standardized curriculum math test | Math | 5 |
| Year 1 Fall | Heat transfer quiz | General science | 7 |
| Year 1 Spring | Unit test on the Earth | General science | 7 |
| Year 2 Fall | Topographic map lab | General science | 8 |
| Year 2 Spring | Heat transfer quiz | General science | 8 |
| Mara | | | |
| Student Teaching | PowerPoint on Renaissance art | Humanities | 9 |
| Year 1 Fall | PowerPoint on World War I | Social studies | 10 |
| Year 1 Spring | Essay on genocide | Social studies | 10 |
| Year 2 Fall | Poster on 20th century ideologies | Social studies | 10 |
| Year 2 Spring | Current events and social action project | Social studies | 10 |
| Mark | | | |
| Student Teaching | District-created history exam | Social studies | 10 |

| Participant | Assessment Task Description | Subject | Grade |
|---|---|---|---|
| Year 1 Fall | Chemistry exam | Chemistry | 11 |
| Year 1 Spring | 50-meter dash poster | Physics | 9 |
| Year 2 Fall | Test on Westward Expansion and the Industrial Revolution | Social studies | 9 |
| Year 2 Spring | No assessment submitted | - | - |
| **Matt** | | | |
| Student Teaching | Math inequalities test | Algebra | 9 |
| Year 1 Fall | Math portfolio | Geometry | 8 |
| Year 1 Spring | Department math exam | Geometry | 10 |
| Year 2 Fall | Math chapter test | Geometry | 10 |
| Year 2 Spring | Math exam | Geometry | 10 |
| **Rachel** | | | |
| Student Teaching | Standardized math curriculum worksheet | Math | 2 |
| Year 1 Fall | Standardized math curriculum worksheet | Math | 3 |
| Year 1 Spring | No assessment available | - | - |
| Year 2 Fall | Textbook assignment on multiplication | Math | 3 |
| Year 2 Spring | Worksheet on subtraction | Math | 3 |
| **Riley** | | | |
| Student Teaching | Poetry anthology | Writing | 5 |
| Year 1 Fall | School wide mid-year math assessment | Math | 4 |
| Year 1 Spring | School wide year-end math assessment | Math | 4 |
| Year 2 Fall | School wide mid-year math assessment | Math | 4 |
| Year 2 Spring | School wide year-end math assessment | Math | 4 |
| **Sonia** | | | |
| Student Teaching | District math worksheet | Math | 4 |
| Year 1 Fall | District mid-year math test | Math | 2 |
| Year 1 Spring | District math final exam | Math | 2 |
| Year 2 Fall | District mid-year math test | Math | 2 |
| Year 2 Spring | District math final exam | Math | 2 |
| **Sylvie** | | | |
| Student Teaching | "Storybuddy" writing graphic organizer | Writing | 2 |
| Year 1 Fall | Project on field trip to State House | Writing (ELL) | 4/5 |
| Year 1 Spring | Biography project | Writing (ELL) | 4/5 |
| Year 2 Fall | Persuasive letter | Writing (ELL) | 4/5 |
| Year 2 Spring | Final project | Writing (ELL) | 4/5 |

As can be seen from Table 4.1, these beginning teachers in this study provided a variety of types of learning opportunities in a variety of formats for their pupils. These tasks ranged from the writing of an analytic literary essay to the creation of a PowerPoint presentation on World War I to the completion of a district end-of-year math exam and a science quiz on ecosystems. The formats for these assignments included written essays, multiple choice and short-answer responses, and hands-on group projects.

The assessments that the beginning teachers who participated in this study submitted reflect the learning objectives they had at different points in time during the preservice period and first two years of teaching. Individual teachers tended to submit the same type of assessment throughout this period. For instance, Elizabeth, a secondary English teacher, submitted five essay assignments, and Sonia, an elementary teacher, submitted five math assessments, four of which were standardized district mid-year or end-of-year exams. Matt, a secondary math teacher, submitted four similar math tests and one different math portfolio assignment where students were asked to write paragraphs about their experiences in math class. Matt's interview data confirm that the tests were the type of assessment that Matt used the most. He indicated that he discontinued the use of the portfolio assignment and relied, instead, on tests as a way to measure pupil performance. Only a few teachers submitted an assortment of assessment types. For instance, Mara, a secondary social studies teacher, varied her instruction by asking pupils to do assignments such as PowerPoints, posters, and essays. However, for the most part, these beginning teachers tended to implement the same type of assessments in their first two years of teaching.

Some of the range and variation of the assessment tasks presented in Table 4.1 can be explained by the QCS research protocol and procedures themselves. QCS participants were not told what type of assessments[5] to submit for the TAPL interviews. They were asked only to submit a "culminating" assessment, which was defined as a summative assessment used at the end of a teaching segment or unit. No additional details were given about the nature of the assessment and participants were not restricted to submitting a particular type of assessment.

One result of this is that not all assessments would necessarily be considered "summative" or "culminating." For example, Craig submitted a short quiz for his Year 2 Spring TAPL interview which was not an assessment that ended a unit. He was not prepared to submit an assessment at the time of the interview and simply submitted the assessment that he used in class close to the day of the interview. This formative quiz was not intended to assess pupil learning for the entire unit and is therefore less extensive than a culminating assessment. The QCS study operated under the assumption that since teachers center instruction around the learning objectives of a culminating assessment, the assessment would serve as a representation of major goals for the teaching segment or unit and reflect lessons leading up to the assessment. A non-culminating assessment, on the other hand, is more likely to be shorter in length, less complex, and non-representative of the larger unit learning goals. Preliminary analysis suggested that culminating assessments may provide more opportunities for authentic intellectual work than non-culminating assessments (Gleeson, Mitchell, Baroz, & Cochran-Smith, 2008).

---

[5] Participants were asked to submit one culminating and two lead-up assessments at each TAPL interview. The two lead-up assessments are not included in this study.

Other participants submitted assessments that were convenient as well. Sylvie, for instance, did not have access to a photocopier and submitted assessments that were displayed on the bulletin board that researchers could take pictures of. This contextual information needed to be taken into account in evaluating the quality of the assessments and is considered in the internal TAPL analysis in Chapter 5.

A second result of the research protocol instructions for the collection of assessment tasks is that assessments in this study differ with regard to who created the assessment. Some participants submitted assessments that they themselves created, which was the original intent of the QCS study. For example, Amanda created a final exam for her 12th grade humanities class and developed the essay questions on her own. Other participants, however, sometimes submitted assessment tasks that they co-constructed by working with other teachers, modifying other teachers' lessons, or modifying packaged curriculum. For example, on a quiz about heat transfer, Lola created some of the questions but also used questions from a state-wide standardized assessment. In addition, she also developed some of the questions with another eighth grade teacher and special education specialist. In cases like this, the assessment represents participants' ideas about assessment and authentic intellectual work but includes the input of other teachers as well. When participants submitted modified assessments, it was difficult to separate participants' input from the work of others. Finally, some participants submitted assessments that they had no role in creating such as a district-wide standardized exam or a textbook worksheet. Sonia submitted a math exam that her school district created as a final exam for all second graders. Sonia did not create this assessment, but it was an

assessment that she used with her pupils as a part of her math curriculum. The fact that participants did not always create the assessments they submitted made it difficult to determine participants' abilities to construct authentic assessments because the assessment did not reflect participants' own work.

Despite the fact that assessments contained different characteristics, all of the assessment tasks that were submitted were analyzed in this study. This decision was a deliberate one. The purpose of this study was to examine the extent to which beginning teachers who were prepared in a teacher education program that emphasized social justice and democratic practices engaged their pupils in authentic intellectual work. Therefore, the assessments that teachers used in their classroom, regardless of how they were created and why they were used, reflect the types of learning opportunities that were presented to students in their classroom. Teachers used those assessments. A teacher could still emphasize authentic intellectual work even though she did not create the assessment if the assessment engaged pupils in authentic intellectual work. Simply the act of choosing to use that assessment signifies that the teacher felt the assessment covered material that she felt was important. It becomes more difficult to ascertain teachers' understandings of authentic intellectual work when they submitted assessments that they did not create or that were mandated for use by their schools, but those assessments remain important to the study because they indicate what was actually happening in the classroom.

In order to address teachers' beliefs about authentic intellectual work in cases where the assessment task was not created solely by participants, this study used the qualitative interview data in the internal TAPL to gain insight into teachers' thoughts and

beliefs about the assessments they used and their pupils' learning on those assessments. In the interviews, participants were asked to discuss the assessment that they submitted and talk about their learning objectives. As a result, it was possible to ascertain some information about teachers' own thoughts related to authentic intellectual work even when they submitted assessments that they themselves did not create. The collection of multiple assessments over time and use of multiple data sources to account for authentic intellectual work provides an opportunity to triangulate data and determine how representative an assessment is of typical practice and assess teachers' understandings about the nature of authentic intellectual work. These data were used to help explain whether or not participants thought about teaching and learning in terms of authentic intellectual work even when they submitted assessments that they did not themselves create. Chapter 5 integrates the qualitative interview data with the quantitative analysis of the assessment tasks to explain why assessments were used. Chapter 4 focuses primarily on the quantitative data.

To compare the different types of assessments in the external TAPL analysis, researchers evaluated the level of authentic intellectual work for each assessment by using a pre-existing rubric (RISER, 2001) designed to measure authentic intellectual work. Assigning each assessment task a numeric score based on the rubric provided a way to quantitize the artifacts and use statistics to analyze the characteristics and differences among and between the assessments. Analyses were conducted to examine how the assessments addressed the different criteria of authentic intellectual work and to examine whether academic discipline, grade level, or school context influenced the

quality of authentic intellectual work. In addition, since not all participants created the assessments they used, an independent samples *t* test was conducted to determine whether there were statistical differences between the quality of authentic intellectual work in the assessments that participants created and the assessments that they did not create.

As mentioned in Chapter 3, the rubric used to measure the authentic intellectual quality of assessment tasks is comprised of three criteria: 1) construction of knowledge, 2) elaborated written communication, and 3) connection of students' lives. The rubric guidelines were designed to evaluate the subject areas of writing, science, math, and social studies across elementary and secondary grade levels to see the extent to which assessments in these disciplines met the qualities of authentic intellectual work. To illustrate what the rubric criteria for authentic intellectual work entails, Table 4.2 provides an excerpt of the standards for assessment tasks in writing. As seen in the table, the standard construction of knowledge focuses on how pupils interpret information, the standard elaborated written communication focuses on how pupils support their arguments in an extensive way, and the standard connection to students' lives focuses on how the assessment calls for pupils to address significant topics.

Table 4.2

*Authentic Intellectual Work Standards for Assessments in Writing (RISER, 2001)*

| Standard | Rubric Description |
|---|---|
| Standard 1: Construction of Knowledge | The assignment asks students to interpret, analyze, synthesize, or evaluate information in writing about a topic, rather than merely to reproduce information. |
| Standard 2: Disciplined Inquiry: Elaborated Written Communication | The assignment asks students to draw conclusions or make generalizations or arguments and support them through extended writing. |
| Standard 3: Value Beyond School: Connection to Students' Lives | The assignment asks students to connect the topic to experiences, observations, feelings, or situations significant in their lives. |

*Note.* Reproduced from *Standards and Scoring Criteria for Assessment Tasks and Student Performance*, by the Research Institute on Secondary Education Reform, 2001, Madison, WI: Author.

Furthermore, each of these standards has a list of scoring criteria used to rate assessment tasks. For example, Table 4.3 provides the scoring criteria for the standard elaborated written communication in the subject of writing. As Table 4.3 indicates, elaborated written communication is evaluated on a 4-point scale. The scoring criteria for the other two standards, which are not shown here, are based on a 3-point scale. Scores from each standard are added together for one total authentic intellectual work score. Possible scores for assessment tasks, therefore, range from 3.0 to 10.0, with 3.0 indicating the lowest level of authentic intellectual work and 10.0 indicating the highest

level of authentic intellectual work. All of the assessment tasks in this study were scored according to this rubric in the external TAPL analysis.

Table 4.3

*Scoring Criteria for Standard 2:* Elaborated Written Communication *in Writing*

Elaborated Written Communication

| Scale | Description |
|---|---|
| 4 | Explicit call for generalization AND support. The task asks students, using narrative or expository writing, to draw conclusions or to make generalizations or arguments, AND to substantiate them with examples, summaries, illustrations, details, or reasons. |
| 3 | Call for generalization OR support. The task asks students, using narrative or expository writing, to either draw conclusions or make generalizations or arguments, OR to offer examples, summaries, illustrations, details, or reasons, but not both. |
| 2 | Short-answer exercises. The task or its parts can be answered with only one or two sentences, clauses or phrasal fragments that complete a thought. |
| 1 | Fill-in-the-blank or multiple choice exercises. |

*Note*. Reproduced from *Standards and Scoring Criteria for Assessment Tasks and Student Performance*, by the Research Institute on Secondary Education Reform, 2001, Madison, WI: Author.

My analysis of the assessment task scores suggests that the teachers in this study engaged pupils in some of the skills necessary for democratic education. Across the teachers, they incorporated authentic intellectual work in their assessments, although the quality was not at the highest possible level. Each assessment task was rated according to

the rubric and given a total authentic intellectual work score. The mean of the total

authentic intellectual work score for all of the assessments in this study was 6.59 ($n$=53,

$SD$=1.798), out of a possible score of 10.0. The mean falls just above the mid-point of the

scale which suggests that assessments invited pupils to engage in moderate levels of

authentic intellectual work. Assessments also had moderate expectations for pupils to

interpret information, support their arguments, and analyze significant issues. Table 4.4

lists the authentic intellectual work scores for teachers' assessments from the five

different TAPL interviews during the preservice period and first two years of teaching.

The mean for each participants' assessments and mean for all of the assessments at each

time-point is also included on the table. As shown in Table 4.4, assessment task scores

ranged from 4.0 to 10.0. Some assessments were rated at the lowest end of the scale,

which means there was very little expectation of authentic intellectual work. On the other

hand, some tasks received the highest possible score and thus had high expectations for

authentic intellectual work.

Table 4.4

*Assessment Task External TAPL Scores*

| | Assessment Task External TAPL Scores | | | | | | |
|---|---|---|---|---|---|---|---|
| | Student | | Year 1 | | Year 2 | | |
| Participant | Teaching | Year 1 Fall | Spring | Year 2 Fall | Spring | *M* | *SD* |
| Amanda | 9.0 | 8.0 | 9.0 | 8.5 | 6.0 | 8.10 | 1.24 |
| Craig | 7.0 | 4.0 | 6.0 | 4.0 | 3.0 | 4.80 | 1.64 |
| Elizabeth | 9.0 | 8.0 | 10.0 | 8.0 | 8.0 | 8.60 | .894 |
| Lola | 5.0 | 6.0 | 5.0 | 9.0 | 6.0 | 6.20 | 1.64 |
| Mara | 7.0 | 7.0 | 9.5 | 6.5 | 9.5 | 7.90 | 1.48 |
| Mark | 6.0 | 6.0 | 7.0 | 4.0 | No Assessment[†] | 5.75 | 1.26 |
| Matt | 5.0 | 6.0 | 4.5 | 5.0 | 4.0 | 4.90 | .742 |
| Rachel | 9.0 | 5.0 | No Assessment[*] | 6.0 | 5.0 | 6.25 | 1.89 |
| Riley | 10.0 | 5.0 | 5.0 | 5.0 | 5.0 | 6.00 | 2.24 |
| Sonia | 5.5 | 6.0 | 7.0 | 6.0 | 6.5 | 6.20 | .570 |
| Sylvie | 8.5 | 6.5 | 7.0 | 9.5 | 6.5 | 7.60 | 1.34 |
| *M* | 7.36 | 6.14 | 7.00 | 6.50 | 5.95 | 6.59 | |
| *n* | 11 | 11 | 10 | 11 | 10 | 53 | |
| *SD* | 1.818 | 1.227 | 1.958 | 1.975 | 1.878 | 1.798 | |

[†]Mark only participated in four TAPL interviews because he was not employed as a teacher for a portion of this study.

[*]Rachel's assessment task for Interview 9 was not collected.

As the external TAPL scores for assessment tasks in Table 4.4 suggest, although the authentic intellectual quality of assessment tasks varied across participants, some individual teachers tended to create and use assessments that were of similar quality over the 3-year period. For example, Elizabeth's assessment tasks consistently ranged from 8.0 to 10.0 and were thus at the highest end of the authentic intellectual work scale. The assessment that she used during student teaching was one point above the assessment that she used at the end of her second year of teaching; the difference in quality between the

two is minimal. Similarly, Matt's assessments ranged from 4.0 to 6.0 over the 3-year

period and were toward the lower end of the authentic intellectual work scale. Like

Elizabeth, Matt had a one point difference between his student teaching and Year 2

Spring assessment scores. Neither Elizabeth's nor Matt's assessments changed radically

over time. Other teachers had a greater range in scores. Lola's assessments, for example,

ranged from 5.0 to 9.0. There was less consistency in the quality of her assessments. The

assessment Lola used in student teaching had the lowest score, along with the assessment

she used in the spring of her first year, while the assessment that she used in the fall

semester of her second year of teaching scored the highest. One reason for the variation is

that the high scoring assessment was a science lab assessment where students engaged in

hands-on inquiry; the other assessments were paper and pencil tests or worksheets. As I

describe in Chapter 5, Lola wanted to engage in more critical thinking in her second year

of teaching and had better classroom management which enabled her to implement more

lab activities. This is one example of how contextual factors influenced the quality of

teachers' assessments.

Figure 4.1 illustrates the authentic intellectual quality of each participants'

assessment tasks, based on the mean of the five submitted assessments. As seen in this

continuum, teachers fell into three categories based on these scores. With means of 4.80

($SD$=1.64) and 4.90 ($SD$=.742), respectively, Craig and Matt had the lowest levels of

authentic intellectual work. Five of the teachers provided learning opportunities that were

rated in the middle with means ranging from 5.75 ($SD$=1.26) to 6.25 ($SD$=1.89). Finally,

4 of the teachers had higher levels of authentic intellectual work, with means ranging

from 7.60 (*SD*=1.34) to 8.60 (*SD*=.894). These groupings were used in the qualitative

analysis to compare cases over time in Chapter 5.

## Authentic Intellectual Quality of Assessment Tasks



*Figure 4.1*. Continuum of the authentic intellectual work quality of teachers' assessment tasks.

As evidenced by the scores and the continuum in Figure 4.1, some teachers were

able to utilize assessments that were more authentic than other teachers. What is less

clear about the scores is whether the beginning teachers in this study had different levels

of authentic intellectual work as they progressed through the 3-year period. As seen on

Table 4.4, the means for assessment tasks at each time-point were different. For example,

the mean for all of the assessments at the student teaching period was 7.36 (*n*=11,

*SD*=1.818) and the mean for all of the assessments at Year 2 Spring was 5.95 (*n*=10,

*SD*=1.878), which might suggest that the level of authentic intellectual work decreased

over time. To investigate this further, a repeated-measures analysis of variance

(ANOVA) was conducted on the assessment task scores to determine whether there were

statistical differences in the authentic intellectual work scores for teachers' assessments

during the five different time-points (Student Teaching, Year 1 Fall, Year 1 Spring, Year 2 Fall, and Year 2 Spring). In other words, this analysis was used to detect if there was a general trend in teachers' assessments scores starting from their student teaching into the end of the second year of teaching.

For this repeated-measures ANOVA, teachers were the random factor and the assessment task authentic intellectual work scores for each time-point were the fixed repeated measures factor. Since Rachel and Mark only had scores for four out of the five time-points, missing values were replaced with the mean for their four assessment task scores. In this analysis, the null hypothesis is $H_0$: $\mu_1 = \mu_2 = \ldots = \mu_5$, meaning that there were no differences in assessment task scores for each of the 5 time-points. The alternative hypothesis, therefore, is $H_a$: $\mu_i \neq \mu_i$ for some $i$, $k$, meaning that there is a difference in scores between two of the time-points. The probability level was set at $\alpha =$ .05. Independence of observations and sphericity were met. Results indicate that the assessment task scores were not statistically different over the five time-points $F$ (4, 40) = 2.540, $p$=.055, eta$^2$ = .203 (see Table 4.5) and the null hypothesis was not rejected. The *a priori* power analysis, described in Chapter 3, suggested that a sample size of 21 was needed in order to achieve a medium effect size (.25) and power of .80. Since this repeated measures ANOVA had a sample size of 11 and the actual power was .48, it is possible that the non-significance was a result of the small sample size. This measure should be replicated with a larger sample to test this further.

Table 4.5

*Repeated Measures ANOVA Tests of Within-Subjects Effects on Assessment Task Authentic Intellectual Work Scores*

| | Type III Sum of Squares | *Df* | Mean Square | *F* | Sig. |
|---|---|---|---|---|---|
| Time-points | 17.958 | 4 | 4.489 | 2.540 | .055 |
| Error (Time-points) | 70.710 | 40 | 1.768 | | |

*Note.* See Table 4.4 for the assessment task scores used in this RMANOVA analysis.

The results of the repeated-measures ANOVA suggest that there were no overall trends in teachers' assessment task scores. That is, the beginning teachers in this study, as a whole, probably did not increase or decrease their levels of authentic intellectual work as they developed as teachers from student teaching to the end of their second year of teaching. There does not appear to be a general pattern where teachers at each time-point consistently increased, or consistently decreased the level of authentic intellectual work in their assessments. However, as the differences between means from the assessments at student teaching and the assessments in Year 2 Spring suggest, some teachers experienced change over the three years, a phenomenon that a repeated-measures ANOVA cannot account for fully. Therefore, it is necessary to further explore how individual teachers implemented authentic intellectual work during this period to see if there are additional patterns or trends that may account for teachers' scores. This analysis is presented in Chapter 5.

Although the mean for all of the assessment tasks suggests that these beginning teachers, as a whole, only engaged their pupils in moderate levels of authentic intellectual work, their scores are actually consistent with the scores of veteran teachers as measured

in other studies. In their two studies of veteran teachers who were identified prior to the study as having authentic practices, King, Schroeder, and Chawszczewski (2001) used the same rubric to find that the mean scores for these teachers' assessment tasks were 6.53 (*n*=16, *SD*=1.33) and 7.30 (*n*=35, *SD*=2.09), which the researchers characterized as moderate levels of authentic work. The beginning teachers in this study had a mean assessment score (*M*=6.59) that was similar to the veteran teachers in King et al.'s study. This suggests that beginning teachers from a teacher education program consistent with some aspects of authentic intellectual work were already performing at a similar level as veteran teachers who explicitly attempted to implement authentic intellectual work practices. However, since both the veteran and the beginning teachers in this study had assessment task scores that fell towards the middle of the rubric scale, this suggests that they were not providing the highest possible levels of authentic intellectual work for their pupils.

Other studies that used Newmann's framework of authentic intellectual work found similar results and concluded that teachers, in general, do not create or use authentic assessments. Bryk, Nagaoka, and Newmann (2000) conducted a 3-year longitudinal study of teachers involved in school restructuring efforts that emphasized authentic intellectual work. In the first year of the study, when teachers were introduced to the ideas of authentic intellectual work, the mean scores for teachers' assessment tasks ranged from 3.67 to 6.53.[6] These researchers concluded that teachers had low to moderate levels of authentic work when they started the school reform. At the end of the

---

[6] Bryk et al. (2000) reported the means of teachers disaggregated by subject and grade level.

3-year period of professional development, levels of authentic intellectual work increased, and teachers' assessment scores ranged from 5.19 to 7.77. When compared to the Bryk et al. study, the beginning teachers in my study engaged pupils in higher levels of authentic intellectual work than veteran teachers who were not trained to implement authentic intellectual work prior to professional development on the topic. The beginning teachers in this study also scored higher than some of the teachers in the Bryk et al. study after intensive professional development. This suggests that some of the teachers from this teacher preparation program engaged in practices consistent with the idea of authentic intellectual work at the start of their teaching career that were as high, and often higher than, veteran teachers. These beginning teachers were prepared to utilize and implement these practices when they began teaching.

Overall, my study, along with the King et al. (2001) and Bryk et al. (2000) studies, suggests that teachers, as a whole, do not engage pupils in high levels of authentic intellectual work. Many assessment tasks fail to ask pupils to use higher-order thinking, defend their ideas, or consider significant issues. By not focusing on these skills, teachers are thus not helping pupils learn to make good judgments or evaluate arguments, both of which are deliberative skills that are necessary in democratic societies.

Although teachers in this study, on the whole, engaged pupils in moderate levels of authentic intellectual work, there were certain aspects of this type of work that they were more likely to address than others. As previously mentioned, authentic intellectual work is evaluated on three criteria. As Table 4.6 indicates, assessment tasks were rated

differently, in a statistically significant way (*p*<.05), on each standard. That is to say, assessments attended to particular characteristics of authentic intellectual work more than others. Teachers were more likely to use assessment tasks that emphasized skills related to the standards construction of knowledge and elaborated written communication and less likely to consider the standard connection to students' lives.

Table 4.6

*Mean Scores for Assessment Tasks by Standard*

| Standard | *n* | *M* | *SD* | Range of Possible Scores | Percentile on Scale |
|---|---|---|---|---|---|
| Construction of Knowledge[a] | 53 | 2.20 | .540 | 1 to 3 | .733 |
| Elaborated Written Communication[b] | 53 | 2.61 | .993 | 1 to 4 | .653 |
| Connection to Students' Lives[c] | 53 | 1.78 | .724 | 1 to 3 | .593 |

[a]A paired samples *t* test showed that scores on the standard Construction of Knowledge differed significantly from scores on the standard Elaborated Written Communication, *t*(52)=4.056, *p*<.001, and the standard Connection to Students' Lives, *t*(52)=4.128, *p*=<.001.
[b]A paired samples *t* test showed that scores on the standard Elaborated Written Communication differed significantly from scores on the standard Construction of Knowledge, *t*(52)=4.056, *p*<.001, and the standard Connection to Students' Lives, *t*(52)=5.800, *p*<.001.
[c]A paired samples *t* test showed that scores on the standard Connection to Students' Lives differed significantly from scores on the standard Construction of Knowledge, *t*(52)=4.128, *p*<.001 and scores on the standard Elaborated Written Communication, *t*(52)=5.800, *p*<.001.

The standard construction of knowledge measures the extent to which the assessment "asks students to interpret, analyze, synthesize, or evaluate information in writing about a topic, rather than merely to reproduce information" (RISER, 2001, p. 4). This standard indicates how well an assessment asks pupils to apply their knowledge and construct their own new ideas. As Table 4.6 illustrates, the mean for assessment task

scores in this category was 2.20 (*n*=53, *SD*=.540), where possible scores ranged from 1.0

to 3.0. Based on the rubric, a score of 2.20 suggests that the majority of the tasks had

"some expectation for students to interpret, analyze, synthesize, or evaluate information"

(p. 4). For example, one assessment that received the highest possible score on this

standard asked pupils to choose one philosopher's theory and demonstrate how the theory

applied to a novel. Although pupils had discussed theories and the novel in class, this

assessment required that pupils apply what they had learned to construct their own

argument about the concepts. In contrast, one low scoring assessment asked only recall

questions. The assessment had multiple choice questions where students recalled

definitions. Even the single essay question asked pupils only to "describe the Populist

movement," which required them to reproduce information that they learned in class.

For the most part, however, the beginning teachers in this study created and used

assessments that fell in between these two examples. They used assessments that

moderately addressed higher-order critical thinking. There were times when teachers did

ask pupils to think critically about concepts and evaluate information, qualities that

contribute to democratic citizens' abilities to make judgments and assess issues about the

common good in society. This is a critical aspect of democratic education, one that, as

demonstrated by the mean on this standard, teachers could emphasize at an even greater

rate than they did on these assessments.

The beginning teachers in this study were less likely to engage pupils in

elaborated written communication. This standard measures the extent to which "[t]he task

asks students to draw conclusions or make generalizations or arguments and support them

through extended writing" (RISER, 2001, p. 5). This standard can be met by expanding and explaining concepts that are consistent with the discipline and can take a variety of forms. For instance, one example of elaborated communication is a writing task that asks pupils to write an essay, but another example is a math or science assessment where pupils were asked to create "graphs, tables, equations, diagrams, or sketches" (p. 11). The mean for all the assessment task scores in this study on the standard elaborated written communication was 2.61 ($n$=53, $SD$=.993), out of a possible range in scores of 1.0 to 4.0. My analysis revealed that assessment tasks were primarily short-answer exercises, involved some expectation for pupils to make a generalization about the topic or provide an example in their answers. For example, low scoring assessments on this standard included a fill-in-the-blank science quiz and a multiple choice math exam where pupils were not required to show their work. In contrast, a high score of 4.0 on this standard would involve asking pupils to make a generalization and support their answer. One high scoring example in this study asked pupils to write an essay analyzing a short story and use literary techniques to support their analyses.

Similar to the standard construction of knowledge, assessments on the standard elaborated written communication, for the most part, fell within the range of the two high and low examples described above and were able to address some aspects of elaborated written communication. They offered some opportunities for pupils to defend their positions and provide evidence to illustrate their ideas, skills used in democratic debate, but these opportunities were not offered all of the time and were accompanied by instances where pupils were not required to support their thinking in a systematic way.

The teachers in this study scored lowest on the standard connection to students'

lives, which addresses the extent to which tasks ask students to consider "a concept,

problem or issue that is similar to one that they have encountered or are likely encounter

outside of school" (RISER, 2001, p. 11). This standard measures how well the

assessment relates to important issues where pupils are able to see the relevance of the

work they do in class. The mean on this standard across assessment tasks was 1.78 ($n$=53,

$SD$= .724), out of a possible range of scores from 1.0 to 3.0. According to the rubric, a

score of 1.78 indicates that assessments offer minimal "opportunity for students to

connect the topic to experiences, observations, feelings, or situations significant in their

lives" (p. 5) and does not explicitly call for students to make this connection. The

assessment activities used by teachers in this study were, for the most part, isolated or out

of context and did not stress the relevance to pupils' lives. There were a few high scoring

assessments, such as a social action project where pupils were asked to select an issue

that they felt strongly about and create a project where they became involved in the issue,

but the majority of the assessments presented decontextualized content. For example, the

majority of the math assessments asked pupils to plug in formulas without providing a

scenario for how and why the formulas might be used, while writing assessments asked

pupils to analyze literature without relating the analysis to pupils' own experiences.

By failing to relate assessments to issues beyond the classroom, teachers did not

make explicit connections between what pupils were doing in class and how that material

operates in a "real world" context. Work performed in isolation does not foster pupils'

abilities to grapple with important societal issues when they are faced with making

decisions in a democracy. They are less likely to consider how different factors impact particular issues or may neglect to see the importance or ramifications of the issues they will be asked to consider.

Overall, my analysis suggests that high and low scoring assessment tasks had very different expectations for authentic intellectual work. As the previous discussion suggests, low scoring assessments provided minimal opportunities to critically engage with material and construct knowledge, whereas high scoring assessments challenged pupils to interpret information and formulate their own ideas. The following discussion of two high and low scoring assessments highlights how teachers engaged pupils in aspects of democratic education.

Figure 4.2 is an illustration of one of the low scoring assessments collected for this study. Craig used this quiz on ecosystems with his middle school general science pupils in the spring of his second year of teaching. This quiz was a non-culminating assessment that Craig submitted because he happened to have it available at the time of the TAPL interview. It received the lowest possible score on the rubric - 3.0 (out of a possible score of 10.0) – and was the lowest scoring assessment in this study. As the figure illustrates, this quiz was divided into two main sections. The first section consisted of 12 definitions where pupils chose terms from a word bank and wrote the terms next to the appropriate definition. The second section had five multiple choice questions followed by one final question asking pupils to explain their answer in writing.

**Ecosystem Interactions Quiz**

*Vocabulary.* Write the correct term from the word bank on the blank to the right of each definition.

*Word Bank*

| | | | |
|---|---|---|---|
| Behavior | Prey | Mimicry | Mutualism |
| Carrying Capacity | Predator | Camouflage | Commensalism |
| Defense Strategies | Symbiosis | Toxins | Parasitism |

1. Blending in with the background.

1. _____

2. Adaptations used by prey to avoid predators.

2. _____

3. An organism that is killed and eaten by another organism

3. _____

4. A relationship between two species in which both species benefit

4. _____

5. Chemicals used to irritate or poison a predator

5. _____

6. A defense strategy that involves a group or individual response to avoid a predator

6. _____

7. A relationship between two species in which one organism benefits and the other is unaffected

7. _____

8. An organism that eats all or part of another organism

8. _____

9. The largest population that an environment can support at any given time

9. _____

10. A relationship in which two different organisms live in close association with each other

10. _____

11. A relationship in which one species benefits and the other species is harmed

11. _____

12. Behavior or coloration that copies another organism which is actually more dangerous to the predator

12. _____

*Multiple Choice*

13. Young wasps are eating the tomato hornworm that is their host. What is this an example of?
        A. Commensalisms
        B. Mutualism
        C. Parasitism
        D. Competition

14. A bird eats a worm. Who is the predator?
        A. The worm
        B. The bird
        C. Both the bird and the worm
        D. Neither the bird nor the worm

15. Limiting factors determine an ecosystem's  carrying capacity for a population because
       A. The number of animals is limited.
       B. Ecosystems are small.
       C. Organisms need resources to survive.
       D. The number of animals is unlimited.

16. Two members of the same species fight over who gets a certain food. Members of different species try to take over a certain nesting area. These are both examples of
       A. Community
       B. Competition
       C. Mutualism
       D. Commensalism

17. In which type of symbiosis do organisms help each other?
       A. Parasitism
       B. Mutualism
       C. Community
       D. Commensalism

18. Certain ants take a sweet liquid called honeydew from tiny insects called aphids. In exchange for the honeydew, the ants protect the aphids from predators. What kind of relationship is this? Explain your answer.
_____
_____
_____

*Figure 4.2.* Low scoring assessment task. Craig's middle school general science quiz on ecosystems.

On the standard construction of knowledge, Craig's assessment received a score of 1.0, the lowest possible score. The majority of the questions on this assessment asked pupils to recall and reproduce previously learned definitions.  In the first section, pupils were asked to select the correct definition from a word bank; there were 12 questions and 12 words to fill in the blanks. Pupils simply recalled information and, since words could only be used once, pupils had a higher chance of guessing the appropriate answer when they did not know the answer. The multiple choice questions required a little more thinking. For example, in question 14 ("A bird eats a worm. Who is the predator?"), pupils had to apply their understanding of the concept "predator" to the situation of a bird eating a worm. The core of this question, however, is a basic definition.

On the standard elaborated written communication, Craig's quiz also received a score of 1.0, the lowest possible score. Question 18 is the only question on the quiz where pupils were asked to explain their thinking in writing. The remainder of the assessment consisted of 12 fill-in-the-blank and 5 multiple choice questions; the dominant expectation of this task was for pupils to answer with pre-stated answers. Pupils had no opportunity to express their understanding of the concepts or provide nuanced interpretations about how the concepts operate.

On the final standard, connection to students' lives, this task received another score of 1.0, the lowest possible score. Some of the questions represented scientific terms in isolation from the broader context. For example, in question 1, the definition of "camouflage" was described as "blending into the background." This definition was not specific to ecosystems nor did it suggest why camouflage is an important concept. Not all of the questions were as non-specific as question 1. For example, question 18 described a scenario related to the relationship between ants and aphids, a relationship that does occur in nature and one pupils could come into contact with outside of the classroom. However, this question still earned the lowest score according to the scoring guidelines. The rubric explicitly states that general disciplinary content knowledge does not automatically guarantee it is connected to pupils' lives "unless the task requires applying such knowledge to a specific problem likely to be encountered beyond the classroom" (RISER, 2001, p. 19). Question 18, therefore, is not explicit enough nor does it provide enough opportunity for pupils to apply information to be considered connected to pupils' lives. Overall, this task does not demonstrate the relevance or significance of ecosystems.

Craig's ecosystems assessment provided almost no opportunities for pupils to engage in authentic intellectual work. Rather, the assessment asked pupils to recall discrete pieces of information, without explanation. Pupils did not need to apply information or make a judgment. As a result, pupils did not produce high levels of authentic intellectual work in response to this assessment. Although a thorough explanation of pupil work is provided later in this chapter, it is worth noting here that the mean score for the pupil work on this assessment was one of the lowest out of all of the submitted assessment tasks. Pupils responded in the way they were asked to respond. In the first two parts, pupils filled in the blanks and chose answers from the multiple choice, sometimes selecting the wrong answers. For the written answer in question 18, most pupils responded by writing only a few words. Pupils did not elaborate in their answers and had few opportunities to demonstrate their understandings. This example suggests that assessments with low levels of authentic intellectual work provide fewer opportunities for pupils to demonstrate authentic intellectual work.

In contrast to low scoring assessments, assessments that received high scores on the external TAPL evaluation offered more opportunities for pupils to engage in skills necessary for democratic education. High scoring assessments gave pupils opportunities to create new knowledge by synthesizing and interpreting information; asked pupils to provide explanations for their thinking and to support their opinions with evidence; and were relevant to the world beyond the classroom, all skills that are necessary for democratic education.

Figure 4.3 provides an example of one high scoring assessment. Riley used this assessment as a fourth grade writing assignment during her student teaching semester. This assignment received a score of 10.0, the highest possible score and was one of the highest scoring assessments in this study. For this assignment, parts of which are not shown in Figure 4.3, pupils were asked to write at least six different types of poems which were then compiled into a poetry anthology booklet. In addition to creating their own poems, pupils were also asked to write extended passages about poems they had read in class and provide supporting evidence about their favorite poem based on their analysis.

<div style="border: 1px solid black;">

## *Poetry Anthology*

*Description:* Students wrote approximately six poems and compiled their poems into a Poetry Anthology booklet. Each poem was styled differently based on the students study of poetry forms. One type of poem in the Anthology was the Pyramid poem. Students had to create a poem that had a certain number of words for each line. The Anthology also included a worksheet where students wrote about their favorite poems.

*Pyramid Poem:* Pupils were given a worksheet in the shape of a pyramid. Pupils were told to write a word on each line. The first line is the topic of the poem and each line has one more word than the previous line. On the left is the pyramid worksheet that pupils were given. On the right is an example of a pyramid poem written by a pupil for this assessment.

|  |  |
|---|---|
| *Assignment* | *Pupil Example* |

Guitar
Frets Chords
red black white
Dad listening Mom listening
[Chad Bach Town] Music School
Eat a darn good breakfast early
Practicing over and over and over again
I touch strings my guitar and my bed
I smell my guitar music and my cool room
Purple Haze, Fire, Sunshine of Your Love, Stairway to heaven

*Poem Worksheet*

### Favorite Poem Worksheet*

Poem Title:

Poet:

Tell what the poem is about.

What did you notice about this poem?

Why did you choose this poem as one of your favorites?

</div>

*Note.* The original worksheet for this assessment had lines where pupils could write their answers to the questions. There was enough room for pupils to write their answers in paragraph form.

*Figure 4.3.* High scoring assessment task. Riley's fourth grade poetry anthology.

Riley's assessment received the highest score for each standard. On the standard construction of knowledge, this assessment task received a score of 3.0. The dominant

expectation for this assignment was for pupils to create new knowledge as they created their own poems and interpreted their understandings of the poem. For example, as Figure 4.3 illustrates, pupils were asked to write a pyramid poem where they chose their own topic and wrote a poem about it using a certain number of lines per stanza. Here pupils applied what they had learned in class about pyramid poems and what they knew about a particular topic to create a new poem.

The fact that pupils were asked to write many different poems and write a reflection also contributed to the assessment earning the highest score, 4.0, on the standard elaborated written communication. The poem worksheet, shown at the bottom of Figure 4.3, asked pupils to choose a poem that they had studied and write sentences explaining what the poem was about, what they noticed about the poem, and why they chose the poem as their favorite. This assessment asked pupils to demonstrate their interpretation of a poem they had studied, which went beyond recall, while supporting their argument behind why they liked the poem.

On the final standard, connections to students' lives, this task also received the highest score, 3.0. Pupils were asked to write poems about their own experiences and feelings, directly relating the poem to their lives, as the pupil's response in Figure 4.3 illustrates with the pyramid poem about this pupil's guitar. Here the pupil created a poem about his life outside of school and about the role of his guitar in his life. He used the poem assessment to reflect on something important in his life.

With this task, pupils had the opportunity to construct knowledge, elaborate and support generalizations, and draw upon important experiences in their own lives. Here,

particularly in the section where pupils were asked to provide a justification as to why they liked one poem better than another, they had the opportunity to make a judgment and use their knowledge about the topic to assess that conclusion. These opportunities led to higher levels of authentic intellectual work than those produced in response to Craig's ecosystem worksheet. In this case, the mean for pupil work scores was 7.33 (*SD*=.577), out of a possible score of 12.0. They created their own poems, critically reflected on why they liked the poems, and provided examples. As mentioned above, pupil work is examined in more detail later in this chapter, but it is worth noting here that this assessment illustrates that when pupils were given opportunities to engage in authentic intellectual work, they were able to produce higher quality work. This relationship between the assessment task and pupil performance is examined in the last section of this chapter.

Riley's assessment offers more opportunities to support democratic preparation by focusing on critical thinking and the construction of ideas than Craig's assessment which assesses pupils' basic understanding of concepts. This is not to say that content knowledge is not related to democracy. On the contrary, democratic deliberation requires conceptual understanding of the issues involved. However, the application and synthesizing of knowledge is imperative for making the best decisions for society and the common good because in a democracy citizens need to evaluate competing ideas to decide which action is the best to take for all citizens. Craig's assessment does not emphasize these qualities. Analysis of the mean score (6.59 out of 10.0) of all assessments in this study indicated that they fell somewhere in between Riley's and

Craig's assessments, engaging pupils in some aspects of the skills necessary for democratic participation.

The previous discussion illustrates that the authentic intellectual quality of assessment tasks that teachers used in their classrooms varied. Although the following chapter presents an analysis using the qualitative internal TAPL data to examine the contexts and conditions of these variations, there are variables that occur across all of these assessment tasks that begin to shed light onto which conditions and contexts influence the ability of beginning teachers to implement authentic intellectual work. This section of Chapter 4 describes differences in scores according to the following variables: 1) academic discipline, 2) grade level, 3) school context, and 4) assessment creation.

**Academic Discipline**

The assessment tasks submitted for this study included all four of the major academic disciplines included in the rubric: writing, science, math, and social studies. The mean authentic intellectual work scores for assessment tasks in different subjects differed from one another (see Table 4.7), which suggests that academic discipline may be related to the quality of authentic intellectual work.

Table 4.7

*Assessment Task Mean Score by Discipline*

| Academic Discipline | M | n | SD |
|---|---|---|---|
| Writing[a] | 8.23 | 16 | 1.224 |
| Social Studies[b] | 7.07 | 7 | 1.946 |
| Math[c] | 5.55 | 19 | 1.136 |
| Science[d] | 5.73 | 11 | 1.679 |

[a]An independent samples *t* test showed that scores in Writing were statistically different than scores in Math, *t*(32)=6.578, *p*<.001, and scores in Science, *t*(25)=4.468, *p*<.001.
[b]An independent samples *t* test showed that scores in Social Studies were statistically different than scores in Math, *t*(24)=2.58, *p*=.019.
[c]An independent samples *t* test showed that scores in Math were statistically different than scores in Writing, *t*(32)=6.578, *p*<.001, and scores in Social Studies, *t*(24)=2.58, *p*=.019.
[d]An independent samples *t* test showed that scores in Science were statistically different than scores in Writing, *t*(25)=4.468, *p*<.001.

As Table 4.7 indicates, assessments in writing were more likely to engage pupils in authentic intellectual work. The mean authentic intellectual work score for writing assessments was 8.23 (*n*=16, *SD*=1.224), which was statistically different (*p*<.05) from the means for math and science assessments. Teachers who used writing assessments provided more opportunities for pupils to critically analyze information related to important topics. One possible reason that writing assessments scored high is that they frequently required pupils to write extensively, allowing those assessments to earn high scores on the standard elaborated written communication. For example, Amanda and Elizabeth both submitted assessments that required their secondary English and humanities pupils to write essays and use evidence to support their arguments. These writing assessments also provided opportunities to relate material to pupils' lives such as Riley's poetry anthology described previously or Amanda's assessment for the novel, *A*

*House on Mango Street*, where pupils were asked to create personal vignettes and prose poems based on the forms in that work. Here pupils were asked to critically reflect and make judgments in writing.

Similar to the writing assessments, social studies assessments were also more likely to engage pupils in authentic intellectual work than assessments in math and science. The mean score for authentic intellectual work on social studies assessments was 7.07 ($n$=7, $SD$=1.946), which was lower than the mean for writing assessments but higher than mean scores for math and science. However, only the math assessments, and not the science assessments, were lower than the social studies assessments in a statistically significant way ($p$<.05). Many of the social studies assessments also asked pupils to defend arguments in writing, although a number of assessments incorporated multiple choice or short answer exercises as well. For example, the social studies exam that Mark submitted during his second year of teaching included 30 multiple choice questions and a choice of two out of three short answer essay questions. Some of the assessments, such as Mark's social studies exam or Mara's assessment asking pupils to create a PowerPoint presentation about World War I, did not provide opportunities for pupils to connect the topic to their own lives, although there were a few instances where social studies assessments made greater connections. The strongest example from the discipline of social studies is Mara's social action assessment. In this assignment, Mara asked pupils to research an important current event, write an analytical narrative about the topic, and relate the topic to a social action project that pupils designed in an effort to make a difference in their community. This example shows how social studies assessments could

promote civic engagement and responsibility, which are also critical to democratic education.

In contrast, assessments in math and science were less likely to engage pupils in authentic intellectual work, as defined here. The mean for math assessments was 5.55 ($n$=19, $SD$=1.136) and the mean for science assessments was 5.73 ($n$=11, $SD$=1.679). There were no statistical differences between the means for science and math assessments but they were both lower, in a statistically significant way ($p$<.05), than the mean for writing assessments. The majority of the math and science assessments involved short answer and multiple choice questions, providing few opportunities for pupils to conceptualize new ideas or defend their work. Some of the math assessments did ask pupils to elaborate on their answers. Rachel's elementary math worksheets, for example, asked pupils to show their work and include diagrams, and Matt offered a few opportunities for pupils to create graphs and figures in his secondary geometry assignments. These features raised the authentic intellectual work scores on these math assessments. This, however, was not the norm. Only one math assessment – a secondary portfolio assignment where students were asked to write paragraphs about their experiences in math class – called for extensive writing and, as Matt discussed in his interview about the portfolio, he discontinued the assignment. Similarly, only two science assessments involved hands-on laboratory work. In fact, some of Craig's middle school science classes were not even held in a lab classroom. Most of the science assessments asked pupils to recall information.

The differences in the quality of authentic intellectual work for various academic disciplines found in this study are consistent with the findings of other studies (Newmann, Lopez et al., 1998; Newmann et al., 1996). These findings suggest that writing and social studies teachers are more likely to engage their pupils in authentic intellectual work than are math and science teachers. Teachers are more likely to provide authentic intellectual work in the disciplines of writing and social studies. The beginning math and science teachers in this study were less likely to implement authentic intellectual work. This is not to suggest that there are no opportunities for assignments in math and science to provide opportunities for authentic intellectual work. On the contrary, critical thinking and analysis are instrumental in the academic disciplines of science and math. However, the math and science teachers in this study did not emphasize these skills.

Although it is evident from the external TAPL analysis that assessments in writing and social studies engaged pupils in authentic intellectual work, there were other factors that may have influenced the scores outside of the nature of the academic discipline. One consideration is the structure of the rubric. The standard elaborated written communication seems to privilege disciplines like writing and social studies that traditionally involve extensive writing. Even though the rubric allows for graphs, tables, and equations as a source of elaborated work in math and science, assessments in those disciplines, particularly standardized ones, are less likely to ask pupils to demonstrate their understandings in that manner. Many of the math and science assessments in this study were standardized assessments that the teachers in this study themselves did not

create. As the following section demonstrates, standardized assessments did not typically provide much authentic intellectual work. Another explanation may have to do with individual teachers in this study. This is a small sample. All of the science assessments were from two teachers, and there is the possibility that the quality of work is related more to the individual teacher than the discipline. These details are further explored in the qualitative analysis in Chapter 5.

**Grade Level**

One of the variables that did not seem to influence scores for authentic intellectual work on assessment tasks was grade level. There were slight differences in the means for secondary, middle, and elementary assessments. As Table 4.8 illustrates, secondary assessments had the highest scores, followed by elementary and then middle school assessments, although the differences were not statistically significant. Only secondary assessments were statistically higher than middle school assessments at *p*<.05. Since the middle school assessments came from two science teachers, it is possible that the difference can be better explained by the effect of academic discipline or teacher. Those factors are further explored in Chapter 5. More importantly however, is the fact that elementary and secondary teachers in this study had similar expectations for their pupils in terms of engaging in authentic work. The same degree of opportunities was available for younger and older pupils. This suggests that both elementary and secondary teachers perceived that their pupils had the capacity to engage in developmentally appropriate authentic intellectual work. They expected young students to engage in complex thinking as much as they expected older students to do so.

Table 4.8

*Assessment Task Mean Score by Grade Level*

| Grade Level | M | n | SD |
|---|---|---|---|
| Secondary[a] | 7.10 | 25 | 1.780 |
| Middle | 5.33 | 9 | 1.732 |
| Elementary | 6.53 | 19 | 1.620 |

[a]An independent samples *t* test showed that the difference between Secondary and Middle school assessments was statistically significant, $t(32)=2.571$, $p=.015$.

## School Context

The type of school where teachers taught did not seem to influence the level of authentic intellectual work teachers provided for their pupils. As Table 4.9 indicates, the majority of the assessment tasks came from teachers in urban schools. There were 42 assessments from urban schools and only 11 assessments from suburban schools. The mean for assessment tasks in urban school was 6.46 (*SD*=1.744) and the mean for assessment tasks in suburban schools was 7.09 (*SD*=1.998), although the difference was not statistically significant. This suggests that the level of authentic intellectual work provided by the teachers in this study was comparable for teachers across schools and pupil populations. These beginning teachers had the same expectations for their pupils to engage in democratic thinking, regardless of whether or not they were in an urban or suburban context. Although further research is needed to support this, especially based on the limited number of assessments from suburban schools in this study, it appears that teachers across schools that varied in terms of pupils' socioeconomic status and location in urban or suburban areas provided similar levels of authentic intellectual work. This is worth further investigation considering that Newmann, Bryk, and Nagaoka (2001) and

181

Newmann, Marks, and Gamoran (1996) concluded that students' socioeconomic status had no bearing on whether or not teachers provided authentic intellectual learning opportunities. To better understand the relationship between school context and authentic intellectual work, Chapter 5 examines school context through the internal TAPL data.

Table 4.9

*Assessment Task Mean Score by School Context*

| School Context | *M* | *n* | *SD* |
|---|---|---|---|
| Urban | 6.46 | 42 | 1.744 |
| Suburban | 7.09 | 11 | 1.998 |

**Assessment Creation**

One variable that did have an impact on the quality of authentic intellectual work was related to the creation of the assessment. As previously mentioned, participants were only asked to submit a culminating assessment that they had used in their classroom. As a result, teachers submitted assessments that they created completely on their own, assessments that they modified or co-constructed, and assessments that they had no role in creating. As Table 4.10 illustrates, there were differences in the scores depending upon who created the assessment.

Table 4.10

*Assessment Task Mean Score by Assessment Creation*

| Assessment Creation | M | n | SD |
|---|---|---|---|
| Participant-created[a] | 7.31 | 13 | 1.521 |
| Modified/Co-constructed | 6.74 | 23 | 2.163 |
| Other-created | 5.85 | 17 | 1.142 |

[a]An independent samples *t* test showed that the difference between assessment task scores on Participant-created assessments was statistically different than assessment task scores on Other-created assessments, $t(28)=2.625$, $p=.014$.

The assessment tasks that the teachers in this study created on their own received a higher score for authentic intellectual work than assessment tasks that teachers did not create; this difference was statistically significant ($p<.05$). As Table 4.10 shows, the assessment tasks that participants created had the highest mean ($M=7.31$), followed by assessments that they modified ($M=6.74$), followed by assessments that they did not create ($M=5.85$), which had the lowest mean. When participants developed their own assessments, they were more likely to create assessments that had higher levels of authentic intellectual work. For example, Mara created an essay assessment for a unit on genocide for her secondary social studies students, which emphasized critical thinking and allowed her pupils to take a stand on a significant global issue. Conversely, Riley's district-created math exam, which she had no role in creating, was a multiple choice exam where pupils routinely applied math formulas to answer the questions.

This finding suggests several things. First, some of these beginning teachers were able to create authentic assessments that critically engaged their pupils with meaningful material. They considered authentic intellectual work when they created assessments.

Beginning teachers are capable of creating assessments that are more authentic than presumably experienced educators or curriculum developers who constructed the standardized and pre-packaged curriculum. These teachers, who were prepared in a teacher education program that reinforced these skills, were capable of constructing authentic opportunities. In addition, when these beginning teachers created their own assessments, they were more likely to attend to authentic intellectual work than pre-packaged curriculum and standardized assessments.

Based on the external TAPL analysis of assessment tasks, teachers prepared in a teacher education program emphasizing aspects of authentic intellectual work were able to attend to some aspects of democratic education in their practice. They were not able, however, to continually address significant skills at all times, yet their work is consistent with the practice of other veteran teachers, which suggests that these teachers are doing as well as, if not better, than teachers in general in terms of this type of work. To see how their efforts influenced their pupils, this discussion now turns to an examination of pupil learning.

### Pupil Learning and the Quality of Authentic Intellectual Work

My analysis of assessment tasks illustrates that many of the beginning teachers in this study were able to attend to some aspects of authentic intellectual work in their practice. The next important issue is how students responded to these opportunities. To what extent were pupils able to demonstrate authentic intellectual work? This section examines pupil performance on the assessment tasks that teachers submitted, referred to here as *pupil work*.

As part of the TAPL protocol, teachers were asked to submit a class set of pupil work for each assessment task they selected for the five TAPL interviews they participated in. Teachers submitted a class set of pupil work for one assessment during their student teaching, a class set of pupil work for two assessments in their first year of teaching, and a class set of pupil work for two assessments in their second year of teaching. From each class set, 10 samples of pupil work for each assessment task were randomly selected as part of this analysis; in cases where the class set contained fewer than 10 items, all of the samples were scored. A total of 481 samples of pupil work that were produced in response to the 53 assessment tasks that were submitted were examined. This study did not track individual pupils over time because teachers did not always submit pupil work samples from the same students nor did they have the same pupils throughout the 3-year period. However, pupils were assigned identification codes for the larger QCS project and it was sometimes possible to match individual pupils' work to qualitative data such as instances where pupils were mentioned in an interview or classroom observation.

As the previous discussion of assessment tasks illustrated, the nature of assessment tasks varied in terms of grade, academic discipline, and format. Therefore, pupil work samples also varied in type, including written essays, completed worksheets, posters, PowerPoint presentations, short answer responses, and multiple choice or fill-in-the-blank answers. Pupils worked individually and in groups, depending upon the parameters of the specific assessment. For example, in response to Mara's assessment on Renaissance art, pupils worked in groups to create PowerPoint presentations; the

PowerPoint files were submitted and researchers evaluated the text and images that pupils used. On Mark's district-created social studies exam, which consisted of multiple choice and essay questions, the pupil work included answers to the multiple choice as well as the written essays. In some cases, pupils did not complete the entire assignment.

For the external TAPL analysis, pupil work was evaluated in the same manner as the assessment tasks. That is, researchers scored the pupil work samples according to the a rubric for authentic intellectual work (RISER, 2001). According to the RISER rubric, the categories and scoring criteria were different for pupil work than the guidelines for assessment tasks. For pupil work, there were three standards for authentic intellectual work: 1) analysis, 2) disciplinary concepts, and 3) elaborated written communication.[7] To illustrate how the rubric described the standards, Table 4.11 presents the definition for each standard for pupil work in math. The descriptions of the standards in the other disciplines have the same underlying concepts but use discipline-specific terminology. As can be seen, mathematical analysis focuses on how pupils think about math content, mathematical concepts focuses on how pupils understand and make connections to fundamental math concepts, and elaborated written communication focuses on how pupils are able to articulate their understandings in the discipline.

---

[7] The rubric standards for pupil work in the discipline "writing" have different labels than the terms used for the other three disciplines. However, since the concepts behind the terms are consistent with the other standards, they have been grouped together for the purpose of clarity.

Table 4.11

*Authentic Intellectual Work Standards for Pupil Work in Mathematics (RISER, 2001)*

| Standard | Rubric Description |
|---|---|
| Standard 1: Mathematical Analysis | Student performance demonstrates thinking about mathematical content by using mathematical analysis; that is, going beyond simple recording or reproducing of facts, rules, and definitions or mechanically applying algorithms. |
| Standard 2: Mathematical Concepts | Student performance demonstrates understanding of important mathematical concepts central to the assignments; for example, by representing concepts in different concepts in different contexts, or making connections to other mathematical concepts, other disciplines, or real world situations. |
| Standard 3: Elaborated Written Communication | The student's performance demonstrates elaboration of his or her understanding, explanations, or conclusions through extended writing; for example, through diagrams, symbolic representations, or prose that presents convincing arguments. |

*Note*. Reproduced from *Standards and Scoring Criteria for Assessment Tasks and Student Performance*, by the Research Institute on Secondary Education Reform, 2001, Madison, WI: Author.

Each standard has a list of scoring criteria used to rate the pupil work. Table 4.12 provides the scoring criteria for the standard mathematical analysis for pupil work in math. As Table 4.12 indicates, the standard is evaluated on a 4-point scale. The lowest

score (1.0) indicates that pupils did not show any mathematical analysis while the highest score (4.0) indicates that pupils demonstrated mathematical analysis throughout their assessment. The scores for the individual standards are added together for one total authentic intellectual pupil work score. Therefore, pupil work scores could range from 3.0 to 12.0, with 3.0 indicating the lowest level of authentic intellectual work and 12.0 indicating the highest level. All of the pupil work samples in this study were scored according to the rubric in the external TAPL analysis.

Table 4.12

*Scoring Criteria for Standard 1:* Mathematical Analysis *(RISER, 2001)*

| | Mathematical Analysis |
|---|---|
| Scale | Description |
| 4 | Mathematical analysis was involved throughout the student's work |
| 3 | Mathematical analysis was involved in a significant proportion of the student's work. |
| 2 | Mathematical analysis was involved in some portion of the student's work. |
| 1 | Mathematical analysis constituted no part of the student's work. |

*Note*. Reproduced from *Standards and Scoring Criteria for Assessment Tasks and Student Performance*, by the Research Institute on Secondary Education Reform, 2001, Madison, WI: Author.

My analysis of the pupil work scores suggests that pupils in this study produced some authentic intellectual work in response to the assignments and assessments teachers used. However, as was the case with the assessment tasks, the authentic quality of

students' work was not at the highest end of the authentic intellectual work rubric. The mean score for all of the pupil work in this study was 6.28 ($n$=481, $SD$=2.083), on a scale that could range from 3.0 to 12.0. The mean falls below the mid-point on the scale which suggests that pupils produced moderately low levels of authentic intellectual work, as defined in this study. Pupil work scores for individual samples of pupil work ranged from 3.0 to 12.0, indicating that some pupils demonstrated very high levels of authentic work, where they showed mastery of concepts and analysis, while other pupils produced very inauthentic work and had little understanding of the material.

In order to examine how pupils performed on individual assessments, means were calculated for the pupil work scored for each assessment task, which is referred to here as the "class mean." Table 4.13 presents the class mean on pupils' authentic intellectual work scores for each teachers' assessment tasks submitted at the five different TAPL interviews during the preservice period and first two years of teaching. As seen in Table 4.13, the class means for pupil work ranged from 4.35 to 8.75. There is greater range and variation in pupil performance than there was in the assessment tasks and fewer general trends. Some teachers, such as Elizabeth and Mara, had pupils engage in higher levels of authentic work at the end of their second year of teaching than they did in their student teaching, suggesting that they were able to increase pupil performance as they developed as teachers. Rachel, Riley, and Sonia, however, encountered the opposite experience as pupil performance decreased over time. Other teachers, such as Craig, Amanda, and Sylvie, had similar pupil performance over the 3-year period.

Table 4.13

*Pupil Work External TAPL Scores*

| | | | Pupil Work External TAPL Scores (Class Means) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Student | Year 1 | Year 1 | Year 2 | Year 2 | Overall | | |
| Participant | Teaching | Fall | Spring | Fall | Spring | *M* | *SD* | *N* |
| Amanda | 7.85 | 6.30 | 8.35 | 8.50 | 7.10 | 7.62 | 1.842 | 50 |
| Craig | 4.75 | 3.10 | 5.70 | 5.10 | 4.45 | 4.61 | 1.914 | 48 |
| Elizabeth | 6.20 | 6.71 | 8.29 | 8.30 | 7.35 | 7.35 | 2.134 | 44 |
| Lola | 6.05 | 5.83 | 6.00 | 8.75 | 6.25 | 6.50 | 2.298 | 47 |
| Mara | 7.40 | 7.20 | 8.50 | 6.85 | 9.00 | 7.77 | 1.771 | 49 |
| Mark | 5.30 | 6.45 | 5.40 | 5.00 | No Pupil Work Samples[†] | 5.57 | 1.669 | 38 |
| Matt | 7.40 | 6.00 | 5.10 | 7.35 | 4.63 | 6.35 | 2.003 | 37 |
| Rachel | 7.11 | 5.00 | No Pupil Work Samples[*] | 5.35 | 5.90 | 5.81 | 1.490 | 39 |
| Riley | 7.33 | 4.90 | 6.40 | 4.35 | 5.60 | 5.45 | 1.577 | 43 |
| Sonia | 8.00 | 4.70 | 5.80 | 5.50 | 5.80 | 5.96 | 1.987 | 50 |
| Sylvie | 6.20 | 5.06 | 5.25 | 6.15 | 6.00 | 5.69 | 1.375 | 36 |
| *M* | 6.71 | 5.51 | 6.45 | 6.46 | 6.30 | **6.28** | 2.083 | 481 |
| *N* | 95 | 97 | 95 | 106 | 88 | | | |
| *SD* | 2.149 | 1.822 | 2.164 | 2.086 | 2.010 | | | |

[†]Mark only participated in four TAPL interviews because he was not employed as a teacher for a portion of this study.

[*]Rachel's assessment task for Interview 9 was not collected.

Teachers were placed on a continuum of authentic intellectual work based on the quality of their pupils' work scores (see Figure 4.4). Similar to the continuum based on assessment tasks, teachers fell into three groups: low, medium, and high levels of authentic intellectual work. In all but two cases, teachers were in the same groups for assessment task scores and pupil work scores. Craig was at the low end for both continua; Lola, Rachel, Riley, Sonia, and Mark were in the medium range; and Mara, Elizabeth,

and Amanda were at the high end. In these cases, pupils produced work that was similar in quality to the assessment tasks that they were given. This suggests that the quality of the assessments that teachers used influenced the quality of pupils' work. To explore this relationship further, a correlation analysis between the assessment task scores and the pupil work scores is presented at the end of this chapter.

## Authentic Intellectual Quality of Pupil Work
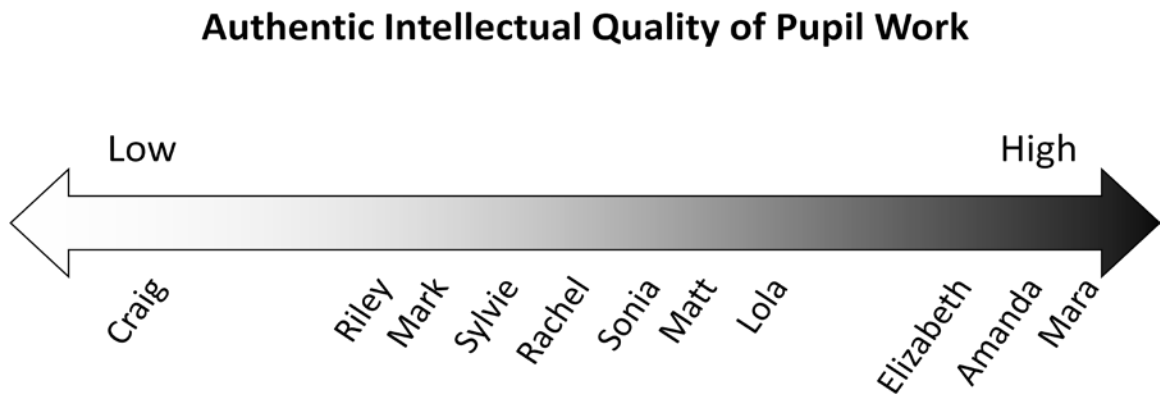


*Figure 4.4.* Continuum of the authentic intellectual quality of teachers' pupil work scores.

Two cases, however, did not fit this pattern. Matt had an assessment task mean that placed him in the low authentic work group for assessment tasks, but his pupil work scores fell within the medium range of the continuum. His pupils performed higher than the assessment tasks that they were given. Although the contextual factors that may have led to this are described in greater detail in Chapter 5, one explanation for this discrepancy is related to the academic level of Matt's students. Matt taught high achieving pupils at an exam school where students were placed into the school based on their academic achievement. It is probable that his pupils were able to produce higher

levels of authentic intellectual work despite the quality of the assessments because they were very high-achieving students. For example, pupils wrote sentences to explain their work at times where they were not asked to do so (see Chapter 5).

The second case was Sylvie, whose assessment task scores placed her into the high category and yet, her pupil work scores fell within the medium range. The population of Sylvie's students may also explain this change. Sylvie taught in a resource room for English Language Learners. According to the RISER rubric, the language ability of students is not taken into account when assigning a score to pupil work samples. The rubric explicitly states: "Scores should assess the quality of the actual written work and not take into consideration possible effects of a student's linguistic background or learning disability" (RISER, 2001, p. 7). However, it is reasonable to assume that her pupils scored lower on their work due, in part, to their language abilities. For instance, Figure 4.5 is a writing assessment completed by one of Sylvie's pupils. Based on the RISER rubric, this pupil work sample would be evaluated as if the piece were written by a fourth grader who is not an English Language Learner. On the standard for writing forms and conventions, this sample would receive a low score due to grammatical errors and sentence structure. The score would not reflect the fact that this particular pupil had only been in the United States for a very brief period of time and was learning a new language. Therefore, some of Sylvie's pupil work scores may not accurately reflect her pupils' understandings. The ELL pupils in Sylvie's class represent an example of how classroom context, here in the form of pupil population, informed the

external TAPL results. The fact that all of the pupils in Sylvie's class are English

Language Learners should be considered when interpreting her pupils' scores.
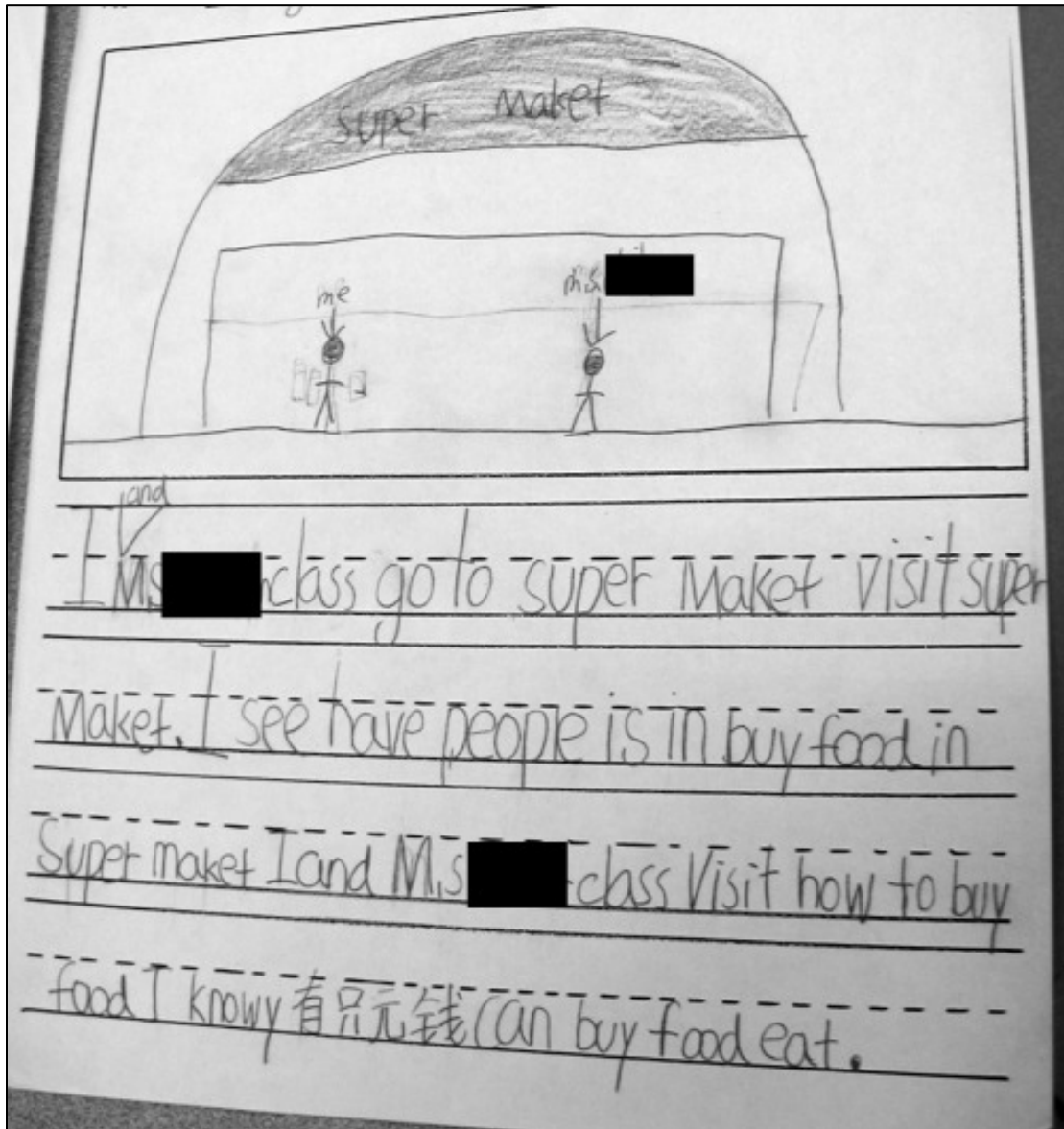


*Figure 4.5*. Writing example from an ELL pupil in Sylvie's class.

As with the analysis of assessment tasks, it is useful to investigate whether the

authentic intellectual quality of pupils' work changed during the 3-year period. A

repeated-measures ANOVA was conducted on the pupil work scores to determine whether there were statistical differences in the authentic intellectual work scores for pupil work during the five different time-points (Student Teaching, Year 1 Fall, Year 1 Spring, Year 2 Fall, and Year 2 Spring). This analysis was used to detect if there was a general trend in the quality of authentic intellectual work for pupil work scores for all of the teachers during teachers' student teaching period through the end of their second year of teaching.

For this repeated-measures ANOVA, teachers were the random factor and the class mean for pupil work scores for each time-point was the fixed repeated measures factor. Since Rachel and Mark only had scores for four out of the five time-points, missing values were replaced with the mean for all of their pupils' work scores on the 4 assessments. In this analysis, the null hypothesis is $H_0$: $\mu_1 = \mu_2 = \ldots = \mu_k$, meaning that there were no differences in pupil work scores for each of the 5 time-points. The alternative hypothesis, therefore, is $H_a$: $\mu_i \neq \mu_i$ for some $i, k$, meaning that there is a difference in scores between two of the time-points. The probability level was set at $\alpha = .05$. Independence of observations and sphericity were met. Results indicate that the assessment task scores were not statistically different over the five time-points $F(4, 40) = 2.282$, $p = .077$, eta$^2 = .186$ (see Table 4.14) and the null hypothesis was not rejected. The *a priori* power analysis, described in Chapter 3, suggested that a sample size of 21 was needed in order to achieve a medium effect size (.25) and power of .80. As with the previous RMANOVA on teachers' assessment task, since this repeated measures ANOVA had a sample size of 11, it is possible that the non-significance was a result of

the small sample size. This analysis should be replicated with a larger sample to test this further.

Table 4.14

*Repeated Measures ANOVA Tests of Within-Subjects Effects on Pupil Work (Class Mean) Authentic Intellectual Work Scores*

|  | Type III Sum of Squares | *Df* | Mean Square | *F* | Sig. |
|---|---|---|---|---|---|
| Time-points | 8.128 | 4 | 2.032 | 2.282 | .077 |
| Error (Time-points) | 35.621 | 40 | .891 | | |

*Note.* See Table 4.13 for the Pupil Work Class Means used in this analysis.

The results of the repeated-measures ANOVA suggest that there were no overall trends in the pupil work scores. That is to say, the beginning teachers in this study, as a whole, had pupils who, most likely, produced the same levels of authentic intellectual work while teachers were student teachers, first-year teachers, and second-year teachers. When taken as a whole, there were probably no differences in pupil performance over time. Most likely, teachers did not have students whose work increased in quality as they developed into teachers nor did they have students whose work decreased in quality. Since it is reasonable to speculate that teachers might utilize and construct higher quality assessments as they became more experienced teachers, this finding suggests that this is not necessarily the case for all teachers. However, at an individual teacher level, there may be instances where individual teachers encountered change over time which the repeated-measures ANOVA does not account for. Chapter 5 examines contextual factors that may help explain these variations between teachers' scores.

It should be noted here that pupil work scores can be limited by the assessment tasks. Pupils, for the most part, can only produce as much authentic intellectual work as they are asked to do. For example, if pupils were given a multiple choice assessment, their work is in the form of multiple choice answers. In that case, all of the pupils on that assessment would earn the lowest score on the standard elaborated written communication because they are only allowed to provide the answer. They are not given an opportunity to explain or support their answer with additional information. None of the pupils would be able to receive the highest score for authentic intellectual work, even if all of their answers were correct. For example, the district-created math assessment that Riley submitted from the fall semester of her second year of teaching received a 5.0 for the assessment task and had a score of 1.0 on elaborated written communication, the lowest possible score for that criterion. The entire test consisted of multiple choice questions. Pupils, therefore, could only earn a 1.0 on that standard on their pupil work. The highest possible score for pupils on that assessment, therefore, was 9.0 instead of 12.0. The class mean for pupil work on that assessment was 4.35 ($SD$=1.292). In this case, pupils were not able to demonstrate certain skills because they were not asked to do so. However, even though students were limited by the assessment task, they were still not engaging in authentic work as defined by this study. A more thorough examination of the relationship between the quality of the assessment task and the quality of pupil performance occurs in the last section of this chapter.

When compared to other studies of authentic intellectual work, the pupils in this study had scores that were almost as high as other pupils. In their study of authentic

intellectual work, described in the previous discussion of assessment tasks, King, Schroeder, and Chawszczewski (2001) found similar results. The mean for pupil work in one of their studies was 7.21 and the mean for pupil work in a second study was 7.47. Their population included pupils with and without disabilities whose teachers explicitly incorporated authentic intellectual work in their instruction. Although the scores in the King et al. study were higher than this study's mean score of 6.28, the scores still fall within the middle of the rubric scale which ranged from 3.0 to 12.0. Pupils in that study did not produce the highest levels of authentic intellectual work. The pupils of beginning teachers in this study were able to perform about as well as the pupils of veteran teachers. The findings of the King et al. study and the findings of this study are supported by other studies showing that, in general, pupils do not produce high levels of authentic intellectual work (Bryk et al., 2000; Ladwig et al., 2007; Newmann et al., 2001).

In this study, pupils were able to demonstrate some aspects of authentic intellectual work better than others, which can be seen through an analysis of pupil performance on the different characteristics of authentic intellectual work. As Table 4.15 illustrates, pupil work had different mean scores on the three different standards and these differences were statistically significant ($p<.05$). Pupils were better able to demonstrate content knowledge and were less able to engage in higher-order thinking or support their opinions.

Table 4.15

*Mean Scores for Pupil Work by Authentic Intellectual Work Standards*

| Standard | $M$ | $n$ | $SD$ |
|---|---|---|---|
| Analysis[a] | 2.14 | 481 | .782 |
| Disciplinary Concepts[b] | 2.20 | 481 | .759 |
| Elaborated Written Communication[c] | 1.95 | 481 | .833 |

[a]A paired samples *t* test showed that the score for the standard Analysis was statistically different than the scores for the standard Disciplinary Concepts, $t(480)=2.462$, $p=.014$, and the standard Elaborated Written Communication, $t(480)=6.515$, $p<.001$.
[b]A paired samples *t* test showed that the score for the standard Disciplinary Concepts was statistically different than the scores for the standard Analysis, $t(480)=2.462$, $p=.014$, and the standard Elaborated Written Communication, $t(480)=7.834$, $p<.001$.
[c]A paired samples *t* test showed that the score for the standard Elaborated Written Communication was statistically different than the scores for the standard Analysis, $t(480)=6.515$, $p<.001$, and the standard Disciplinary Concepts, $t(480)=7.834$, $p<.001$.

In general across the pupil work in this study, pupil work received the highest score for the standard disciplinary concepts. This standard measures how well pupils understand "the fundamental ideas relevant" to the subject matter (RISER, 2001, p. 27). This primarily focuses on content. The mean for this category was 2.20 (*n*=481, *SD*=.759) (out of a 1.0 to 4.0 scale). The mean is slightly below the mid-point of the scale. In terms of the rubric, this indicates that pupils showed an understanding of some concepts that were significant for the topic "but their use is significantly limited and/or shows significant flaws in their understanding" (p. 27). This indicates that pupils had wrong answers, did not show high levels of understanding about the concepts in the field, and did not exhibit extensive content knowledge. Although the score is low, it is the highest out of the three standards and indicates that teachers were the most successful in helping pupils understand the content matter.

Pupils had slightly lower scores on the standard analysis. This standard measures the extent to which work "goes beyond mechanically retrieving or reproducing fragments of knowledge" (RISER, 2001, p. 20). Here the focus is on whether pupils apply or restate information. The mean for this standard was 2.14 ($n$=481, $SD$=.782), out of a possible range in score from 1.0 to 4.0. Similar to the previous standard, the mean for analysis falls below the mid-point of the rubric scale and indicates that pupils only engaged in analysis "in some portion" of their work (p. 22). Pupils were more likely to demonstrate lower order thinking skills. The moderate score on analysis suggests that pupils are not able to critically evaluate information. This does not foster their ability to engage in rational debate about the common good because pupils are not learning how to make good judgments.

Similarly, pupils were least able to support their thinking in an extended way. Pupil work received the lowest score on the standard elaborated written communication. This standard measures the extent to which pupils demonstrated "an elaborated, coherent account that draws conclusions or makes generalizations or arguments and supports them with examples, summaries, illustrations, details, or reasons" (RISER, 2001, p. 8). The focus of this standard is how well pupils are able to support their arguments and explain their thinking. The mean for this standard was 1.95 ($n$=481, $SD$=.833), out of a possible range of scores of 1.0 to 4.0. Similar to the previous two standards, the mean for elaborated written communication falls below the midpoint of the scale and is even closer to the lowest end of the scale than the other two. A score of 1.95 on the rubric indicates that pupils showed "some evidence of elaboration" but only "a small portion of the

student's work comprises an elaborated, coherent account" (p. 8). Pupils did not engage in extended communication about their ideas. This is most apparent in the multiple choice answers that pupils provided in assessments such as Mark's district social studies exam and Riley and Sonia's math exams. When combined with the scores on the standard, analysis, this suggests that the pupils in this study did not produce high levels of authentic intellectual work. They had few opportunities to extensively interpret information or construct and defend arguments, skills that are necessary for participation in a democratic society.

Although it is likely that pupil performance was limited by the assessment tasks that they were given, it is important to note that for all three authentic intellectual work standards, pupil work fell below the mid-point of the rubric scale. That is, although pupils demonstrated some knowledge of content and concepts, they were less able to utilize and synthesize that material. This lack of higher-order thinking suggests that pupils have not developed the skills needed to evaluate competing ideas and make rational decisions.

To further examine differences in the extent to which pupils demonstrated authentic knowledge, it is useful to examine two different pupil work samples from the same assessment. Figure 4.6 shows an excerpt of two pupils' answers to Sonia's 4th grade math assessment, which was part of a pre-packaged curriculum that she was mandated by her district to use in her classroom. Only a portion of the pupils' work on this assessment is shown. This assessment task received a score of 5.5[8], out of a possible 10.0, on the external TAPL analysis. For this portion of the assessment, pupils were asked to describe

---

[8] This assessment task was scored by two raters. The two raters disagreed on the scoring and their scores were averaged to obtain the total assessment task score.

and compare different geometric figures. They were instructed to show their work, using "drawing, tables, and/or computations" and explain their answers. Comparing the responses of two pupils elaborates how scoring was conducted.



*Figure 4.6.* Example of high and low scoring pupil work samples. Sonia's 4th grade math test about polygons. The assessment task scored a 5.5 (out of 10.0).

The two pupils performed very differently on this assessment. Pupil A, whose work is shown on the left hand side of Figure 4.6, received a score of 5.0, out of a possible range of scores from 3.0 to 12.0, on this assessment and is representative of a

low score. Pupil B, whose work is on the right hand side of the figure, received a score of 9.0 and represents a high score on this assessment. As Figure 4.6 illustrates, these two pupils received different scores on each standard. (Each standard is based on a scale of 1.0 to 4.0.)

On the standard mathematical analysis, which measures a pupil's ability to go beyond reproduction of information, Pupil A received a score of 2.0 and Pupil B received a score of 3.0. Pupil A exhibited "some" analysis by interpreting the similarities between the shapes in the two questions but did not enumerate or explain the difference between the two shapes, which was part of the question. Pupil A did not complete the question. Pupil B, on the other hand, compared and contrasted the different shapes. This is a higher level of analysis than Pupil A because it involved making a comparison and contrasting the differences between the shapes.

On the standard disciplinary concepts, which measures pupil understanding of the relevant mathematical concepts, Pupil A and Pupil B exhibited different levels of understanding. Pupil A demonstrated "some" understanding of the concepts by stating that triangles have 180° and that quadrilaterals have 360°; he received a score of 2.0 on this standard. In contrast, Pupil B was able to name the different polygons and used the terms "polygon," "quadrilaterals," and "trapezoid" to describe the shapes, demonstrating an understanding of the different terms. In addition, Pupil B also noted the characteristics of a right triangle. Although Pupil B incorrectly labeled the triangle as an isosceles triangle, the majority of the work (as well as the remainder of the work not shown here)

was correct. Pupil B scored a 3.0 on this standard. Here, Pupil B was able to show that he knew multiple terms and understood what they meant.

On the third standard, elaborated written communication, which measures how pupils articulate their understanding of the material, it is fairly obvious by looking at the two samples of work that Pupil A and Pupil B had differences in their ability to elaborate their thinking. The question on this assessment task specifically asks pupils to show their work using "drawing, tables, and/or computations" and asks pupils to explain their answers. This is an example of how a math assessment could score high on elaborated written communication. Pupil A received a score of 1.0, the lowest possible score, on this standard. For one thing, his answer is incomplete because he only wrote about similarities and not differences as the question asked. In addition, it is not clear from his writing which part of the question he is answering. Pupil B, on the other hand, scored a 3.0 on this standard. He used multiple sentences to explain his answer and labeled the diagrams to further clarify his point. He enumerated the similarities and differences between the shapes (i.e. why "they are the same" and why "they are different") and used mathematical terms to articulate his understandings of the concepts (i.e. "quadrilateral," "right triangle," etc.). In addition, Pupil B went beyond the task by using space on the back of his worksheet to continue his answer to the question. Through elaboration, Pupil B is able to show what he does and does not know about the material.

These two examples of pupil work illustrate the range of pupils' abilities to engage in higher-order thinking. Pupil B clearly articulated and supported his answers whereas Pupil A did not. Pupil B demonstrated authentic intellectual work by exhibiting

such skills as comparing and contrasting and used his knowledge of the content to analyze a situation. Through this type of thinking, Pupil B is developing the ability to analyze information in a systematic and informed way.

This discussion of pupil work illustrates that the authentic intellectual quality of pupil work varied by assessment task and also by pupil. Although Chapter 5 presents an analysis using the qualitative internal TAPL data to examine the contexts and conditions of these variations, it is important to note that there are variables that occur across all of the pupil work samples that begin to shed light onto which conditions and contexts influence pupils' abilities to demonstrate authentic intellectual work. The following section of this chapter describes the range and variation of scores according to the following variables: 1) academic discipline, 2) grade level, 3) school context, and 4) assessment creation.

**Academic Discipline**

Similar to the analysis of assessment tasks, one variable that seems to have a relationship to the quality of pupil's authentic intellectual work is academic discipline. As Table 4.16 shows, pupil work scores were higher in certain academic subjects. The means for pupil work in writing and social studies were statistically higher ($p<.05$) than the means for pupil work in math and science. (There were, however, no statistical differences between the means for writing and social studies or between the means for math and science.) Pupils in this study produced higher levels of authentic intellectual work in social studies and writing than they did in math and science.

Table 4.16

*Pupil Work Mean Score by Discipline*

| Academic Discipline | M | n | SD |
|---|---|---|---|
| Writing[a] | 7.0 | 133 | 1.976 |
| Social Studies[b] | 7.1 | 67 | 2.089 |
| Math[c] | 5.9 | 105 | 2.161 |
| Science[d] | 5.6 | 105 | 2.161 |

[a]An independent samples *t* test showed that scores in Writing were statistically different than scores in Science, $t(236)=5.327$, $p<.001$, and scores in Math, $t(307)=5.194$, $p<.001$.
[b]An independent samples *t* test showed that scores in Social Studies were statistically different than scores in Math, $t(170)=4.484$, $p<.001$, and Science, $t(170)=4.184$, $p<.001$.
[c]An independent samples *t* test showed that scores in Math were statistically different than scores in Writing, $t(307)=5.194$, $p<.001$, and scores in Social Studies, $t(170)=4.484$, $p<.001$.
[d]An independent samples *t* test showed that scores in Science were statistically different than scores in Writing, $t(236)=5.327$, $p<.001$, and Social Studies, $t(170)=4.184$, $p<.001$.

Pupils were more likely to analyze information and elaborate their deep understanding of content in social studies and writing. In contrast, pupil performance in math and science assessments involved primarily recalling information without applying knowledge to new situations. Since assessment tasks in writing and social studies also scored significantly higher than assessment tasks in science and math, the similar trends in pupil work scores suggest that assessment tasks that had higher expectations for authentic intellectual work enabled pupils to produce higher levels of authentic intellectual work. This relationship will be examined in the next section of this chapter.

**Grade Level**

Another variable that had a relationship to the quality of pupils' authentic intellectual work is grade level. Table 4.17 lists the means for pupil work by grade level.

As can be seen, secondary pupils produced more authentic intellectual work than elementary and middle school pupils in a statistically significant way ($p<.05$).

Table 4.17

*Pupil Work Mean Score by Grade Level*

| Grade Level | M | n | SD |
|---|---|---|---|
| Secondary[a] | 6.9 | 226 | 2.062 |
| Middle | 5.7 | 77 | 2.349 |
| Elementary | 5.8 | 178 | 1.711 |

[a]An independent samples *t* test showed that the score for Secondary was statistically different than the scores for Middle, $t(301)=4.891$, $p<.001$, and Elementary, $t(402)=6.194$, $p<.001$.

One may speculate that the differences in the quality of pupil work by grade level is due to developmental factors where younger pupils may not be able to engage in the same level of cognitive complexity as older pupils. This is supported by the finding, as discussed in the assessment task section, that secondary and elementary teachers had the same expectation of authentic intellectual work for their pupils. If they provided the same types of opportunities, it is reasonable to expect that pupils would respond in the same manner. However, one factor that may have more influence on this divide is academic discipline. As the previous section demonstrates, the level of pupils' authentic intellectual work may be influenced by subject matter. Pupils performed higher in social studies and writing than they did in math and science. The majority of secondary pupil work was in writing and social studies, while the elementary pupil work consisted primarily in math and the middle school pupil work was only in science. Therefore, pupils may perform

differently because they are able to demonstrate their knowledge in particular subjects. These issues will be examined in greater depth in the following chapter.

**School Context**

Another variable that was related to the quality of authentic intellectual work is school context. Unlike the scores on the assessment tasks, there was a statistical difference (p<.05) between pupils' performance in urban and suburban schools, as indicated in Table 4.18. The mean for pupil work for pupils in urban schools was 6.19 ($n$=384, $SD$=2.089) whereas the mean for pupil work for pupils in suburban schools was 6.66 ($n$=97, $SD$=2.024). That is to say, pupils in suburban schools were more likely to produce higher levels of authentic intellectual work than pupils in urban schools. This study had a disproportionate number of pupil work samples from the urban context than the suburban context, which means this finding is not generalizable to the greater population, but considering that the difference would most likely be greater, in favor of suburban schools, if there had been more assessments from the suburban context, this raises concerns about the level of authentic intellectual work that pupils in urban schools are producing. The analysis in Chapter 5 examines this in greater detail.

Table 4.18

*Pupil Work Mean Score by School Context*

| School Context | M | n | SD |
|---|---|---|---|
| Urban[a] | 6.19 | 384 | 2.089 |
| Suburban | 6.66 | 97 | 2.024 |

[a]An independent samples *t* test showed that the score for Urban was statistically different than the score for Suburban, $t(479)$=1.996, $p$=0.46.

**Assessment Creation**

Finally, another variable that was connected to pupil performance is the creation of the assessment task. As described previously, some of the assessment tasks that teachers submitted were assessments that they themselves did not create. My analysis of assessment tasks illustrated that the assessments that teachers created themselves had higher scores for authentic intellectual work than the assessments that they did not create. A similar trend was found in the analysis of pupil work. As Table 4.19 illustrates, pupils produced different levels of authentic intellectual work depending upon who created the assessment task. Pupils produced the most authentic work on the assessments that teachers in this study created completely on their own. The mean for pupil work on those assessments was 6.50 ($n$=113, $SD$=2.037). Pupil work on assessment tasks that teachers co-constructed or modified had a mean of 6.39 ($n$=203, $SD$=2.234) while pupil work on the assessments that teachers in this study did not create was 6.00 ($n$=165, $SD$=1.895). There was a statistical difference ($p$<.05) between the mean for pupil work on the assessments that teachers in this study created and the mean for pupil work on the assessments that teachers did not create.

Table 4.19

*Pupil Work Mean Score by Assessment Creation*

| Assessment Creation | M | n | SD |
|---|---|---|---|
| Participant-created[a] | 6.50 | 113 | 2.037 |
| Modified/Co-constructed | 6.39 | 203 | 2.234 |
| Other-created | 6.00 | 165 | 1.895 |

[a]An independent samples *t* test showed that the score for Participant-created was statistically different than the score for Other-created, $t(267)$=2.096, $p$=.037.

Pupils performed higher levels of authentic intellectual work on the assessments that the teachers in this study created on their own than the quality of pupils' work on assessments that the teachers in this study did not create. This suggests that these teachers were able to impact pupil learning in a more authentic way when given the opportunity to make their own decisions about assessment practice. This also suggests that teachers who construct their own assessments are more likely to influence pupil learning outcomes in a more authentic way than standardized assessments. Pupils are more likely to demonstrate higher-order thinking and the construction of knowledge on non-standardized assessments.

Pupils in this study were able to demonstrate aspects of authentic intellectual work at a moderate level consistent with the work of pupils of veteran teachers. These beginning teachers did engage pupils in some level of critical thinking, although the extent to which this occurred varied. To further examine how teachers influenced pupil learning, this discussion turns to the relationship between learning opportunities and pupil learning.

### Relationship Between Learning Opportunities and Pupil Learning

The previous discussion of the assessment task and pupil work scores suggests there may be a relationship between the quality of learning opportunities and the quality of the pupil responses to those learning opportunities. For example, with regard to academic discipline, pupil work scores in social studies and writing had a higher mean than pupil work scores in math and science. The same was true for the assessment tasks; assessment task means in social studies and writing were higher than the assessment task

means in math and science. This suggests that pupils are more likely to produce higher levels of authentic intellectual work when they are given more authentic assessments.

To explore this relationship, the assessment task score ($n$=53, $M$=6.59, $SD$=1.798) and the mean, referred to previously as the "class mean," for all of the pupil work scored for each assessment ($n$=53, $M$=6.28, $SD$=1.317) were correlated. Pearson's $r$ was significant and strong ($r$=.68, $p$<.01). The magnitude of Pearson's $r$ can range for -1 to +1, where the strength of the correlation increases as the value approaches -1 or +1 and the direction of the relationship is indicated by the + or - (Glass & Hopkins, 1996). In this analysis, Pearson's $r$ was a positive .684, which indicates a strong direct relationship and the scatter plot in Figure 4.7 illustrates that this is a linear relationship. That is to say, the authentic intellectual work quality of assessment tasks is highly correlated to the authentic intellectual work quality of pupils' work. Pupils in this study were more likely to produce higher levels of authentic intellectual work when they were given assessment tasks that were a higher authentic intellectual quality. The inverse was also true; pupils demonstrated lower quality authentic intellectual work when they were given assessment tasks that also had low scores on authentic intellectual work. The pupils in this study produced work comparable to the level of work that their teachers gave them. When these beginning teachers provided highly authentic learning opportunities, their pupils responded in highly authentic ways. When these beginning teachers did not provide authentic learning opportunities, their pupils responded in inauthentic ways.
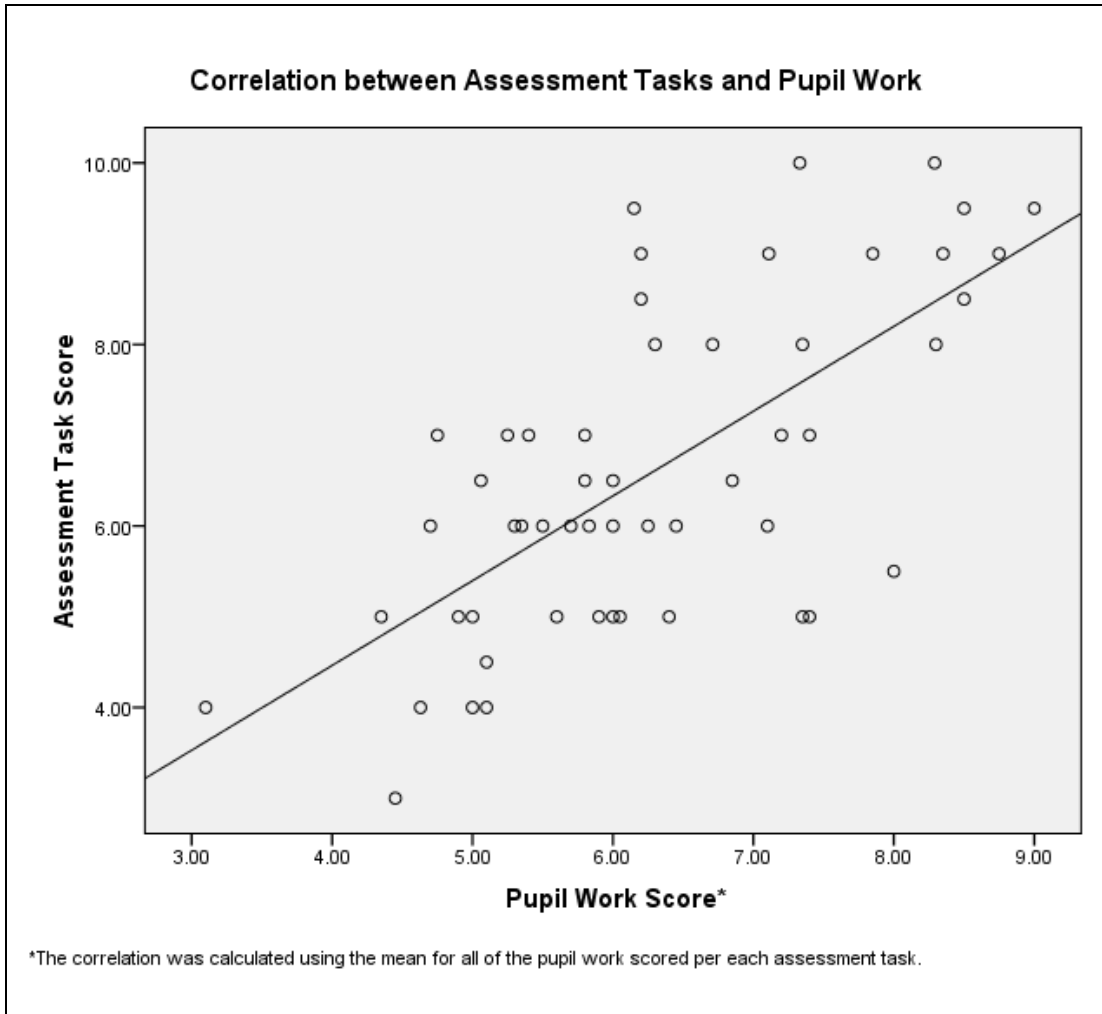
*Figure 4.7.* Correlation between assessment tasks and pupil work scores (using class mean).

The correlation of assessment to pupil work is consistent with the work of a previous study. King, Schroeder, and Chawszczewski (2001) calculated correlations of $r=.62$ and $r=.68$ between assessments and pupil work scores in their two analyses and described the relationship as a "strong" relationship. The King et al. study found the same relationship between the assessment task scores and pupil work. Pupils in that study were

also more likely to produce higher levels of authentic intellectual work when given assessments that engaged them in this type of work.

The correlations between assessments and pupil performance found in this study and the King, Schroeder, and Chawszczewski (2001) study indicate the centrality of high quality assessments to pupil learning. The construction and the creation of the assessment is a vital factor in pupils' abilities to produce high quality work. Pupils who are not given opportunities to engage in authentic intellectual work are not likely to demonstrate authentic learning. In order to create a classroom for authentic intellectual work, teachers, therefore, need to utilize or construct assessments and learning opportunities that emphasize authentic intellectual concepts. Teachers who create opportunities for pupils to engage in higher-order thinking will have pupils who are able to think critically.

As the previous discussion revealed, the beginning teachers in this study who created their own assessments had pupils who demonstrated higher levels of authentic intellectual work than pupils who were given assessments such as standardized exams or pre-packaged curriculum that someone else created. The correlation between assessment tasks and pupil work illustrates that, when given the opportunity, these teachers were able to create assessments that enabled pupils to produce high levels of authentic intellectual work. When teachers used their knowledge to create assessments, their pupils demonstrated higher levels of authentic intellectual work. These teachers were able to influence pupil learning.

In some of the cases where teachers submitted standardized assessments, it was due to the fact that they were mandated by the school or district to use that material in

their classroom. As the internal TAPL analysis in the next chapter illustrates, the teachers in this study felt pressure to utilize mandated curriculum and to teach to the test. They found it challenging to incorporate authentic intellectual work in favor of standardized curriculum. However, since other studies have demonstrated that when pupils are exposed to authentic intellectual work, they still perform higher on standardized assessments than pupils who are not given authentic intellectual work (Newmann et al., 2001), the findings of this study suggest that teachers can impact pupil learning, in terms of general knowledge and as outcomes on standardized tests, when they create their own assessments.

The external TAPL analysis illustrates that the beginning teachers in this study were able to emphasize aspects of authentic intellectual work and their pupils were able to demonstrate some level of authentic intellectual work as well. Although teachers and pupils exhibited moderate levels of authentic intellectual work, their scores were comparable to the scores of teachers and pupils of studies of veteran teachers. The level of authentic intellectual work was influenced by such factors as academic discipline and the creation of the assessments. Teachers who created their own assessments created more authentic intellectual work opportunities than pre-packaged curriculum and standardized assessments. This is significant, especially as this study found that the quality of assessments influences the quality of pupil learning. The teachers who created their own authentic assessments were more likely to have pupils who demonstrated higher levels of authentic intellectual work.

The external TAPL analysis drew primarily from an evaluation of assessment tasks and pupil work samples based on a rubric for authentic intellectual work. The next chapter uses qualitative data from the internal TAPL analysis to examine teachers' thoughts about assessment and pupil learning in relation to authentic intellectual work and incorporates the external TAPL findings to construct a more developed understanding of teachers' practices and their pupils' learning.

**CHAPTER FIVE: THE AUTHENTIC INTELLECTUAL QUALITY OF TEACHERS' UNDERSTANDINGS OF ASSESSMENT AND PUPIL LEARNING**

In this study, Newmann's (1996) framework of authentic intellectual work is used in the Teacher Assessment/Pupil Learning (TAPL) protocol as a way to examine the learning opportunities beginning teachers provide and the way they think about their goals for assessment and pupil learning. In the previous chapter, I presented the results of the "external TAPL," an analysis in which teachers' assignments/assessments and their pupils' performance on those assignments were evaluated according to a rubric for authentic intellectual work (RISER, 2001). This rubric quantitized teachers' assessments and pupil work in order to describe and compare the quality of teachers' practices and pupil learning in a uniform manner across all teachers in this study. As that chapter illustrated, the teachers and their pupils, on the whole, utilized assessments and produced work that could be thought of as having "moderate" levels of authentic intellectual work.

The external TAPL analysis, however, only provides one vantage point into teachers' practices. The second component of the TAPL protocol – the "internal TAPL" – provides a deeper look into the reasons teachers made assessment decisions and how they understood pupil learning by focusing on descriptive qualitative data. Mixing quantitative and qualitative methods allows researchers to get at issues in ways that one method alone cannot do and enables the triangulation of findings (Greene et al., 1989). In the internal TAPL analysis, teachers' understandings and goals related to assessment and pupil learning, as described through interviews, were examined to see the extent to which their thoughts and goals aligned with the underlying concepts of authentic intellectual work. In this chapter I present the findings of the internal TAPL, an analysis of the qualitative data

that gets at teachers' perspectives. Drawing on interview data, I address the following

question: How does the way that teachers talk about their goals and understandings of

assessment and pupil learning over time reflect the concepts that are central to the idea of

authentic intellectual work?

This chapter also focuses on the relationship between the qualitative and

quantitative analyses, or what Onwuegbuzie and Teddlie (2003) refer to as the *data*

*comparison* and *data integration* stages of mixed methods data analysis. First, through

data comparison, I use the findings of the quantitative external and qualitative internal

TAPL analyses to compare the ways in which teachers talked about their goals and

understandings of assessment and pupil learning to the scores their assessments and pupil

work samples received when evaluated based on the rubric for authentic intellectual

work. This analysis provides a way to triangulate the findings of the external analysis.

Then, through data integration, I use both the quantitative external and qualitative internal

TAPL data to examine the factors and conditions that influenced how beginning teachers

in this study provided authentic learning opportunities for their pupils. Finally, the

chapter concludes with two case studies of two very different classrooms to illustrate how

teachers' use of authentic intellectual work was influenced by their understandings of

assessment and pupil learning and different contextual factors. Only one of these teachers

was able to engage pupils in high levels of authentic intellectual work while the other

offered few opportunities for pupils to do so.

My analysis suggests that the degree to which the beginning teachers in this study

were able to construct learning opportunities that emphasized important concepts of

authentic intellectual work, such as construction of knowledge, disciplined inquiry, and value beyond school, was influenced by their beliefs about teaching and learning and the ways these beliefs interacted with certain contextual factors in their classrooms. Teachers whose goals for students' learning included critical thinking and constructing arguments, were more likely to use assessment tasks that received high scores on the rubric for authentic intellectual work. That is to say, teachers who thought about concepts related to authentic intellectual work, created learning opportunities that emphasized authentic intellectual work. Conversely, teachers whose goals for assessment and pupil learning were primarily lower-order and had very little connection to authentic intellectual work, were more likely to have assessment tasks and pupil work that received low scores on the external TAPL analysis. With a few exceptions, the teachers who did not consider goals related to authentic intellectual work rarely provided learning opportunities for their pupils to engage in authentic learning.

In this chapter I first present an analysis of the ways in which teachers talked about their goals and understandings of assessment and pupil learning to see the extent to which their thoughts are consistent with the framework of authentic intellectual work. Based on the things they mentioned when talking about their goals for assessment and pupil learning, I placed teachers on a continuum according to how their learning goals aligned with the framework of authentic intellectual work. Then, I explore different conditions and contexts that influenced teachers' abilities to implement authentic practices which include time, school culture/accountability, student ability and motivation, behavior/classroom management, and content area. Finally, I present two

case studies – Matt and Mara – to illustrate differences between teachers who are able to provide authentic intellectual work for their pupils and teachers who do not engage pupils in this type of learning.

## Teachers' Goals for Assessment and Pupil Learning

As described in Chapter 3, the teachers in this study took part in five TAPL interviews during the course of their preservice period and first two years of teaching (Student Teaching, Year 1 Fall, Year 1 Spring, Year 2 Fall, Year 2 Spring). Participants were asked to bring one assessment task and one class set of pupil work on that assessment task to each interview; these were the artifacts evaluated in the external TAPL analysis presented in Chapter 4. This chapter looks at the internal TAPL analysis, an examination of what teachers said in the interviews about their assessments and pupil work. During the TAPL interviews, participants were asked questions about the creation and implementation of the assessment that they brought; their learning goals for pupils and evaluation of pupil learning; how well the assessment worked and might be modified in the future; and what they considered "high," "medium," and "low" examples of pupil performance. Table 5.1 presents the questions in the portion of the interview that is referred to as the "TAPL Interview Protocol." The purpose of the TAPL protocol within the larger QCS study was to see how the beginning teachers in this study thought about assessment and pupil learning by talking about specific examples of their assessments and pupil work. It is important to note that the TAPL interview protocol was embedded within interviews that also asked additional questions related to teachers' experiences in the teacher education program and teaching as described in Chapter 3. (See Appendix A

for the full interview protocols.) Occasionally, responses to those questions were also

considered in this analysis.

Table 5.1

*Teacher Assessment/Pupil Learning (TAPL) Interview Protocol*

| Interview Question | Probe |
| --- | --- |
| 1. OK, let's take a look at the assignment you brought. Although we only have one assignment, it would be helpful if you could walk me through the larger unit it draws from. You could work backwards and describe the larger unit or you might want to move chronologically through the unit and describe the pieces that led up to this final assessment. | • How does it fit into a larger unit? <br> • Was this something you devised yourself? <br> • Was any part of this lesson from a preexisting lesson that you adapted? <br> • Why did you decide this lesson/assignment/assessment would be appropriate? <br> • How much autonomy did you have in creating the lesson or assignment? |
| 2. What did you want students to get out of this activity? How do you know whether or not students accomplished what you wanted them to get out of this activity/lesson/unit? | • How did you evaluate these assignments (rubric, scoring, etc.)? |
| 3. Let's now look at your examples of a high, a medium, and a low-level response. How do these samples compare to the overall class? | • Is this work representative of the class? <br> • Is this what you expected? |
| 4. Did the students who completed these examples meet your expectations? Why or why not? | • What might you do differently in the future for each of these students? |
| 5. Why did you choose these? | • Tell me about these three students (SPED, ELL, Bilingual). |

As seen in Table 5.1, the TAPL interview protocol did not ask questions about the

framework of authentic intellectual work per se. Instead, the questions addressed

teachers' goals for specific assessments and their understandings of pupil learning on

those assessments. However, for the qualitative analysis, I used Newmann's (1996)

framework of authentic intellectual work as a lens to analyze teachers' responses to these questions. Teachers' responses were coded according to how they thought about teaching and learning, with particular attention to how those ideas related to the concepts behind authentic intellectual work. Codes included terms such as critical thinking, interpretation, evaluation, supporting generalizations, and connecting to students' lives. The codes were then conflated into categories that represented the three different criteria for authentic intellectual work: construction of knowledge, disciplined inquiry, and value beyond school. I used the descriptions of each standard as delineated by the RISER rubric to place the codes into the three criteria for authentic intellectual work.

Construction of Knowledge included codes such as interpretation, analysis, synthesis, evaluation, as well as the absence of construction of knowledge such as recall, memorization, and lower-order thinking. Disciplined Inquiry included codes such as elaborated written communication, generalizations, supporting arguments, and conceptual understanding, as well as the absence of disciplined inquiry such as multiple choice and factual recall. Value Beyond School included codes such as connections to students' lives, community, and issues outside of the classroom, as well as the absences of value beyond school such as isolated information. For example, when teachers talked about wanting students to support their arguments with evidence, those examples were included within the category elaborated written communication because supporting generalizations is part of Newmann's definition for that criterion. Codes that did not fit within the three criteria, such as issues related to "time," were examined and incorporated into the analysis of contextual factors.

During the analysis of teachers' responses, it became clear that teachers could be grouped into three categories representing the extent and frequency with which they talked about their goals for assessment and pupil learning in relation to authentic intellectual work. Teachers were placed into these three categories along a continuum based upon the quality and quantity of their coded responses and the degree to which their responses reflected the definition of authentic intellectual work. In the first category, teachers who consistently spoke about authentic concepts such as critical thinking and supporting arguments were considered "high" on the continuum, meaning that they thought about assessment and pupil learning in ways that reflected the concepts of authentic intellectual work. In the second category were teachers who occasionally described their goals for pupil learning in terms related to the three criteria of authentic intellectual work, such as higher-order thinking and conceptual understanding, but they also had goals that were less authentic, emphasizing ideas such as factual recall. Teachers in this category were labeled as "moderate." Teachers in the third category were considered "low" because they rarely talked about their goals for teaching and learning in terms of authentic intellectual work. Instead, these teachers expected pupils to reproduce information or did not think deeply about their assessment practices and pupils' learning.

The following example illustrates how teachers were placed into the three different categories. As part of the TAPL interview protocol, teachers were asked to explain why they chose to utilize a particular assessment. Mara's description of what she wanted her pupils to learn in social studies was considered a "high" response:

> [One goal] is just understanding how things happen.… you don't have to memorize exact dates but understand how they relate to each other and understand

how it makes sense that the American Revolution happened before the French Revolution because the French look[ed] at what the Americans [did]. (Mara, Interview 8[9])

In this instance Mara wanted her students to focus on higher-order conceptual understandings of how historical events relate to each other, instead of having students focus on lower-order memorization of historical dates. This aligns with disciplined inquiry and disciplinary concepts. Teachers at the high end of the continuum consistently referred to their learning goals in authentic terms.

Teachers in the "moderate" category were less consistent; there were instances where these teachers stressed authentic intellectual work, but, also felt it was important to emphasize lower-order skills. For example, Rachel wanted her elementary pupils to understand multiplication at a conceptual level. She explained that her goal was for students "to understand both that 4 times 3 means four groups with three things in it. And for them to understand that then you can add three, plus three, plus three, plus three, and to understand that you can skip count three, six, nine, twelve, and it's all the same thing" (Rachel, Interview 8). However, at the same time, another goal that she had was "to help [students] with their math fluency…knowing their facts really quickly" (Rachel, Interview 11). Here she expressed the need for pupils to be able to access their multiplication tables in a rapid manner from memory. These two examples represent two competing views of learning multiplication because one emphasizes conceptual understanding while the second emphasizes recall. Instances where teachers expressed these competing ideas were considered moderate on the continuum.

---

[9] As mentioned previously, to maintain the integrity of the larger Qualitative Case Studies project within which this study is embedded, the interview titles from the larger study were used.

Teachers who were placed "low" on the continuum were ones that made little to no reference to concepts related to authentic intellectual work. For example, Craig's goal for a middle school science assessment he used, where pupils created "trading cards" to describe the different planets, was for students to know "the unique characteristics of the different planets" and use the cards to "revie[w] for the test" (Craig, Interview 8). For Craig, the purpose of the assessment was to help students recall facts about the planets and memorize that information for a future assessment, all of which are lower-order skills.

Figure 5.1 illustrates where each of the teachers in this study were placed on the continuum of how their understandings of assessment and pupil learning reflected the concepts of authentic intellectual work. Teachers were placed into one of the three categories based on a holistic analysis of all of their interview data. Teachers within each category were not ranked, and their names in Figure 5.1 are listed alphabetically per category. Amanda, Elizabeth, and Mara were placed in the "high" category; they continually considered authentic learning. The teachers in the "moderate" group included Lola, Rachel, Riley, and Sonia; they addressed authentic issues but also considered lower-order skills, such as recall, as significant aspects of learning. Teachers at the "low" end of the continuum included Craig, Mark, Matt, and Sylvie; these teachers' understandings of assessment and pupil learning rarely reflected the principles of authentic intellectual work.

**Teachers' Understandings of Assessment and Pupil Learning in Relation to the Concepts of Authentic Intellectual Work**

| Low | Moderate | High |
|-----|----------|------|

| Craig | Lola | Amanda |
| Mark | Rachel | Elizabeth |
| Matt | Riley | Mara |
| Sylvie | Sonia | |
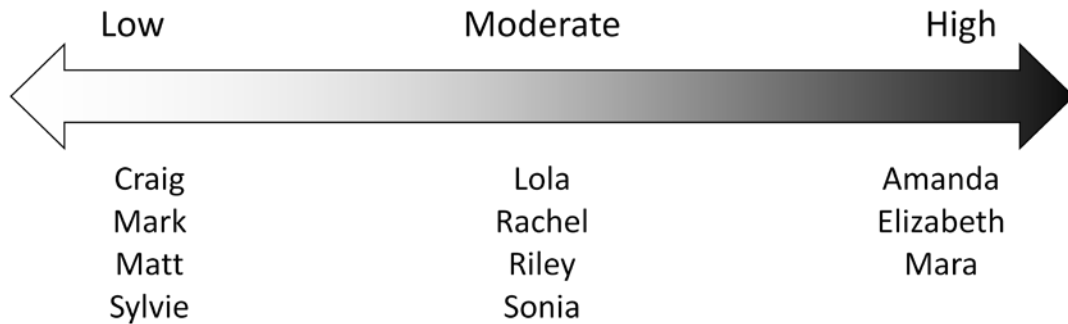
*Figure 5.1*. Continuum of teachers' understandings of assessment and pupil learning.

The following section of this chapter describes the general ways that teachers in the three different categories on the continuum – high, moderate, and low – talked about their goals for and understandings of assessment and pupil learning in relation to authentic intellectual work.

**Teachers Whose Goals Reflected High Levels of Authentic Intellectual Work**

Amanda, Elizabeth, and Mara were the three teachers who consistently used concepts related to authentic intellectual work when they talked about their goals and understandings of assessment and pupil learning. They consistently talked about students' learning and assessment in terms of the three criteria for authentic intellectual work: construction of knowledge, disciplined inquiry, and value beyond school. It is worth noting that Amanda and Elizabeth were secondary English teachers while Mara was a secondary social studies teacher. As seen on the external TAPL analysis, these were subjects that appeared to offer more opportunities for authentic intellectual work.

These teachers consistently talked about concepts related to the construction of knowledge. For example, they continually referred to wanting their pupils to become critical thinkers. When Elizabeth was first interviewed at the beginning of her teacher education program, she described her goals as a teacher as "keeping expectations high and pushing [students] . . . so they're learning to be critical. . . . looking at things analytically, critically, asking questions about what they read" (Elizabeth, Interview 1). Throughout her first two years she repeatedly asked students to analyze texts and construct their own interpretations. She described how she sought ways to scaffold critical thinking, saying that her cooperating teacher helped her to start thinking about how "to break down an assignment that requires high-level thinking but not make it overwhelming" (Elizabeth, Interview 9). Her assignments reflected this scaffolding approach; they often consisted of multiple parts where pupils would do things such as fill out a graphic organizer to identify particular literary elements and supporting textual evidence before writing an essay about their ideas. She focused on how pupils could construct their own interpretations and arguments.

Similarly, Amanda and Mara frequently spoke about how important it was for their pupils to know that there were no "right" answers, emphasizing that pupils should think for themselves. For instance, describing the way she graded a persuasive essay that she assigned pupils during the student teaching period, Amanda said, "…as long as they support [their argument], then I gave them credit for that. Because I felt like that is the point of a persuasive essay. You don't necessarily have one right answer. It's just more like who can argue the best" (Amanda, Interview 5). Although developing arguments

requires more than just "who can argue the best," Amanda at least focuses on having

pupils construct an argument. She went on to explain that she did not want to "censor"

pupils' answers by evaluating the content and, instead, focused on how pupils' supported

their opinions. On another occasion, Mara emphasized that pupils had to come up with

their own interpretation. Describing one lesson from her first year of teaching where she

asked students to determine which causes of World War I were the most important, she

explained:

> And then I asked them [which cause] do you think is the most important? And it
> was funny because I had them doing it in a group activity and some of
> them,…Well, the group can't decide.  And I was like, "Well it's okay if you put
> down two." So it was really funny they were so freaked out about the right
> answer.  And I'm like, "There is no right answer." And they were like, *There's
> no right answer*!" (Mara, Interview 8)

Mara reinforced critical thinking by creating lessons where pupils debated topics and

voiced their opinions on particular issues. Here pupils were challenged not to just repeat

what they had learned about the causes of World War I but to really think about what

cause was the most important. She provided them with a format to evaluate evidence and

come up with their own interpretations instead of just choosing one pre-decided answer.

One aspect of construction of knowledge in the framework of authentic

intellectual work is for pupils to interpret, synthesize, analyze, and evaluate knowledge.

These teachers talked about ways that pupils could use knowledge and not reproduce

information. For example, Amanda's description of a vocabulary assessment illustrates

her interest in having pupils know more than just definitions. She said, "On the quizzes I

always have them use…their words in a sentence. That, they find that to be the most

challenging thing more so than definitions…memorizing definitions is not that hard…it's

more like using it" (Amanda, Interview 8). Amanda was more interested in assessing

pupils' abilities to demonstrate their understandings than to repeat a definition.

Describing the assignments that she used in her second year of teaching, she said:

> So I felt like this was a really appropriate way to assess their understanding of the philosophers and of the book because they had to draw examples from both sources and in order to write the paper they had to understand where the philosophers were coming from…and they had to understand…what is Rousseau's view of the natural state of man and they had to know … how does that apply to the book…and to today's society.  So they had to think of examples…of our own government system or they could even do, I said they could even do outside research if they…weren't sure and they wanted to talk about another country's government. (Amanda, Interview 10)

In this example, Amanda thought directly about the ways pupils could apply their

understandings and create new knowledge, wanting pupils to make important

connections. She went on to explain that she wanted students to "demonstrate that you've

learned not just history and ELA content but also the ability to analyze literature and tie it

to historical themes and literary themes" (Amanda, Interview 11). Her goal was for pupils

to interpret literature. Similarly, Elizabeth continually stressed literary analysis. She

wanted pupils "to learn how…to make sure that they understand what these literary tools

are…[and] how to use them.  And how to show how they do help support the theme [and]

to write a true analytical paper" (Elizabeth, Interview 10).  Her goals went beyond

reading comprehension to interpretation. Fostering the skills of evaluating and

interpreting information, essential skills for participation in democratic society, were

central to these teachers' goals.

    These teachers also emphasized the value of having students explain their

understandings in an elaborated way, which is consistent with the standard disciplined

inquiry: elaborated written communication. Extensive writing was a central component in each of their assessments and all three teachers talked about ways in which writing enabled pupils to demonstrate their conceptual understandings and conduct higher-order thinking. Across the five TAPL interviews, Amanda submitted four essay assessments, Elizabeth submitted five essay assessments, and Mara submitted two essay assignments, two PowerPoint assessments, and one poster assignment, where both the PowerPoint and posters required extensive writing. Not only did they assign writing assignments but they spoke about the importance of allowing pupils to express their knowledge in detail. At the beginning of her student teaching experience, Amanda described the type of assessment that she preferred to use as follows:

> I think what you get from an essay or a piece of writing is …a better judgment…of what the students know than multiple choice because multiple choice is simply recognition.  And that's the less thinking involved than putting pencil to the paper and thinking.  So I think back to assessments, I, my own style would probably be more leaning toward giving them papers and creative writing or a portfolio even.  Like a portfolio of all their journal entries and I'd rather do that than a quiz or a test… (Amanda, Interview 3)

In this example, Amanda talked about the importance of having pupils describe their understandings in writing as opposed to a multiple choice assessment. Considering that assessments that focus on "narrative or expository" writing have a higher value than multiple choice assessments on the RISER rubric, Amanda's own beliefs about writing and assessment, as seen here, are aligned with higher levels of authentic intellectual work. Similarly, Elizabeth did not consider multiple choice questions to be the best way to engage pupils in higher-order thinking and saw writing as a way to allow students to develop and demonstrate these skills. Elizabeth said, "So, higher level thinking –

definitely more so in the essay I see that more, but I make sure that I give them some questions on the test that make them think a little bit, they're not just multiple choice questions" (Elizabeth, Interview 8). She saw writing as the best place for pupils to engage in higher-order thinking, although she admitted that she tried to incorporate higher-order thinking in other types of assessments. These teachers saw writing as a way to support pupils' construction of their own knowledge and illustrate pupils' conceptual understandings. For instance, Mara described an assignment where students had to create definitions of political theories: "And they had to read those and then each group had to come up with their own working definition, two quotes that illustrated what the essential meaning of the [theory] was, and then they had to illustrate a pictogram to show that they really understood what it meant" (Mara, Interview 10). As this example suggests, Mara also allowed pupils to demonstrate their understandings in an elaborated way through a graphic representation, which is also emphasized in the rubric. Overall the three teachers who were placed "high" on the continuum preferred to engage pupils in extensive writing as a way to truly assess pupil learning and is indicative of authentic intellectual work.

This group of teachers also considered the standard elaborated written communication through their emphasis on having pupils use evidence to support generalizations, a quality that is valued on the RISER rubric. These three teachers thought about ways to emphasize these characteristics in their assessments. "So I wanted to choose an inflammatory article," Amanda explained, "to get them into that mode of…'I have an opinion about this and I can support it with this evidence'" (Amanda, Interview 5). She also valued this idea when evaluating assessments and graded pupils on

their ability to support their answers in a persuasive essay instead of judging the position they took. As English teachers, Elizabeth and Amanda routinely stressed the importance of using evidence from a text to support a written argument. Describing her goals for one assessment, Elizabeth explained,

> I was hoping they'd sharpen their analytical skills a little more. I really wanted them to, again, be able to pull things, significant passages out of text and then analyze them. I just take that for granted 'cause I can do that real quick,…[b]ut for them, they just really like to talk in first person and say what they think….I just really wanted them to just pick a passage and just to be able to analyze it and just show the reader, and show me, that they understand how they can connect the literary device to the theme… that was the … hardest part, to show how that quote actually ties into what they're trying to prove their thesis. (Elizabeth, Interview 5)

Elizabeth challenged her pupils to support their ideas. Similarly, one goal that Amanda had for her students "was to be able to back up their main ideas with textual evidence" (Amanda, Interview 5). She explained that she "did a lot of direct instruction" where she selected a page from the text and asked pupils to identify "a good example to back up your main ideas" (Amanda, Interview 5). Both Elizabeth and Amanda incorporated the same graphic organizer into their assessments as a way to have pupils organize their writing around a main idea, evidence, analysis, and a linking sentence. Similarly, Mara wanted her pupils to realize that "when you're writing a history paper, you need to use history to prove what you're saying" (Mara, Interview 9). Their emphasis on supporting generalizations and ideas suggests that these teachers valued the process of evaluating information and constructing rational arguments, all of which are central to preparing pupils to develop and critique arguments in the public sphere outside of the classroom.

In addition to creating assessments that emphasized supporting generalizations, Amanda, Elizabeth, and Mara also evaluated pupil work based on how well pupils were

able to do this. Describing her thought process behind grading assignments, Elizabeth made comments such as "some of her textual evidence is a little skimpy," referring to a student who received a low grade (Elizabeth, Interview 8). Here is how Elizabeth described one pupil's essay on the novel *The Things They Carried* by Tim O'Brien:

> This student got a C-. … So her strengths were she had a very unique thesis.… she argued that soldiers avoid emotion to avoid attachment and signs of weakness yet such lack of emotion leads to isolation. … So good thesis. …But she's got some good examples to show how soldiers used defense mechanisms to kind of avoid confronting their emotion but her analysis of each piece of textual evidence is really weak. In some cases she just like ends the paragraph with a quote without doing anything with it. … You're also arguing that the lack of emotion leads to isolation of a soldier. I don't see echoes of this point until the second to last paragraph when you explore O'Brien's reunion. (Elizabeth, Interview 11)

Her evaluation of this pupil's work was based on the extent to which the pupil used evidence from the text to support the thesis. Describing a different pupil's work on another assignment, she said, "I see a thesis statement at the bottom of her introductory paragraph. She has a nice way of incorporating quotes because it's not abrupt, and the way …she analyzes it, it makes sense because [it] connects to the thesis" (Elizabeth, Interview 5). In these excerpts, Elizabeth defined pupil learning by how well pupils could construct and support a thesis. Similarly, Mara evaluated students on their abilities to support their point of view. Explaining how she graded an essay about different options the United States could consider to confront genocide throughout the world, Mara explained, "[This student] did a good job with supporting her option and talking about it. And then she takes it a step further where she was critiquing her own ideas and saying the faults in it and really getting to the heart of the issue" (Mara, Interview 9).The ability to support generalizations and evaluate arguments, concepts central to the framework of

authentic intellectual work, were key components for these teachers in assessing pupil learning

The teachers on the high end of the continuum also talked about assessment and learning relative to the standard value beyond school. They connected their lessons and assessments to their pupils' interests and their community. In her assignment on the novel *The Death of a Salesman*, Amanda brought in music that her pupils listened to as a way to make a connection between them and the novel's theme of "the American Dream." As part of this lesson Amanda incorporated an article that she described as "perfect because it related to something students could relate to [because] it was about how the new American dream is that everyone wants to get famous really quick" (Amanda, Interview 5). The article used the television shows "American Idol" and "Survivor" – shows her pupils watched – as illustrations of the American Dream and provided a contemporary example that pupils could relate to and analyze.

Another goal Amanda had for an assessment was for pupils "to be passionate or interested…about what they're writing" (Amanda, Interview 8). For that assignment she asked pupils to write personal prose poems that allowed them to express feelings and experiences that were significant in their own lives. Pupils were able to gain disciplinary content knowledge about prose poems and vignettes in a more meaningful and authentic way by creating intensely personal pieces of their own writing. Similarly, Elizabeth sought ways to connect curriculum to issues that pupils could connect to outside of school or in their personal lives. Drawing on pupils' backgrounds, she chose to teach the novel *Interpreter of Maladies* because she felt that her pupils could identify with its

immigration story. "I wanted to teach that book because I thought it was culturally relevant to them," she said, "not so much because they were Indian but the fact that this idea of the immigrant experience and appreciating the difficulties, like, all of the complexities involved in relocating and all of the issues that come with it" (Elizabeth, Interview 10). This emphasis on relating material to pupils' lives created a more genuine learning experience for pupils.

These teachers also made explicit the relevance of schoolwork and created situations for the work that were similar to "real world" experiences. Amanda had students write about an experience in their lives to present to first graders. She described the experience as follows:

> But I think it was a pretty positive experience overall because they finally were able to have an audience, like, for their work. They hadn't been able to read it to their classmates as much as I wanted them to, you know, like, only some of them got to do the paired reading based on the time restraints we had. So, you know, for them to have that audience and be feeling like, oh, you know, maybe I'm making a difference for these kids. Maybe I can actually teach them, like, a real skill. You know. I thought that was really important because they need to realize that they, they are role models. They are leaders. You know, for their community, you know, for, for the younger kids. (Amanda, Interview 9)

Amanda fostered ideas of community with her students by helping them take a leadership role in their school community, while creating a meaningful assessment that had a concrete purpose. Similarly, Elizabeth also designed an assessment that involved having pupils actually engage in the type of work that occurs outside of the classroom. As part of her assessment for the book *Interpreter of Maladies*, she asked students to interview immigrants so they could understand experiences and she "thought a lot about that when I wanted to do this book….[t]hey can go out and explore and like do kind of an inquiry

type thing" (Elizabeth, Interview 10). In this experience, pupils were able to understand the themes in the novel by interacting with people who shared experiences similar to the characters in the novel.

In addition to making connections to pupils' personal experiences, these teachers also specifically addressed issues related to their pupils' own communities. Mara's goals for her social action project included to foster "[a]wareness about the world that they live in" and that even though students "may seem insignificant," they are able to make an impact (Mara, Interview 11). Mara wanted pupils to recognize that they were part of their community. Similarly, Amanda asked her pupils to challenge the things they saw in their urban community. She assigned the book *Fast Food Nation*, which examines America's food industry, and asked students to relate the book to what they saw in their neighborhood. She explained:

> We actually just had a discussion about this the other day and they were saying how they noticed there's a lot more fast food places in urban areas and they always wondered why? You know, they were like, "Why is it that [fast food places] decide to come to urban areas, you know? Is it because there are kids who are more likely to eat fast food maybe because their parents can't cook for them? Is it, does it have something to do with the education level?" (Amanda, Interview 10)

These two examples illustrate that these teachers engaged pupils in a critique of societal structures outside of their classroom. This is consistent with the framework of authentic intellectual work but also reflects their teacher preparation program's emphasis on teaching for social justice.

As teachers in urban schools, Elizabeth and Amanda both saw connection to pupils' lives as an important "hook" for students to buy into the work that they were

doing in school because they felt their pupils needed to be motivated to learn beyond the

promise that they would need the knowledge and skills for college. Amanda described

her use of work that demonstrated the value beyond school as follows:

> Homework, too, is a very, like, problematic thing, which doesn't usually get done
> and so, so it's, so it's a lot of things and I guess another thing I've learned
> working with diverse populations is that you've got to be able to make everything
> relevant to their, somehow tie into what they already know.  You know, you've
> got to have that hook because they're not necessarily going to think of college as
> the hook.  Like, that's not, for all, not all of them will think, oh, like, I need to
> read, like, *Tale of Two Cities* because I might see a text like this in college, no,
> they need something else to dig into.  You know what I mean?  And so I've
> realized you've got to make every assignment relevant to their lives and that's,
> you know, a lot of, that's just their age, too.  They like to talk and write about
> themselves so you got to find assignments that allow them to express themselves,
> who they are. (Amanda, Interview 9)

Here Amanda sees connecting material to pupils' lives as a necessity in holding pupils'

interests.

It is unclear from the data in this study, whether Amanda believes that this need is

unique to pupils in urban schools. Her comments could suggest that she believes her

students will not go to college, and therefore need a different reason to do their work in

school. However, her comments could also illustrate the argument Newmann (1996)

makes in the framework for authentic intellectual work: all students need to see the

relevance of the work they do in school. Truly authentic experiences involve situations

that actually occur in the real world. Elizabeth, made this point when she discussed one

goal that she had for her third year of teaching:

> … my next goal for next year is making sure that they understand why each piece
> of literature that we read is really relevant to now because it really is still
> pertinent.  We can always see a reflection of issues or ourselves in what we read.
> I did a small unit on Genocide recently with the seniors and … I just needed
> something to kind of rope them in for a couple of weeks and I gave them some

local examples of teens that are making a difference in their community. And so in that regard I think in terms of teaching, making them aware that they have some sort of responsibility or they are able to do things for others. That came out in that kind of particular unit but whether or not students can kind of see how they can make sense of the literature they read in their own lives I don't know if they're there yet. I don't know if I'm there yet, if I'm doing that really. (Elizabeth, Interview 11)

Elizabeth wanted pupils to be able to use what they learned in school to apply it to their own learning outside of school.

These teachers mentioned their teacher education program as a source they drew from as they considered assessment and pupil learning and these elements are related to the concepts of authentic intellectual work. For example, one of the things Elizabeth said she took away from the program was "[t]he constant reminder to be pushing for higher-level thinking" (Elizabeth, Interview 9). That sentiment was echoed by Mara as she continually referenced her social studies methods course and its emphasis on higher-order historical thinking skills. Many of Mara's assessments, such as the PowerPoint on Renaissance art, the "Take a Stand" activity where pupils had to state their opinion about a particular topic or question, and the social action project where pupils became involved in a community issue that was of interest to them in order to make a difference were all assessments introduced in the methods course and were all considered to be highly authentic activities. Amanda and Elizabeth also used materials from their coursework in the preparation program such as a graphic organizer that helped pupils learn how to support arguments with textual evidence from their readings. Since these activities attended to aspects of authentic intellectual work, this suggests that the teacher

preparation supported these teachers goals to utilize assessments that aligned with the framework.

One final point about teachers in this "high" category is that two out of the three teachers made comments about preparing pupils for participation in a democratic society when they talked about assessment and pupil learning. For instance, Amanda challenged her students to think about the United States' "history of institutionalized racism and segregation" and to question their schools and communities. She felt that "critical thinking is really important in terms of social justice" (Amanda, Interview 6) and that social justice involved "preparing kids to succeed in [and] be an active participant in a democracy…in this country" (Amanda, Interview 9). One of the things that became clear for Mara during her student teaching was the relationship between teaching and democracy. She said, "I never really thought about it until even teaching this course, how much of it is preparing them to be citizens in this country, never really thought about that or thought how I was going to address it, but every day we talk about even current events or even huge things…" (Mara, Interview 6). Mara's views on teaching grew to incorporate the idea of preparing students for democratic life. When asked about teaching for social justice at the end of her teacher preparation program she said, "I really, really see the whole citizen aspect of it…just how important it is to make these kids aware" (Mara, Interview 6).

As these examples and excerpts demonstrate, the teachers in the "high" category continually thought about assessment and pupil learning in ways that were consistent with the framework of authentic intellectual work. Their emphasis on concepts related to

the framework stands in sharp contrast to the goals and understandings of teachers who had moderate or low levels of authentic intellectual work.

**Teachers Whose Goals Reflected Moderate Levels of Authentic Intellectual Work**

There were four teachers in this study – Lola, Rachel, Riley, and Sonia – whose talk about assessment and pupil learning reflected "moderate" levels of ideas and concepts related to authentic intellectual work. These four teachers occasionally mentioned the concepts of authentic intellectual work, but they also stressed goals that were less consistent with the framework, such as lower-order thinking, because they felt that their pupils needed to know certain core knowledge or they faced challenges implementing authentic learning opportunities.

There were some instances where these four teachers talked about their goals for assessment and pupil learning in ways that aligned with the standards for authentic intellectual work. For example, Sonia's comments reflected the standard construction of knowledge. She explained that she wanted her pupils "to learn to be thinkers" (Sonia, Interview 1) and described a successful assessment as one that fostered "really good higher-order thinking or just critical thinking" (Sonia, Interview 5). In another example, Riley's goals for reading aligned with the criterion disciplined inquiry by stressing conceptual understanding and elaborated written communication:

> One of the things I really try to emphasize with the reading is just comprehension and talking about the books and kind of understanding it at a higher level…rather than just the right there in the text questions, kind of like the why, the meaning behind it and also teaching them to respond to what they've read in writing. I think it's important to be able to communicate what you know in writing. (Riley, Interview 11)

238

Here she wanted her students to describe their understandings of their reading comprehension in writing. At other times Riley talked about the importance of writing. "[W]ith an essay you can write what you know. It's not like with multiple choice," she said. "You can't really show the teacher what you know" (Riley, Interview 1). Riley frequently mentioned the importance of writing as a way to explain understanding, even in subjects like math. For example, describing one assessment she said,

> Rather than just give a worksheet and see what they got for an answer, I asked them for one of their homework problems to write out in words how they solved the problem. And I think that's a really good way to assess because they have to really think about the process. And you kind of get into their head of what the process that they're going through as they solve the problem. (Riley, Interview 5)

Riley focused on having pupils explain their thinking and wanted them to use writing. She saw this as a way to get pupils to "really think." In another example of the standard elaborated written communication, Lola talked about adding opportunities for students to provide "explanations… in multiple choices so that you know they get some credit for the multiple choice question itself but then they have to explain why they chose that" (Lola, Interview 8). She came to realize the importance of allowing student to express their ideas beyond multiple choice answers. These are just a few of the ways they expressed authentic intellectual work. Some of these responses were similar to the teachers in the "high" category, but these teachers had fewer instances where this was the case.

Although there were moments, like those described above, when these teachers talked about goals and learning in ways that were consistent with authentic intellectual work, there were also times when their goals were less than authentic. The teachers in this category talked about the basic skills their elementary students needed. They stressed

things such as having students recall social studies content knowledge, memorize basic

math facts such as multiplication tables, and focus on specific writing skills such as

"focus correction areas" for grammar. For example, Sonia's description of the goals for

the social studies PowerPoint presentation she assigned on the ancient Chinese dynasties

illustrates this focus:

> But basically I wanted them to have, the basic information that I wanted them to
> to be able to present was the name of your dynasty. If you knew the name of some
> of the emperors that's fine. I wanted them to have a map of where their dynasty
> ruled. I wanted them to know from what year to what year the dynasty
> ruled…[and] four important things…about their dynasty. (Sonia, Interview 5)

Her objective for this assessment was for pupils to be able to include specific facts in

their PowerPoints to demonstrate their knowledge of the Chinese dynasties. This goal

was primarily factual recall of content knowledge and did not engage pupils in

interpreting or synthesizing information. Oftentimes these teachers sought to assess

pupils' knowledge of specific content or that they were able to demonstrate certain skills.

That is not to say that these teachers saw lower- and higher-order thinking as an

either/or situation. More often than not, these teachers grappled with how to balance

basic skills and analysis. All four of the teachers at the moderate level talked about the

importance of balancing skills and higher order thinking. For example, Riley spoke about

helping her elementary pupils understand a social studies textbook. She said,

> So I think you need a balance.  Because I remember specifically [in] my social
> studies [methods] class  [the professor] was making a point by saying [the
> textbook]  isn't the only thing you would teach them history with because it has
> its biases and, it's not as rich as maybe reading a primary source document, kind
> of them putting the pieces together on their own.  So I kind of think you need a
> balance because you do need that skill though, too. (Riley, Interview 9)

Riley wanted pupils to be able to analyze primary sources, which is an authentic practice for historians, but at the same time recognized that her students needed to have the skills to read a social studies textbook, which is a less authentic task. Similarly, Sonia also mentioned the need to balance lower-order and higher-order thinking with her pupils. Discussing her views on teaching math she said,

> …some sort of those assessments that, yes, are very time consuming but are a little closer to assessing the real knowledge that children bring,… kids need experience with manipulatives, they need to see how the math works and it needs to be concrete before it becomes abstract. So I mean I think that a balanced approach… you need to drill certain facts but they need to, or I mean, I think kids are fine if they understand the algorithm for adding and subtracting. They just need to understand it. (Sonia, Interview 9)

Sonia stressed that "concrete" ideas were necessary before pupils could understand "abstract" concepts. She continued to consider these issues in her second year of teaching as she thought about ways to engage pupils in higher-order thinking as they did lower-order tasks. As she made curricular decisions related to the books she used in class, Sonia said, "I was thinking…when do I ask just comprehension questions, like when do I ask higher level, like more analysis or you know just going through those levels [Bloom's Taxonomy] or just what do I need to have the kids do before I need to have them understand the story first" (Sonia, Interview 11). Sonia considered levels of complex thought and sought ways to balance reading comprehension with deeper analysis so that pupils could engage in higher-order thinking.

Some of the emphasis on balancing skills and authentic intellectual work for this group is due to grade level. It is significant to note that all four teachers in this category were prepared as elementary teachers. Riley, Rachel, and Sonia all taught elementary

during their first two years of teaching, and although Lola taught middle school science for her first two years, she spent the preservice period in an elementary classroom. Only one of the elementary teachers in this study was not placed in the "moderate" category, Sylvie, whose ideas about assessment placed her in the "low" category. The teachers in this category reconciled their limited emphasis on higher-order thinking in terms of what was developmentally appropriate for young pupils. They reasoned that elementary pupils needed basic skills before they could advance.

Here the interview data provides a deeper understanding of how elementary teachers considered authentic intellectual work that was not apparent in the scores from the external TAPL. There were no statistical differences in the quality of elementary teachers' assessments when compared to secondary teachers' assessments, although there were slight differences in pupil work scores between the two. However, based on the interviews, elementary teachers thought more about assessing pupils' basic knowledge. Riley constructed activities where her pupils were encouraged to formulate their own opinions because she felt they were able to do so. Describing a successful debate she had in class, she said, "So I think those conversations are good to have with [pupils] because I think fourth grade is kind of the age where they start to be able to think more critically about things" (Riley, Interview 11). She felt that fourth graders could begin to handle more critical thinking. Similarly, Sonia described how the fourth graders in her student teaching classroom were learning how to construct and support arguments with evidence. "I think they did well," she said, "and they've never done that before. And so they weren't very good at supporting their opinions with evidence, but they tried" (Sonia,

242

Interview 5). This was new to fourth graders. The belief here is that pupils need basic skills before they can move to more complex thinking.

This attempt to balance lower- and higher-order thinking suggests that elementary teachers felt that their pupils needed to learn basic knowledge and skills as a foundation before pupils could go further and they needed to provide this base knowledge. This does not necessarily contradict the principles of authentic intellectual work. Newmann (1996) acknowledges that one can "not expect all classroom activities to meet all three [authentic intellectual work] standards all of the time" and that "repetitive practice, retrieving information, and memorization of facts or rules may be necessary to build knowledge and skills as foundations for authentic performance, or to prepare for unauthentic tests required for advancement in the current educational system" (p. 27). However, Newmann emphasized that the goal is "to keep authentic achievement clearly in view as the ideal valued end" (p. 27). The previous examples showed that these teachers sometimes valued authentic achievement as an end goal but that was not always the case. These teachers struggled to balance authentic and inauthentic work.

Further demonstrating the influence of grade level on authentic intellectual work, Lola demonstrated a shift in the way she stressed concepts of authentic intellectual work from her elementary preservice experience to her middle school science classroom as she shifted her focus from skills to a more critical perspective. During her preservice period, when she taught in an elementary classroom, Lola described her desire to make sure pupils had basic math skills:

> [B]ecause I had been working with [pupils] all year and I knew where they
> typically fell out, in terms of how fast they caught on to things or their

243

computation because you have to do computation in all the other units too, this is just like focused on computation, this unit.… this is the one thing I feel all kids need to be able to do, you know? And it is more just skills, basic skills, and then you can fly once you have the basic skills, but you can't without it. (Lola, Interview 5)

Here she focused on basic math computation to serve as a foundation for future learning. This happened in other instances as well. Emphasizing lower-order thinking, she later said in the same interview, "I don't care if they can really talk about it if they can't do it yet, because they have nothing to talk…I mean it's good to know, but you need to be able to do it" (Lola, Interview 5). Lola did not see conceptual understanding as uppermost and considered basic skills as important building blocks that would help pupils engage in higher-order thinking at a later point in time. Lola focused at this point in time on the lower-order skills. She said,

> I think the take-away is that especially when, any information that's new like this and still forming in the brain, you need to keep doing it. And maybe that makes me old-fashioned but I feel like the more that you do these types of skills, the math skills, it will sink in. And, by not doing it for those three weeks they just started to lose it and if they'd had even a week of [practicing] religiously for a week, I'm sure they would have done much better. So the take-away is that you can't let up like that. (Lola, Interview 5)

She focused on having her students practice these skills so that they would remember these facts by using memorization and drills which represent inauthentic pedagogy.

However, by the second year of teaching, when she was teaching middle school science, Lola focused more on critical thinking skills, in part after attending a National Science Teachers conference. She began to think more about critical thinking more but felt that she was still constrained by the curriculum she was expected to cover in her class that focused on specific content knowledge. She said,

> …kids don't know how to critically think and science is a subject …where you should be learning to critically think. But it's really hard when there's certain content that we're supposed to cover and there are things – There are definitely critical thinking exercises that you can do or labs that you can have but what do you do when your kids don't even know enough information to be able to like I think there are legitimate times when they don't know enough or have enough background knowledge to be able to make a plan to address a critical question. (Lola, Interview 10)

Lola wanted to have student do more critical thinking but also found it difficult to do so because her pupils did not have enough background knowledge. This illustrates the necessity for certain skills before pupils can move to more complex thinking. As a result, Lola still stressed skills in her teaching that year, but she also incorporated more labs. She described an activity where pupils took the role of environmental engineers to look at topography and analyze and interpret maps. Her goals were for pupils "[t]o be able to ask good questions, to be able to be given a problem and not shut down immediately, to be able to critically think and give it your best shot" (Lola, Interview 11). She began to see teaching and learning as a more complex balance between content knowledge and higher-order thinking:

> I really, even though I think that you shouldn't teach to a test and that I am totally into these critical thinking skills and fostering that as a science teacher, I also feel there is a certain amount of curriculum that students should understand and that when they sit down for a subject matter test, they should, if I have done my job they should feel competent on most of that material. (Lola, Interview 11)

In an attempt to do this, Lola sought to integrate more critical thinking than lower-order recall. One of the assessments that she submitted during her second year of teaching was a lab, something that she had not submitted previously. It received the highest score on the rubric for authentic intellectual work out of all of her other assessments. She mentioned that she submitted the lab for that particular TAPL interview because in the

past she "had done mostly quizzes and tests" and that the lab illustrated that she had "branched out more" during that year (Lola, Interview 11).

Teachers whose understandings for assessment and pupil learning were "moderately" related to the concepts of authentic intellectual work balanced lower- and higher-order thinking. Their decisions were motivated, in part, by what they perceived to be the developmental needs of their elementary pupils to learn basic core knowledge and skills before they could engage in more complex thought. The illustrations here, combined with the examples from secondary teachers in the "high" group raise questions about the differences between how elementary and secondary teachers view and implement authentic intellectual work.

**Teachers Whose Goals Reflected Low Levels of Authentic Intellectual Work**

Finally, there were teachers in this study whose ideas about assessment and pupil learning had very few connections to the underlying principles of authentic intellectual work. Instead of talking about concepts such as construction of knowledge or elaborated written communication, these teachers' goals for assessment and pupil learning were simplistic. The teachers in this group were Craig, Mark, Matt, and Sylvie. Unlike the previous two categories, this group had teachers from different grade levels and subject areas. Craig taught middle school science, Mark taught secondary social studies and science, Matt taught secondary math, and Sylvie taught in an elementary resource room for English Language Learners.

Their goals for assessment lacked strong beliefs behind instructional choices. In some cases, their goal for an assessment was simply that the assessment was required or

listed as part of a prescribed curriculum. For instance, as described previously, Craig assigned pupils an activity where they listed basic facts about the different planets in the solar system on "trading cards" that they then used as a study guide to review for a test. When asked what he wanted pupils to get out of the solar system assignment, Craig replied that "one of the frameworks…was understanding the unique characteristics of the different planets" and "so that was kind of the target" for the assignment (Craig, Interview 8). Craig's assessment practices were guided by a curriculum binder created by the district to which he often adhered. Similarly, Mark also had a scripted curriculum when he taught secondary science. During his first year of teaching Mark accepted a teaching position in science, a field he was not licensed to teach. He commented that he liked the inquiry-based curriculum, not because it fostered higher-order thinking, but because it made his role as a teacher easier. He said, "There's a certain amount of uncertainty built into the curriculum so that sort of makes it easy for me to not exactly know things because the students are to discover the underlying principles involved in physics and chemistry" (Mark, Interview 7). Mark used construction of knowledge as a reason not to know all of the material that he should. It is ironic and disconcerting that Mark used an authentic practice to avoid being an expert in the discipline. Although the assessments may have been authentic, Mark did not have the necessary knowledge to engage with the material. His own work was very inauthentic.

There were instances where these teachers utilized authentic assessments but did not consider the concepts of authentic intellectual work as their learning goals. For example, Mark developed an assessment that he gave his pupils every Friday. Mark's

description of the assignment was aligned with authentic intellectual work because he said he wanted pupils "to be able to read something, take evidence from it and use it…to synthesize the information and use it for their own purposes" (Mark, Interview 5). Synthesizing information and using evidence to support a generalization aligns with the criteria construction of knowledge and elaborated written communication. However, Mark went on to explain that "the real motivation for continuing the essays" on Fridays was "a classroom management issue" because on Fridays his pupils "were just wild" and when he assigned the essay, the class wrote the essay and "was never so quiet" (Mark, Interview 5). Mark did not provide feedback on the essays back to his pupils and only checked the assignment off for completion. In this example, he assigned an authentic assessment, but did not think about the authentic value of the assignment. Again, this is a troubling practice and an ironic situation. On the surface it appeared that Mark was engaging pupils in authentic work, but the reality was very different. It is important to note that Mark's practices did not fit the pattern of other teachers in this study and were considered troubling.

Sylvie conducted her classroom similar to Mark and was also considered different than other teachers. For example, Sylvie's assessments had high external TAPL scores. One of her assessments was a field trip to the State House where pupils were asked to do research before they went on the field trip, write about their trip for the school newspaper, and present their articles to their classmates. Although the assessment is highly authentic, Sylvie did not talk about goals that were consistent with authentic intellectual work. Instead, her reasons for creating the assignment stemmed from what

was written in the curriculum she was given. For one thing, field trips were a part of her class, as a way to help English Language Learners become acclimated to American culture and institutions. For another, her decision to have students research the State House was due to the fact that "when it came time for the field trip they were ready to do research projects" (Sylvie, Interview 8). She had no particular inclination to develop this into a more authentic experience. More worrisome, Sylvie did not engage in teaching practices to engage students in an authentic experience with the work. She allowed pupils to directly copy from other sources. Describing one pupil's work on an assignment where Sylvie asked pupils to create a timeline, she explained:

> She copied it all. She copied it, but the point was she was able to pick out what…she understood the concept of a timeline. She understood the concepts that it only had the important things…chronological order. She understood that in a timeline you don't write the whole page of the whole paragraph. So I mean, yes, she copied it, but she understood what to copy. (Sylvie, Interview 9)

Although it could be argued that the pupil understood concepts related to a timeline as Sylvie described, Sylvie did not find it problematic that her pupils did not understand the material on their own. Copying was an accepted practice. Sylvie spent little time in the classroom working with students on their assignments, oftentimes choosing to sit at her desk while students worked on their assignments. Although her assignment tasks reflected authentic intellectual goals and received high external TAPL scores, in actuality, she did not follow through to make the students' interactions with the assessment authentic. Her practices here are troubling. The differences between her scores and thoughts about assessment make her an outlier and demonstrate the importance of using both the internal and external data to understand teacher and pupil

learning. Based on these examples, there are questions about the quality of Mark and Sylvie's teaching practices.

Some of the teachers in this group provided contradictory responses when asked about their practices. For example, Mark often talked about the importance of having students construct their own knowledge, a seemingly authentic goal. Describing one assignment he said,

> I want them to be able to look at this material and look at history, interpret it for themselves, decide what they think happened and what people's motivations were but the real object here is to get them, to sort of figure out what, what they think or what they want to be able to do, and to look at things like a newspaper article and interpret it and what's really going on here. Do I believe this is what really happened? Do I think there is a bias? (Mark, Interview 5)

In this excerpt, Mark stressed his desire for pupils to interpret information. However, he later described his goal for the same assignment in a very different way:

> I mean, if you look at it, it's easy stuff to get out of. I mean, I wasn't looking for, I almost wasn't looking for any interpretation of this stuff, I mean like really just find the information and regurgitate it. Like, ya know, make it, ya know and offer some assessment of it, that's it, not difficult, and they all did really well on that part. (Mark, Interview 5)

His first description is consistent with the criterion, construction of knowledge, but in his second statement, he admits that he was not looking for interpretation and instead was asking student to recall information. Looking at the assignments confirms the later description. Most of the questions on his assessments emphasized lower-order questions such as asking pupils to write essays where they "Describe the Populist Movement" or "Outline the changes in the United States during the industrial revolution;" his assessments did not engage pupils in higher-order thinking.

The four teachers who were placed "low" on the authentic intellectual work continuum did not talk about assessment and pupil learning in a manner that was consistent with the concepts of authentic intellectual work. It is interesting to note that two of these teachers – Craig and Matt – taught science and math, respectively, subjects that were shown in the external TAPL analysis to provide less opportunities for pupils to engage in authentic intellectual work. These two teachers had a very traditional and straightforward approach to teaching which focused primarily on content knowledge. They were considered, for the most part, adequate teachers in their schools. The other two teachers – Matt and Sylvie – had very questionable teaching practices and were not considered by their peers or administrators as good teachers. To get a better sense of how their understandings of assessment and pupil learning compared to the quality of their assessments and pupil learning, I now turn to an examination of this relationship.

**The Authentic Intellectual Quality of Teachers' Assessments/Pupil Work and**

**Teachers' Understandings of Assessment and Pupil Learning**

The previous section focused on teachers' ideas about assessment and pupil learning based on interview data. However, a more detailed interpretation of whether and how their views were reflected in their practice is necessary to understand the whole picture. Mixed methods analysis allows for a more detailed and complex perspective on a particular phenomenon. In this part of the study I compare the qualitative internal TAPL data described in this chapter with the external TAPL analysis described in Chapter 4. Comparing the results of the two analyses, I examine the following: How did the way that teachers talked about their goals for assessment and pupil learning compare to the quality

of their assessment tasks and pupil work as scored on the rubric for authentic intellectual work?

One way to look at this is to compare the two continua for the assessment task and pupil work scores on the external TAPL, as seen in the previous chapter, and the one continuum for their understandings of assessment and pupil learning from the internal TAPL analysis presented in this chapter. Looking at the three continua together provides an entry point into possible relationships between teachers' practices, as defined by the quality of their assessment tasks and pupil work, and their understandings of teaching and learning, as defined by the analysis of their interview data, in a very general sense.

Figure 5.2 shows the two continua of teachers based on the scores their assessments and pupil work received when they were evaluated with the rubric in the external TAPL at the top of the figure. The bottom of the figure shows the continuum of where teachers fell according to the extent to which their understandings and goals for learning were consistent with the concepts of the framework of authentic intellectual work as described in the internal TAPL analysis. In the bottom continuum, each band of teachers is listed alphabetically and not in hierarchical order.
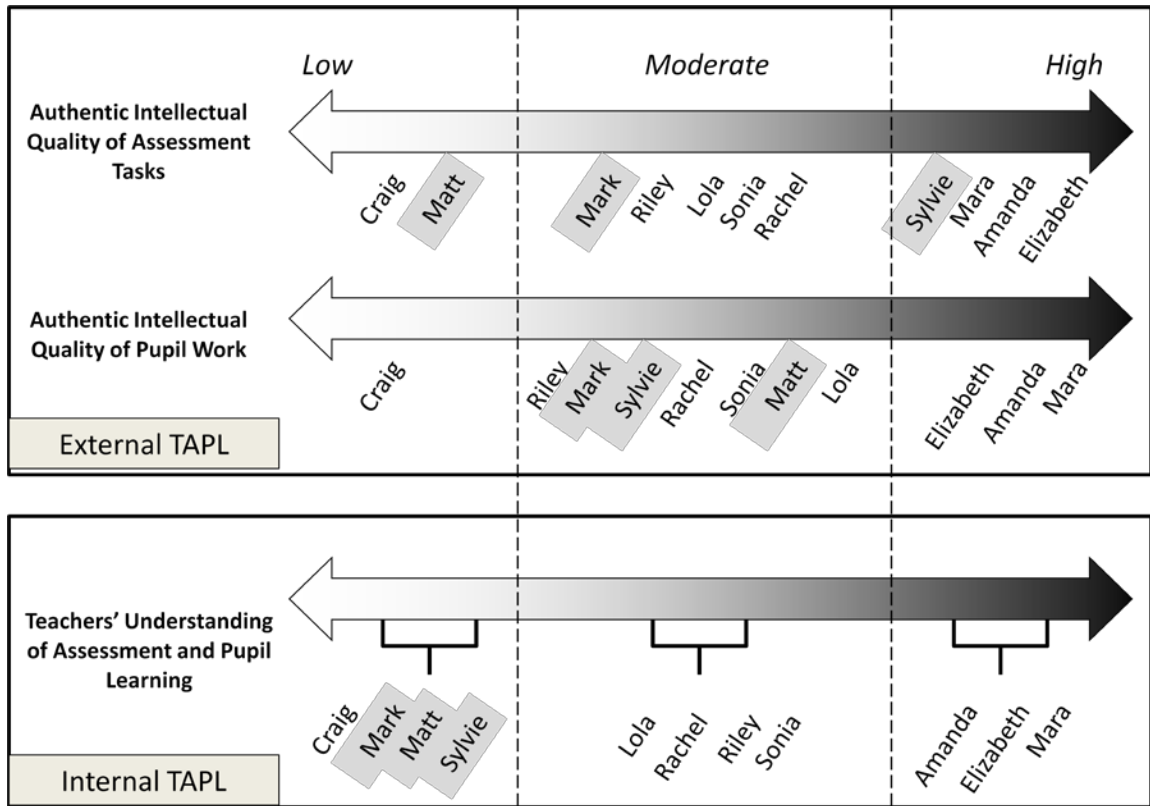
*Figure 5.2.* Continua for teachers' scores and thoughts related to authentic intellectual work.

As can be seen in Figure 5.2 above, the teachers in this study, for the most part, fell into the same band of the continua for both the external and internal TAPL analyses. That is to say, teachers whose goals for assessment and pupil learning reflected high levels of authentic intellectual work were more likely to have high quality assessment tasks and high pupil work scores as evaluated on the RISER rubric. Only three out of the eleven teachers (whose names are highlighted in Figure 5.2) did not match up in all three continua: Mark, Matt, and Sylvie. Although details about these three cases are provided in greater detail in what follows, a general explanation is offered here. For all three of these teachers, there were significant inconsistencies between practice and the way they

talked about teaching and learning. These teachers did not engage in practices that aligned with their ideas related to assessment and pupil learning. With the exception of these three teachers, the consistency of the other eight teachers in this study suggests that teachers' understandings of teaching and learning, in this case how understandings related to the framework of authentic intellectual work, were consistent with teachers' practices. This consistency strengthens the validity of the two separate analyses. This makes sense, in a general way, since it is reasonable to assume that teachers' goals for assessment should be reflected in the assessment that they utilize.

The teachers who had understandings that were highly consistent with authentic intellectual work in their interviews (Amanda, Elizabeth, and Mara), received high scores on their assessment tasks and pupil work. The three teachers who were categorized as "high" in the internal TAPL also had high scores in the external TAPL analysis. As Table 5.2 illustrates, The scores for their assessment tasks ranged from 6 to 10 (out of a 10-point scale) and the mean for all of their assessment tasks was 8.2 (n=15, SD=1.177). The external TAPL scores for their pupil work class mean ranged from 6.2 to 9 (out of a 12-point scale) and the mean for all of the pupil work scores was 7.59 ($n$=143, $SD$=1.907). These were the highest scores for the entire study.

Table 5.2

*External TAPL Scores for Teachers in the High Band on the Internal TAPL Continuum*

| Participant | Student Teaching | | Year 1 Fall | | Year 1 Spring | | Year 2 Fall | | Year 2 Spring | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Amanda | | | | | | | | | | |
| Assessment Task | 9 | | 8 | | 9 | | 8.5 | | 6 | |
| Pupil Work | 7.9 | 2.11 | 6.3 | 1.16 | 8.4 | 1.13 | 8.5 | 1.83 | 7.1 | 2.08 |
| Elizabeth | | | | | | | | | | |
| Assessment Task | 9 | | 8 | | 10 | | 8 | | 8 | |
| Pupil Work | 6.2 | 1.69 | 6.7 | 2.14 | 8.3 | 2.14 | 8.3 | 2.52 | 7.4 | 1.73 |
| Mara | | | | | | | | | | |
| Assessment Task | 7 | | 7 | | 9.5 | | 6.5 | | 9.5 | |
| Pupil Work | 7.4 | 1.51 | 7.2 | 1.89 | 8.5 | 2.07 | 6.9 | 0.82 | 9 | 1.64 |

*Note.* Pupil Work "class mean scores" are presented.

Similarly, the teachers whose goals and understandings of assessment and pupil learning were categorized as "moderate" in the internal TAPL analysis, had tasks and pupil work that received scores in the moderate range of the scale for authentic intellectual work on the external TAPL analysis. The teachers whose goals and aims for learning had moderate levels of authentic intellectual work (Lola, Rachel, Riley, and Sonia), also had tasks and work that were in the moderate range of scores. As Table 5.3 suggests, their assessment tasks received external TAPL scores that ranged from 5 to 10 (out of a possible 10) and the mean for all of their assessment scores was 6.16 ($n$=19, $SD$=1.546). Their class mean for their pupil work scores ranged from 4.7 to 8.75 (out of a possible 12) and the mean for all of their pupil work scores was 5.95 ($n$=179, $SD$=1.912).

Table 5.3

*External TAPL Scores for Teachers in the Moderate Band on the Internal TAPL Continuum*

| Participant | Student Teaching | | Year 1 Fall | | Year 1 Spring | | Year 2 Fall | | Year 2 Spring | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Lola | | | | | | | | | | |
| Assessment Task | 5 | | 6 | | 5 | | 9 | | 6 | |
| Pupil Work | 6.1 | 2.65 | 5.8 | 1.58 | 6.0 | 2.05 | 8.8 | 0.46 | 6.3 | 2.80 |
| Rachel | | | | | | | | | | |
| Assessment Task | 9 | | 5 | | No Assessment | | 6 | | 5 | |
| Pupil Work | 7.1 | 1.76 | 5.0 | 1.72 | No Pupil Work | | 5.4 | 0.94 | 5.9 | 0.32 |
| Riley | | | | | | | | | | |
| Assessment Task | 10 | | 5 | | 5 | | 5 | | 5 | |
| Pupil Work | 7.3 | 0.58 | 4.9 | 1.45 | 6.4 | 1.78 | 4.4 | 1.29 | 5.6 | 0.84 |
| Sonia | | | | | | | | | | |
| Assessment Task | 5.5 | | 6 | | 7 | | 6 | | 6.6 | |
| Pupil Work | 8.0 | 2.05 | 4.7 | 1.70 | 5.8 | 1.99 | 5.5 | 1.33 | 5.8 | 1.40 |

*Note.* Pupil Work "class mean scores" are presented.

It is significant to note that the majority of the assessment tasks submitted by three elementary teachers in this group (Rachel, Riley, and Sonia) were district-created standardized tests or district-mandated pre-packaged curriculum. Lola also submitted an assessment from standardized curriculum during her preservice period, although the assessments from her first two years of teaching when she was in a middle school science classroom were assessments that she created. Although both Riley and Sonia may have been influenced to submit math assessments due to an interest in math on the part of one of the QCS study researchers, the abundance of elementary math assessments reinforce the centrality of math instruction in the elementary classroom. This reflects the

accountability focus trend in Riley and Sonia's scores where elementary pupils were expected to pass statewide math assessments. Interview data supports the emphasis on math and literacy skills in their classrooms. Since the external scores represent, for the most part, assessments that these teachers did not create, the internal data provides a more nuanced understanding of how these teachers viewed assessment and pupil learning. As shown in the previous section, these teachers did, at times, talk about assessment and pupil learning in a more authentic way. These teachers did try to think about assessment and learning authentically. However, if these teachers had been evaluated on their assessments in the external TAPL alone, it would be a one-dimensional view.

To illustrate this further, I use an example of Lola's fifth grade math assessment, which she used with pupils while student teaching. This assessment was from a pre-packaged curriculum that she did not create. The assessment received a score of 5.0 (out of 10.0) on the external TAPL evaluation and was one of the lowest scoring assessments in this study. The assessment task was not a reflection on Lola's ability to create authentic learning opportunities because she did not create the assessment. In addition, the pupils' work score for this assessment was 6.0 (one of the lowest), a situation in which Lola also did have a role. From the TAPL interview data, a slightly different picture emerged. First, it was not surprising, but very apparent, that the exigencies of classroom life, over which student teachers have almost no control, played a strong role in pupil performance on the low-rated assessment task. Due to spring break and statewide testing, Lola's pupils had a three week break between the time of the unit and the final assessment used in the TAPL evaluation. Lola commented on the disappointment both

she and her pupils felt due to the dramatically lower performance on this assessment in

comparison to previous assignments. Lola felt this large gap in time between instruction

and assessment was a pivotal factor in pupils' performance. Second, we can gain valuable

insight into Lola's understandings of pupils' learning in her description of the pupils'

work. In her discussion of one example of pupils' work, for instance, Lola provided a

detailed explanation of the strengths and weaknesses of this pupil's mathematical

understandings (see Figure 5.3).

---
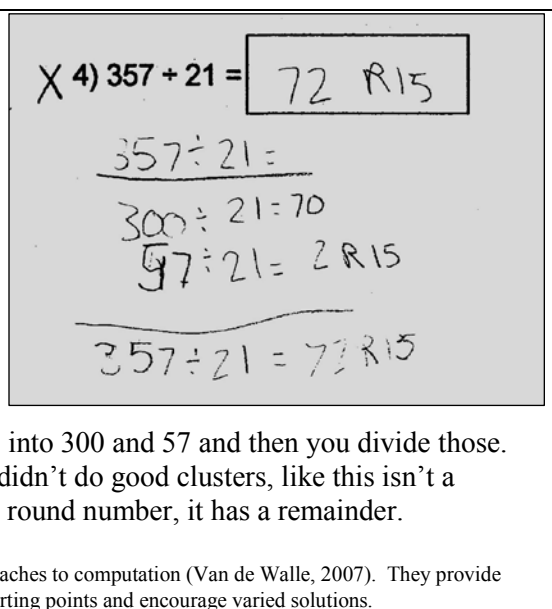
*Lola:* And then this is someone who got into the medium group. She was able to do the multiples and she could count backwards. She understands subtraction, although she got the wrong answer. Same thing with division. She kinda gets clusters* but not the right answer.…

*Interviewer:* Ok, so this problem is 357 divided by 21.

X 4) 357 ÷ 21 = [ 72 R15 ]

357 ÷ 21 =
300 ÷ 21 = 70
57 ÷ 21 = 2 R15

357 ÷ 21 = 72 R15

*Lola:* And so here she tried to break down the 357 into 300 and 57 and then you divide those. But the problem with what she did is that she just didn't do good clusters, like this isn't a friendly cluster. 300 divided by 21 doesn't equal a round number, it has a remainder.

\* Cluster problems are related problems that highlight flexible approaches to computation (Van de Walle, 2007). They provide students with opportunities to think about problems with different starting points and encourage varied solutions.

*Figure 5.3.* Example of how internal TAPL data informed external TAPL data.

Lola recognized and articulated that this pupil understood division but chose an

inappropriate method by using a number that could not be easily divided. Lola accurately

identified areas where pupils were struggling and pinpointed the nature of these struggles.

Further, Lola reflected upon her efforts to assist pupils in their learning by discussing her

instructional methods and the various successes she saw, especially on the quizzes she

administered during her instructional sequence. She commented, "And they really got it and they were excited for themselves, and they really, their own opinions of their ability to do math increased so much" (Lola, Interview 5). These insights into Lola's practice and understanding of pupil learning would be masked if only the external TAPL ratings had been used. This example also illustrates that there were contextual factors, such as the school district accountability context and the developmental level of their pupils, that influenced teachers' implementation of authentic intellectual work. These contextual factors are examined in greater detail in the next section.

Finally, teachers who were categorized as failing to think about assessment and pupil learning in an authentic way in the internal TAPL analysis, were not always consistent across all three continua (see Figure 5.2). As Table 5.4 illustrates, Craig was the only teacher who had the lowest scores on both assessment tasks and pupil work on the external TAPL analysis and was categorized as "low" in his interviews. Mark and Matt, however, had students who produced moderate levels of authentic intellectual work. Mark's assessment task scores were also considered "moderate" while Matt's assessment tasks were considered "low." Sylvie had "high" assessment task scores but "moderate" levels of pupil learning. The external TAPL scores for these teachers' assessment tasks ranged from 3 to 9.5 (out of a scale of 3 to 10) and the mean for all of their assessment tasks was 5.76 ($n$=19, $SD$=1.670). The external TAPL scores for their pupil work scores ranged from 3 to 11 (out of a scale of 3 to 12) and the mean for all of their pupil work scores was 5.49 ($n$=159, $SD$=1.869).

Table 5.4

*External TAPL Scores for Teachers in the Low Band on the Internal TAPL Continuum*

| Participant | Student Teaching | | Year 1 Fall | | Year 1 Spring | | Year 2 Fall | | Year 2 Spring | |
|---|---|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | M | SD | M | SD |
| Craig | | | | | | | | | | |
|    Assessment Task | 7 | | 4 | | 6 | | 4 | | 3 | |
|    Pupil Work | 4.75 | 1.16 | 3.1 | 0.21 | 5.7 | 2.62 | 5.1 | 2.08 | 4.45 | 1.62 |
| Mark | | | | | | | | | | |
|    Assessment Task | 6 | | 6 | | 7 | | 4 | | No Assessment | |
|    Pupil Work | 5.3 | 1.77 | 6.45 | 1.38 | 5.4 | 1.71 | 5.00 | 1.69 | No Pupil Work | |
| Matt | | | | | | | | | | |
|    Assessment Task | 5 | | 6 | | 4.5 | | 5 | | 4 | |
|    Pupil Work | 7.4 | 2.76 | 6.0 | 0 | 5.1 | 0.97 | 7.35 | 1.25 | 4.63 | 1.25 |
| Sylvie | | | | | | | | | | |
|    Assessment Task | 8.5 | | 6.5 | | 7 | | 9.5 | | 6.5 | |
|    Pupil Work | 6.2 | 1.92 | 5.06 | 1.21 | 5.25 | 1.16 | 6.15 | 1.53 | 6 | 0.71 |

*Note.* Pupil Work "class mean scores" are presented.

As suggested in the previous section there were specific factors that help explain the inconsistencies between quality of assessment tasks, quality of pupil work, and the degree to which Matt, Mark, and Sylvie's ideas coincided with the ideas that are central to  authentic intellectual work for Matt, Mark, and Sylvie. In the case of Matt, who had moderate pupil work scores but low assessment task scores and goals, one likely explanation is that Matt had some high-achieving pupils who were able to produce authentic work, even though the assessment tasks asked them to do only a little. A detailed explanation of this is presented in Matt's case study at the end of this chapter (see Figure 5.4). Matt's goals for assessments and the actual assessments that he created were actually consistent and were at low levels.

Mark and Sylvie, on the other hand, had assessment task and pupil work scores that were more authentic than their goals for teaching and learning. As the previous section demonstrated, Mark and Sylvie exhibited questionable teaching practices and were considered by the external researchers as ineffective teachers. They utilized assessments that were somewhat authentic, and in Sylvie's case highly authentic, but implemented the assessments in a very inauthentic manner. For example, as described previously, Mark gave pupils an authentic essay assessment on Fridays, but he utilized it to keep students quiet and did not even grade or respond to pupils' work. In his first year of teaching, Mark accepted a job teaching science, a subject for which he was not licensed to teach. He followed a scripted curriculum, which accounts for some of the moderate scores for the assessment tasks. In this class he engaged pupils in "inquiry," but said he liked this approach because he did not need to be as knowledge in the subject and pupils would learn the material on their own. This shows how his work appeared authentic but his beliefs did not support those practices. Similarly, Sylvie did not require that pupils follow through on the assessments. Classroom observations suggest that she often sat at her desk and let pupils attempt to do the work with little guidance and, as was already mentioned, allowed pupils to copy their answers from other sources. Pupils in their classes did not engage in authentic instruction.

When data from both the internal and external analyses are taken into account, it is apparent that teachers' ideas about assessment and learning, for the most part, were often related to their assessment practices in a consistent way. The analysis also suggests that there were contextual factors that influenced teachers' abilities to engage pupils in

authentic intellectual work. The next section examines these contextual factors in greater detail.

**Contexts and Conditions Related to Authentic Intellectual Work**

In making sense of teachers' assessment task and pupil work scores on the external TAPL and their understandings of assessment and pupil learning through the internal TAPL, it is clear that there were factors that aided or detracted from teachers' implementation of authentic intellectual work. This section examines factors from the qualitative data that influence how teachers considered authentic intellectual work. This differs from the factors discussed in the previous chapter where I looked at the impact of variables such as grade level, school context, and academic discipline on the quality of assessment task and pupil work scores. Those variables were ones that had been examined in other studies and could be measured quantitatively. In contrast, in this section of Chapter 5, I examine specific factors that influenced particular cases in this study and integrated the results from the external TAPL data as a way to further support what teachers said in their interviews. This analysis shows that teachers' abilities to attend to the concepts of authentic intellectual work were related to specific classroom contexts and their thoughts about learning that go beyond the factors analyzed in Chapter 4.

Based on an analysis of what teachers said about their classroom experiences and teaching and learning, there were five factors that seemed to influence teachers' implementation of authentic practices. These themes were: 1) time, 2) school culture/accountability, 3) student academic ability and motivation, 4) behavior/classroom

management, and 5) content area. This section of the chapter describes how these themes interacted with teachers' implementation of authentic intellectual work.

**Time**

One factor that influenced the degree to which teachers implemented authentic intellectual work was time. That is to say, the beginning teachers in this study made decisions about engaging pupils in authentic learning based on the amount of time it took to implement authentic assessments and teaching activities.

The teachers in this study demonstrated that creating and evaluating authentic assessments is time-intensive, something that they considered in choosing what assessments to use. The assessment tasks that received the highest external TAPL scores were ones that teachers took the time to create themselves, followed by assessments that they modified from others. Teachers who created assessments were more likely to engage pupils in more authentic work than teachers who used assessments that were created by others, and therefore were less time consuming to develop.

The teachers were cognizant of the amount of time it took to construct and evaluate authentic materials. For example, Sonia, in a discussion about more authentic math assignments, stated, "…some sort of those assessments that, yes, are very time consuming but are a little closer to assessing the real knowledge that children bring" (Sonia, Interview 9). She recognized that more meaningful assessments took longer to implement but that they were better assessments. The time factor played a part in some teachers' reluctance to undertake the time-consuming assignments. For example, Mark developed a project on the Great Depression during the student teaching period where

pupils used multiple sources to create a presentation, but said he would not do another

project like that one because it took too much time to plan. Describing the project, Mark

explained:

> [The project] actually went really, really well…but I would sort of step back from doing it again. I mean it was a lot of work for me, because I went, I probably spent four hours in the library gathering books for this thing, so it wasn't like I wouldn't do it every week, but I would definitely I'll definitely do it again, at some point. (Mark, Interview 5)

Mark did not consider the time that it took to put the assignment together as worth the

effort. In fact, the following year he accepted a teaching position where he was given a

scripted curriculum where he did not have to spend a lot of time creating assessments. He

used less authentic assessments because they were less time-consuming to create.

In addition to the time it took to develop authentic assessments, teachers also

considered the amount of time it would take to grade authentic work. Although she

wanted her pupils to do more formal writing, Elizabeth was hesitant to assign more of

this type of work, and acknowledged, "That's going to suck for me to grade" (Elizabeth,

Interview 9). Evaluating pupils' written work took up a lot of time. Similarly, Matt was

also very aware of how long it took to grade good assessments. In his first year of

teaching Matt explained, "…the better I make [assessments] the harder it's going to be

for me to grade. You know because truly you know the best assessments are tough to

grade because you can't just check, check, check, check, and that's what I did a lot of

times" (Matt, Interview 9). Most of his assignments were math problems where he could

just check off if students had the right answer. In fact, when he did implement a more

authentic assessment that incorporated writing, he stopped using it the following year

because he did not have time to grade all of the assignments. Some teachers were persuaded to use multiple choice assessments, a less authentic format, because it saved them time in grading, even if they disagreed with the fundamental idea of this type of assessment. "I also worry about …. I also don't think multiple choice tests are a good way of testing people," Mark stated at the beginning of his teacher education program. However, he then explained, "I'm more interested in short answer and essay questions. And then I look at it as that's a lot of work. That creates a lot of work for me" (Mark, Interview 1). It is not surprising then that both of the social studies assessments that Mark submitted, including the one used at the end of his second year of teaching, were comprised primarily of multiple choice questions and a few short answer/essays. Some teachers strayed away from more authentic assessments because they took too much time to grade.

The majority of the teachers who had the highest external TAPL scores, such as Amanda, Elizabeth, and Mara, were teachers who spent the most time at school, to the detriment of their personal lives. They all worked long hours – Amanda worked from 6:45 am to 10 pm every day – and were concerned about their lack of a personal life. These teachers began to streamline and reduce some of the authentic intellectual work that they were doing in their second year of teaching. For example, although Amanda still assigned writing journals in her second year of teaching, she chose not to evaluate as much as she did the first year. She explained:

> …I don't sit and read every single entry….usually I'll make a comment on two or three per notebook and then that'll be it….I'll make sure that they have …the entries but I might not sit and read every page because that, that was very time consuming….Last semester, like, you know, it would take me a whole weekend to

do a set of notebooks.  So now I just find that in the effort of time…I don't read the entire notebook. (Amanda, Interview 8)

She had a better defined sense of what she needed to look at and what she could spend less time on and developed a better balance between work and life. However, she did lower the amount of formative feedback that she provided to pupils and, in a sense, limited authentic instruction. Riley and Elizabeth had similar experiences and did things such as "put more emphasis on, like, making sure I have a social life even during the week sometimes or on the weekends" (Elizabeth, Interview 10). Both Amanda and Elizabeth, two teachers who had the highest external TAPL scores, had lower TAPL scores at the end of their second year of teaching; they utilized less authentic assessments at the end of their second year of teaching than they did during their first year of teaching.

Finally, in addition to the time it took to grade and implement authentic intellectual work, teachers felt their curriculum constrained them from being able to do as much authentic work as they would have liked to do. They frequently talked about how many other things they were expected to cover. Some of the teachers felt constrained by what was included in their school curriculum and that they might have incorporated more authentic learning if there was more room in the curriculum. Mark felt that in the district he taught, that the district curriculum and tests emphasized that "the kids just have to memorize stuff" and felt that students were just recalling information (Mark, Interview 10) for which he blamed the district's pacing guide. Teachers struggled with wanting to add more authentic work that was not related to the curriculum. Riley explained,

Well, it's kind of a mix.  We have certain curriculum but I mean you can sort of [alter the curriculum] when things come up. For example like you know we were reading *Time: For Kids* this week and it was about the [presidential] election and I

feel like you need to sort of seize some of those teachable moments. And so yeah you do have the discussion about voting or whatever. And yeah, that's not part of the fourth grade curriculum but should you not talk about it? You know? (Riley, Interview 10)

Riley wanted to incorporate important current events, like the presidential election, but the topic was not in the curriculum. Although, in this case, she did include the discussion, she was continually aware of what she needed to cover. Similarly, Matt stated that "time issues" prevented him from introducing literacy into his math classroom (Matt, Interview 10), and although Riley enjoyed doing a "really in-depth project" she felt "there isn't time do real over the top stuff" (Riley, Interview 10). As a result, she said, "We do little things" (Riley, Interview 10). Riley, and many of the other teachers, did not pursue all of the work they wanted to do because they had too much to cover in the curriculum. The curriculum limited the level of authentic intellectual work.

**School Culture and Accountability**

Teachers' abilities to engage in authentic intellectual work were also influenced by the larger context of the culture of the schools in which they taught, particularly related to issues of accountability. Some of the teachers in this study faced barriers as they attempted to implement authentic intellectual work.

They faced constraints from their school contexts, which made it difficult to implement higher levels of authentic intellectual work. Many of the beginning teachers in this study focused their assessments on preparing their pupils for the State Assessment tests that the pupils would have to take. Rachel's school, which was very concerned with preparing pupils for the test, emphasized a skills-based approach. Rachel described her math lessons as follows:

Yeah, I mean we're doing — We just started multiplication and they're starting to make flashcards tomorrow and I feel like something I've never been very good at is figuring out how to manage them memorizing facts and that's a big thing at the school is math fluency. And so I don't feel like I did a great job of it for addition and subtraction which is what I talked to a lot of parents about just practicing at home so I'm trying to be on top of it so that I can give them the materials that they need to practice at home but then also manage them practicing in the classroom because they need to know those. (Rachel, Interview 10)

She felt a lot of pressure to have pupils memorize math facts because her school emphasized those skills so that pupils would pass the statewide standardized math test. Although many of the teachers in this study disagreed with the format of the State Assessment, they recognized and resigned themselves to preparing their students for the assessment. "I don't necessarily agree with the State Assessment…," Rachel said, "but that's what it is right now and so it's my job too" (Rachel, Interview 10). The state assessment influenced the type of assessments she utilized at the expense of authentic intellectual work. All of her assessments were from a pre-packaged math curriculum that did not have the highest external TAPL scores. As Mark explained about the district-wide final assessment for his history students, "I mean I'm not a huge proponent, I'm not a huge advocate for multiple choice testing. I don't think multiple choice really tests whether kids know stuff, but they have to pass these tests. And so, on my exams I was giving them multiple choice tests" (Mark, Interview 5). Although the previous discussion showed that Mark's decisions were also influenced by time, he justified his assessment practices by relating his choices to the State Assessment. Standardized assessments informed teachers' assessment practices.

They also added elements to their assessments that reflected the State Assessment, often at the expense of authentic intellectual work. Teachers who had strong authentic

intellectual work practices moved towards incorporating elements of the state

standardized assessments into their own assessments. This example from Elizabeth

illustrates how her instruction changed to reflect aspects of the State Assessment:

> [L]ast week we did a lot of State Assessment multiple choice practice, short answer. They also, all 10th graders have a State Assessment prep class this term… so we've been kind of making sure that they're getting some practice. So I kind of hit them hard last week with a lot of State Assessment stuff and I think I'll try to do a little bit every day, just so they don't feel surprised or anything. It takes a lot of time but I didn't want a high number of my students having to retake it. I just – I want them all to be able to pass. (Elizabeth, Interview 8)

Elizabeth changed her instruction in order to prepare her students for the specific types of

questions that they would encounter on the State Assessment. The following year,

Elizabeth said almost the same thing:

> … we spent last week going over State Assessment. I gave them practice multiple choice and strategies and then open response. We haven't looked at long comp[osition] yet. But the writing certainly, definitely thinking about State Assessment. I want them all to pass. Like now we're, we just started *To Kill a Mockingbird* so, you know, how might I develop, like, sample open response questions as we're, as we're reading along. So, like, they're having writing practice and "okay, remember what were your State Assessment steps - use it to answer these questions." (Elizabeth, Interview 10)

All of the assessments Elizabeth submitted for the TAPL interviews were essays and she

stated a preference for more authentic assessments. However, she added less authentic

multiple choice questions during the period before the State Assessment her assessment

to reflect the types of questions that students were most likely to see on that test.

**Students' Academic Abilities and Motivation**

Another factor that influenced teachers' use of authentic lessons and assessments

was their pupils' academic abilities and motivation. Some teachers moved away from

authentic intellectual work because they thought their pupils did not have the necessary

academic skills to engage in higher-order thinking. In addition, high rates of pupil absenteeism and unwillingness to complete assignments limited pupils' production of authentic intellectual work.

Teachers moved away from authentic intellectual work when they thought that pupils lacked basic academic skills or content knowledge. For instance, Lola explained that there were "times when [pupils] don't know enough or have enough background knowledge to be able…to address a critical question" (Lola, Interview 10). As mentioned previously in this chapter, the elementary teachers in this study struggled to balance providing young students with a foundation of basic skills and knowledge with higher-order thinking skills. Some of the secondary teachers began to incorporate more basic skills instruction into their teaching, particularly in their second year, because they discovered that their pupils did not have the prior background knowledge or skills.

Amanda, Elizabeth, and Craig all talked about changes that they made to focus on skills. For instance, in her second year of teaching, Elizabeth focused on basic writing skills in her Advanced Placement (AP) English course, skills that pupils were expected to have as a prerequisite to take an AP course, because her pupils had very low writing skills. She explained, "I'm definitely going slower than I probably should.…I'm not going to rush through something when the students don't even know how to write a freaking paper – because they don't know how to – and they're juniors in an AP class" (Elizabeth, Interview 10). Elizabeth's frustration with her pupils' lack of basic skills is clear in this excerpt. Throughout her experiences at her urban school, Elizabeth criticized the school system for allowing pupils to lag behind academically, and felt that pupils in a

suburban school would have had the necessary skills to better engage with the content she felt pupils needed. Elizabeth, therefore, made a decision to sacrifice AP-level content, which one might assume to have higher levels of authentic intellectual work, in order to help pupils gain basic writing skills.

In another example, Craig focused on teaching his middle school science pupils study skills because he did not think his pupils knew how to study for a test. He said,

> Well, obviously, the content is a starting point, but I decided I need…to focus on study skills of various types. And this was something…one of the professors at [teacher preparation program] kept harping on is you can't just teach them the content. You have to teach them what you want them to do, not just what you want them to know…So you have to teach them how to take notes. You can't just say, "Take notes."…You can't just say, "Study for a test."…So I've been starting to do a lot more…when we have vocabulary, I'll get some flashcards…and then…a couple days before the test, we'll review things, put them all together, and that's their study packet to take home…So I've tried to get a little bit more explicit about what I expect them to be doing rather than just say study for the test. (Craig, Interview 11)

Craig made changes from his previous instruction. These were skills that pupils needed to do well in school and are important, although they do not necessarily represent authentic intellectual work.

The degree to which pupils produced authentic intellectual work was influenced by pupils' actions as well. Many teachers faced high levels of absenteeism. Elizabeth mentioned that 11 of her pupils were failing her class because of attendance issues; those pupils were not in class enough of the time to do the required work. Attendance was a problem throughout Elizabeth's high school. She described a biology teacher who had 19 students in one class fail. Elizabeth expressed her frustration:

> [W]hen you look at attendance records and then you look at suspensions and then you look at all the other factors and they you look at them telling you that they

only do a half hour homework a night, it's like, well how much of this is my fault? Maybe I am not making the material very accessible but they're not meeting me at any certain point to say, "I am not getting this," until it's too late. (Elizabeth, Interview 8)

In this excerpt Elizabeth raises a point about pupil responsibility and legitimately questions how much she as a teacher can really do if pupils do not come to class and do not do their work. Getting pupils to do their work was a challenge for many of the teachers in this study. Elizabeth described instances where more than half of her class did not hand in their assignments. Matt stated that only about 20% of his pupils did their homework and often described his pupils as "lazy." Similarly, Mark also had pupils who refused to do work in class. For instance, Mark described the packaged curriculum that he used in his secondary science class as "active physics" that was designed to have pupils do classroom activities. He was unable to get pupils to do the work. He said, "[K]ids don't even do it….They're not even pretending. They're just like, 'No, I'm not doing anything'" (Mark, Interview 8). Mark became frustrated with his pupils and said he "[didn't] know what to do with them" (Mark, Interview 8). These teachers struggled with getting pupils to do work.

The pupil work analyzed in this study support these teachers' statements. Many samples were incomplete where pupils did not even attempt to answer questions. In a few instances, there were only a few samples of work from a class because only a few of the pupils handed in the assignment. This impacted pupil work scores on the external TAPL. The RISER rubric takes into account pupil completion of the task in the standard, analysis, by considering the extent to which pupils engage in analysis throughout the entire assessment. Most likely, pupils who do not complete the assessment would have a

lower score on that standard than pupils who completed the assessment. Since the other two standards, disciplinary concepts and elaborated written communication, only require raters to score the portion of work that pupils actually complete, there is a chance that pupils who do not complete the assessment will not have the lowest score. However, it is reasonable to assume that pupils who were not in class most likely did not gain a deep enough understanding of disciplinary concepts to score very high on assessments. When pupils are not in school or do not do their work, the quality of assessments is not a factor in their performance because they were not present for instruction.

It is unclear from data in this study whether pupil engagement was related to pupils' level of productivity. That is to say, it was unclear why pupils disengaged from school and whether the level of authentic intellectual work was related to pupils' disengagement. However, there was evidence to suggest that attendance and getting pupils to do work was a larger, school-wide problem at some of the schools where these teachers taught.

**Behavior and Classroom Management**

Teachers' abilities to implement authentic intellectual work were also influenced by how well they were able to manage their classroom and students' behavior. A few of the beginning teachers in this study described instances where they refrained from doing more authentic hands-on activities because their students misbehaved. Teachers described instances where students destroyed materials for hands-on activities or engaged in such disruptive behavior that they were unable to complete the assessment.

As a beginning elementary teacher, Rachel struggled with classroom

management. There were times when she wanted to do math activities with manipulatives

but found that her class was disruptive and did not pay attention. She described one

attempt from the fall of her first year of teaching as follows:

> This [lesson] has a lot of activities that involve sitting in a circle and counting
> round the class actually.  And I barely did any of them because it's sort of
> management.  Like I've just learned there's a lot of stuff in [Math Curriculum]
> that if I could get my kids to listen and cooperate a little more it would be helpful
> but some of them I just know aren't going to work and so to try it would be a
> waste of time.  So the one day, I think twice I actually got them to like do these
> kinds of activities and both of those were days when I only had like 15 kids there.
> (Rachel, Interview 8)

Rachel had trouble managing her students and was only able to do activities that were

more engaging when some of her students were out of the classroom. She decided not to

try to do certain activities because she knew her students would not cooperate. However,

as she progressed in her first year of teaching, her ability to manage her class improved

and  she was able to do more. Toward the end of her first year she said,

> … one of the better things we've done in math in a while…the first lesson of the
> unit was supposed to be each [student] with a container of the shapes and they're
> sorting them out and figuring out what's the same and what's different. But first
> of all, my kids with manipulatives and stuff like that, the spheres would've been
> rolling and they would've been, like –  I just at this point I don't even go there. I
> know it's a bad idea.  But they were really engaged on the carpet, most of them,
> some of them are just, at this point just don't even try anymore.  But a good
> majority of the class was, like, really figuring out, like, what the difference was
> between the ones that I was holding up and learning those new words and then
> they got really invested in thinking about where they see those shapes in the real
> world.  Just the way they were working together, which they usually don't do very
> well.  I felt like that, for the first, I mean, I'd say the first half hour of math, that
> lesson on the carpet then seeing them work together, even some of the lower ones
> that usually, lower ones or the ones that act really disinterested in math…it was
> just nice to see them really invested in that. (Rachel, Interview 9)

Once her pupils were better behaved, Rachel could do more hands-on activities.

Similarly, the three science teachers mentioned that there were times that they did not do labs because of their students' behavior. As a science teacher, Mark had difficulty managing his class. He blamed his students for his instructional choices:

> I've, like, altered the curriculum significantly and I'm not doing any of the cool little physics projects that are in the book, we're not doing any of that. It's like they sit, they write and that's it. No group work. They have to work quietly. It's terrible. …I mean I can't even, literally, if I turn my back the room is, like, crazy. It's not even like, like, you know, kids talking. It's like, you know, kids are falling on the floor, like, stuff's getting broken, stuff is disappearing. …I don't think we're going to do any of those, not one of those experiments. Maybe, I've done a couple demonstrations and that's gone well. I may try to, like, keep doing that. But as far as them doing, it'll never happen. (Mark, Interview 8)

As a result of his inability to manage his class, Mark stopped doing demonstrations that would provide students with a better understanding of science concepts and he decided not to let the students do labs on their own. This severely limited students' opportunities to engage in authentic work in science by doing actual science experiments and partake in the scientific process.

Lola had similar problems with her middle school students. Although she did a few "mini-lab type of demonstrations," she focused primarily on "a lot of book work and notes" because she had not "figured out a way to teach [her students] with the materials and have them a) learn, and to b) not be social, and c) not throw stuff across the room" (Lola, Interview 7). She found that her students were not able to manage the materials and were destructive. As she attempted to implement authentic hands-on experiences, Lola had difficultly effectively managing student behavior and learning. During her first year of teaching, her beliefs related to authentic science instruction shifted as a result of her experiences:

And I've learned, I believe a lot still in hands on science but I also learned that it's sometimes easier said than done to actually have learning come out of the hands on stuff. Like you really have to manage it well and I'm definitely still working on that. (Lola, Interview 9)

Managing her class for an effective lab was central to doing labs for Lola. As she gained better classroom management skills in her second year of teaching, she was able to do more. One of the assessments that she used in that year was a lab assessment, described in a previous section, and was the highest scoring assessment that she submitted throughout this study.

This shows how important classroom management is. When teachers did not have control, they were more likely to engage in less authentic work. As new teachers became better managers, they had more opportunities to implement more authentic work.

## Content Area

Some of the beginning teachers in this study expressed particular beliefs about the content area they taught and how the subject matter contributed to their ability to address concepts related to authentic intellectual work. These teachers tended to see social studies and writing as subjects that supported authentic work unlike math and science that was more inauthentic.

Riley, Sonia and Sonia felt there were more opportunities for engagement and hands-on learning in social studies than math. The elementary math curriculum that Riley and Sonia were required to teach was prescriptive and did not allow for a lot of creativity in how to teach the materials. As a high school math teacher, Matt did not want his pupils to engage in "creative" thinking in math and focused primarily on having student recall information.

The three science teachers in this study appeared to be the most concerned about their subject area and pupil performance. They felt that science was not taken seriously in their schools. Mark was hired as a high school science teacher even though he was not certified to teach the subject. Mark saw science as a static subject and not as engaging as history, the subject he was certified to teach. He explained, "That you can't [debate] and I mean obviously that's one of the fun things about teaching History. Chemistry obviously like a mole is 6.02 X 10 to the $23^{rd}$ molecules. That's it. There's not debate" (Mark, Interview 9). It is not surprising that Mark did very few hands-on and authentic activities in his science classroom. Craig's middle school science class was not considered a "core subject." His classes only met four days a week while classes for the other subjects met five days a week. He traveled between two different schools and one of his classes did not even meet in a laboratory room so they were unable to do hands-on labs. Craig wanted to take his pupils on a field trip but was told he was not allowed to do so because science was not considered a main subject. Lola explained that science was often seen as a very lower order subject: "I don't think that in a lot of schools especially in the cities, science and especially earth science they're taught like textbook courses…With very little critical thinking involved and a lot of memorization. And kids are used to doing well in science by doing not much of anything."

As the external TAPL demonstrated, the quality of social studies and writing assessments and pupil work was more authentic than work in science and math. I suggested in Chapter 4 that some of the differences in quality between the disciplines may be related to the ways in which the framework of authentic intellectual work appears

to privilege writing and therefore disciplines that involve a lot of writing – such as social studies and English – would be more likely to score higher on the rubric. However, teachers' thoughts about the type of work that they do in these subjects may suggest that differences could also be related to the ways that these subjects have traditionally been taught in schools.

As can be seen time, school culture/accountability, student abilities/motivation, behavior/classroom management, and content area influenced teachers' willingness or ability to utilize some authentic assessments more than other teachers. These factors influenced teachers in different ways. This is clear by examining two different cases. These two cases show how these factors interacted with teachers' ideas about assessment and pupil learning and influenced the amount of authentic intellectual work that teachers used in their classrooms.

### Teaching for Authentic Intellectual Work: Matt and Mara

The degree to which the beginning teachers in this study were able to implement authentic intellectual work in their classrooms was influenced by the ways teachers' beliefs about assessment and pupil learning interacted with particular contextual factors that they faced in their classrooms. Teachers whose understandings of assessment and pupil learning were consistent with the concepts underpinning the framework of authentic intellectual work sought to emphasize authentic learning. In contrast, teachers who were less concerned with authentic intellectual work were less likely to utilize assessments that emphasized these concepts. Drawing on both the quantitative external and qualitative internal TAPL analyses, I present two case studies to underscore how all of these features

came together in two different classrooms. First, I present the case of Matt, a secondary

math teacher. Matt's assessment tasks received some of the lowest scores in the external

TAPL evaluation and were not considered to be authentic. His pupils performed slightly

better, and produced moderate levels of authentic intellectual work when their work was

evaluated with the RISER rubric. In addition, when Matt spoke about his assessment

goals and pupil learning, he rarely mentioned concepts related to the framework of

authentic intellectual work. At the other end of the continuum is the case of Mara. Mara,

a secondary social studies teacher whose assessments and pupil work were rated very

highly in the external TAPL, expressed a number of concepts related to the framework

when she talked about teaching and learning. These two case studies illustrate how

teachers' beliefs informed their practices and how teachers were constrained by or sought

to go beyond contextual factors that influenced their implementation of authentic

intellectual work. As a result, only one of these teachers attended to ideas that fostered

authentic intellectual work, and, in turn, democratic education.

Both Matt and Mara were highly qualified teachers according to the federal

mandate. They both had undergraduate degrees in the subjects they taught, they

completed the same teacher education program, and acquired teaching positions in

prestigious, high-achieving high schools. Matt taught secondary math at a highly-

selective, urban "exam" school, where students were required to pass an entrance exam to

be admitted to the school, and Mara taught secondary social studies in a large, affluent,

suburban school, one of the more academically successful schools in the region. Neither

Matt nor Mara faced pressures related to the state assessment exams because students in

their schools typically did well on these. Both had a certain degree of freedom to create their own instructional activities around a general curriculum framework.

Although their backgrounds and teaching contexts were similar, Matt and Mara had very different pedagogical practices. Matt relied on traditional math exams to assess pupil learning whereas Mara utilized a variety of innovative assessment practices including PowerPoint presentations and community action projects. In terms of the way these teachers talked about goals for assessment and pupil learning, their responses were very different. As shown in Figure 5.1 at the beginning of this chapter, Matt was placed in the "low" category and did not refer to concepts related to authentic intellectual work. Mara, on the other hand, was placed in the "high" category because she consistently referred to authentic ideas.

There were also differences in the authentic intellectual quality of their assessments and pupils' learning as measured by the RISER rubric in the external TAPL analysis. As Table 5.5 demonstrates, Matt and Mara had scores that were statistically different from each other ($p<.05$). The mean for Matt's five assessment tasks was 4.9 ($SD=.742$), while the mean for Mara's five assessment tasks was 7.9 ($SD=1.475$). Mara routinely utilized highly authentic assessment tasks, while Matt had some of the lowest scores in the study. Similarly, the mean for all of the samples of pupil work that Matt submitted for this study was 6.35 ($n=37$, $SD=2.003$), while the mean for Mara's pupil work samples was 7.77 ($n=49$, $SD=1.771$). Mara utilized more authentic assessments and her pupils produced higher levels of authentic intellectual work than Matt's students. This supports the correlation between learning opportunities and pupil learning that was

presented in Chapter 4. Assessment tasks that had higher scores were more likely to result in higher pupil work scores.

Table 5.5.

*Mean External TAPL Scores for Matt and Mara*

|  | M | n | SD |
|---|---|---|---|
| Matt |  |  |  |
|     Assessment Tasks | 4.9 | 5 | .742 |
|     Pupil Work | 6.35 | 37 | 2.003 |
| Mara |  |  |  |
|     Assessment Tasks | 7.9 | 5 | 1.475 |
|     Pupil Work | 7.77 | 49 | 1.771 |

*Note.* An independent samples *t* test showed that the difference between Matt and Mara's Assessment Task scores was statistically different, $t(8)=4.064$, $p=.004$.
An independent samples *t* test showed that the difference between Matt and Mara's Pupil Work scores was statistically different, $t(84)=3.464$, $p=.001$.

The differences in the way in which Matt's and Mara's assessment tasks and pupil work were scored according to the rubric in the external TAPL can be explained, in part, by an analysis of how their goals for their pupils' learning reflected the concepts of authentic intellectual work and how they interacted with classroom contextual factors. Matt did not make authentic intellectual work a goal in his classroom, whereas Mara made that type of work a priority.

**Matt**

Matt took a very traditional approach to teaching mathematics. As a student, Matt excelled in math and, as a teacher, his pedagogy emulated teachers from his childhood whom he considered to be the most successful. In his instruction he lectured, gave students practice problems in class, and checked to see if students completed their homework assignments. His assessments reflected this approach. Four out of five of these

assessments he submitted for this study were math tests, with questions copied directly from a textbook or that were created collectively by the department faculty at his school. The only assignment that differed from the tests was a portfolio where students were asked to write a few paragraphs about their experiences in math throughout the semester and keep a collection of their major semester assignments so that they would have materials to study from for future exams. Other homework assignments came directly from text- or work-books and he rarely introduced any other type of alternative assessment or project.

Matt did not think highly of his experience in the teacher education program. He thought that the program placed too much emphasis on theory, at the expense of practice, and questioned the value of writing papers related to teaching. He wanted the program to provide more direct answers about practical and specific things to do in the classroom. Describing his math methods course, he explained,

> Well for methods I just -- I think would have been so much more helpful to actually teach lessons and to actually talk about -- You know like take a chapter of an Algebra book and say how would you set up your lesson plans for this? How would you set up your unit? How would you teach these units? What would you do if the kids came in the next day and totally didn't understand what you had taught them the day before?...Like those things. And we talked nothing about that stuff. (Matt, Interview 8)

He did not try to make connections between theory and practice and incorporated very little from the program into his teaching. He wanted the teacher preparation program to help him use the textbook, which is what became his central resource for developing assessments.

Matt's experience in the teacher education program did not alter his views related to teaching math. He had a very straightforward view of math and placed boundaries on the type of thinking that he wanted students to engage in. His beliefs were inconsistent with the fundamentals of the authentic intellectual work criterion construction of knowledge. For example, although he stated that math involved higher-order thinking, he thought there were some types of thinking that did not fit within the math classroom:

> There aren't a lot of crazy, crazy things you can do  because there's some pretty concrete things you have to get across.  There's not a, there's not a lot of, I don't wanna say high level thinking, because there is a lot of high level thinking. There's not a lot of…creative thinking because with math, there is one answer that you get to.  There's not like there's a couple different opinions you could have.  Two plus two is four… (Matt, Interview 2)

For Matt, there were certain ways that students should think about math. Matt's resistance to "creative" thinking in math remained constant throughout his first year of teaching and he continued to believe in the importance of providing structure to pupils' thinking. Describing his actions in the classroom he said,

> You know Math – it's hard to be creative.…. Sometimes I try [not to] stifle a little bit of creativity in the Math classroom because it's nice for students to try and make things their own.  And I'd love them to find ways to memorize or remember theorems or formulas or whatever.  But somewhere along the line … the pedagogy became let's let kids try and work things out on their own.  And I am not a subscriber to that at all.  I think that it -- I think it frustrates them more than having something that they have to do that's long and drawn out. (Matt, Interview 8)

Matt preferred direct instruction over inquiry-based learning. His learning goals focused on how well students knew content and their ability to use formulas to solve math equations. He did not stress broader conceptual understandings. For example, in a manner that is typical of his descriptions of other assessments, he described the goals for one of

his assignments as follows: "For quadrilaterals, I wanted them to be able to understand what each of the ones were parallelograms, rectangles, rhombus, square, trapezoid. I want them to understand the difference between interior and exterior angles because we use them for different things" (Matt, Interview 10). By focusing primarily on the properties of geometric shapes, his goals consisted of lower-order recall and comprehension. His goals did not go beyond concrete and discrete pieces of information.

Ironically, Matt's students were actually able to do more authentic intellectual work than Matt asked them to. Matt was the only teacher in this study whose pupils had higher pupil work scores on the external TAPL than his scores on the assessment tasks. As the correlation in Chapter 4 illustrated, when all of the assessment tasks and pupil work scores in this study were correlated, there was a direct positive relationship between the quality of the assessment task and pupil work. That is to say, pupils who were assigned assessments that were scored higher produced higher levels of authentic intellectual work than pupils who were assigned assessments that had lower scores. In Matt's case, however, his pupil work scores were higher than the scores on his assessment tasks. One possible explanation is that the students in his school had strong academic backgrounds and were able to demonstrate higher levels of understanding in spite of the assessments they were given. For example, Figure 5.4 shows one pupil's response on a geometry exam question. The assessment, as a whole, was scored a 5.0 (out of 10). Questions on this exam had only moderate expectations for construction of knowledge or elaborated written communication, and they had no connection to the world outside of the classroom. However, this sample of pupil work was scored a 9.0 (out

of 12). As can be seen from the figure, this pupil correctly solved each part of the

problem, drawing a graph to help explain her work. The question asked pupils to "justify

your answer" and this pupil did this by writing sentences to explain her thought process.

This pupil demonstrated mathematical analysis, disciplinary concepts, and elaborated

written communication in this answer.
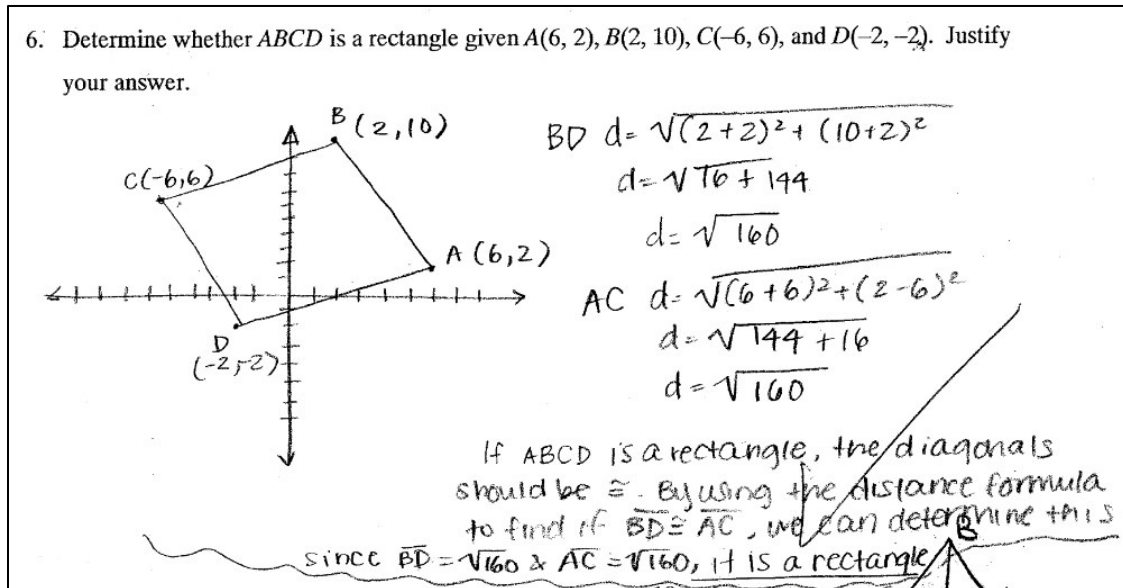


*Figure 5.4.* Example of high scoring pupil work on low scoring assessment task.


Ironically, Matt did not think his pupils were high-achieving and tended to blame

them for their inability to engage in what he considered to be more authentic work. Matt

described the pupils as not being able to engage in critical thinking. When asked about if

he helped pupils to critically think, Matt replied, "I do my best and they fight it. You

know they're all about what can I memorize and what can I plug in" (Matt, Interview 8).

His assessments, however, did not encourage critical thinking and often asked students to

"plug in" numbers and formulas. For example, in the assessment he used at the end of his

second year of teaching, an assessment Matt helped to design, pupils were asked to apply

formulas. Figure 5.5 illustrates one of the questions. To answer this question about the

angles of a triangle, pupils had to put the correct numbers into the formula for an

isosceles triangle. Therefore, pupils had to memorize that all the angles in a triangle add

up to 180° and that they should subtract the two angles from 180 to find the third angle.

In this case they had to recognize that angles E and D were the same because it was an

isosceles triangle and then subtract 122 from 180 and divide by 2. Pupils simply had to

plug in the formula. Matt did not challenge his pupils to think critically in his assessments

and pupils, therefore, had no motivation to engage in critical thinking.



Use the information in the figure below to find $m\angle D$.

[A] 61°     [B] 28°     [C] 151°   [D] 29°
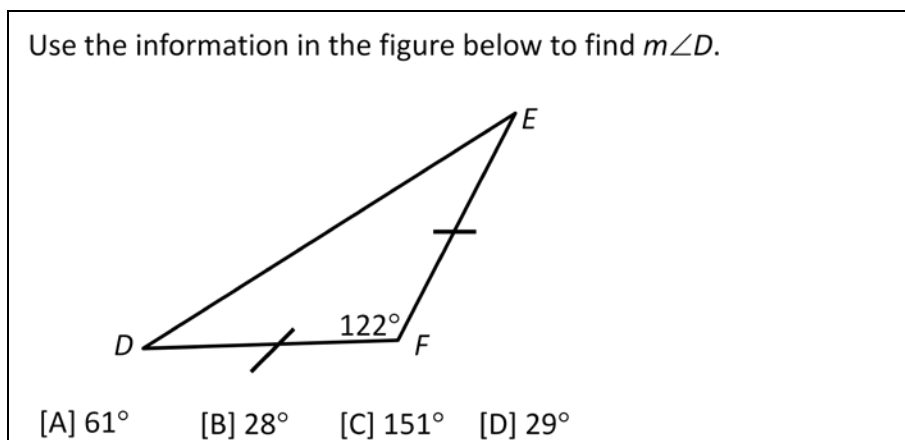
*Figure 5.5*. Isosceles triangle assessment question.

Matt did not connect his own assessment practices to his pupils' abilities to think

critically. Instead, he saw pupils' lack of critical thinking as a school-wide problem:

> You know I've had conversations about this with fellow teachers in the Math and
> teachers in other disciplines and critical thinking is I just think at an all-time low.
> And we agree … I sat with one of the Chemistry teachers one day and we were
> just talking and saying truly their critical thinking skills just are so below average

and when we try to do proofs in the class they fight it so badly.  You know, they don't want to -- They don't want to try to make a plan and follow the plan and see what they can do.  So it's tough. (Matt, Interview 8)

He continued to blame the lack of critical thinking on his pupils and considered this as a reason that pupils were not able to engage in authentic intellectual work. This can be seen by Matt's comments about the work of one of his pupils on a geometry test. The test question, shown in Figure 5.6, was a word problem that related geometry to a real world situation, one of the only times where Matt made connections beyond the classroom. Pupils were asked to determine how far fans were standing away from the stage at a concert. Matt commented that his students did not want to think about a word problem:

> You know and we did this problem in class so many times but so many kids skipped it.  And they were just you know they looked and they said oh, you know four lines of words.  I can't read all that.  I can't do this… You know and they can't pull that information out of here and draw themselves a picture.  I tell them all the time, I say draw a picture. (Matt, Interview 8)

Matt attributed his pupils' lack of authentic work on their motivations. He believed that he was incorporating things like critical thinking into his assessments but that it was the pupils' fault that they were unable to do well on the assessment. He frequently blamed pupil performance on not studying. "I truly don't believe they study at home", he said, "like 75% of them – I don't think they study at home" (Matt, Interview 8). He distanced himself from his pupils' learning.

---

7. Tyson is standing at the edge of a 15-foot stage for a concert. He is six feet tall and can see his two biggest fans in the crowd. One of them was lucky and got into the front row. His line of sight to her has an angle of depression of 75. The other fan got stuck in the back row and Tyson's line of sight to her has an angle of depression of 30. How far apart are these two fans to the nearest foot?

---

*Figure 5.6*. Geometry word problem.

Matt did, however, begin to take some responsibility for critical thinking in his second year of teaching. He realized that he was limiting pupil experiences:

> I don't think we do much critical thinking here at all. I think it's, especially the math classrooms. I think it's very, learn the theorem, copy down the theorem, memorize the theorem, regurgitate it for a test and forget about it. … I'm guilty a little bit of that because I don't necessarily teach a lot of lessons that are really critical thinking. A lot of times when I get the problems, especially with my 9th grade Algebra 2 class this year, whenever we would get to a word problem that involved a couple extra steps, I skipped them because it literally took an entire 45 minute class to do one of those problems. They were, they were nearly helpless when it came to trying to pull the math information out of the problem and then realize what to do with it. (Matt, Interview 11)

Yet, Matt used time as a reason to limit authentic intellectual work. Word problems were too time-consuming and so he moved away from introducing those types of problems in order to cover additional content. He did not consider certain authentic learning experiences as worth the time to do in class.

The time that it took to plan, implement, and evaluate authentic intellectual work was a factor that influenced the type of assessments that Matt chose to use. Matt considered balancing his personal and work life as extremely important and always tried to make time for his personal life, even at the expense of work. He chose to stop work in the evenings in order to play and coach basketball. Matt struggled with planning during student teaching and the first two years of teaching, frequently falling behind. Part of this stemmed from the conscious decision he made to create a balance between his work and life. He recognized that his decision to "have a life" put him in a "bad position when it came to grading and that he "need[ed] to find balance" (Matt, Interview 9). He fell behind in grading and sought assessments that involved less work. Describing his work in the first year of teaching Matt explained,

I'd really just, you know, my main thing was go through each chapter, go through each section, you know, I'd give them notes about it. I'd give them examples and I'd let them practice. I think, I think probably the biggest thing … was I would let them practice and a lot of times let them practice with their group members because I just felt like some of this was probably a little bit of convenience and it made it easier for me that I didn't have to make specific lessons and teach specific things to each person. (Matt, Interview 9)

He chose not to differentiate instruction because it was easier and less time-consuming not to. He made decisions based on how much time he had. At the end of his first year he reflected that he had wanted to "develop projects and stuff like that" but found he "didn't have the time for it" the first year (Matt, Interview 9). He did not incorporate additional projects in his second year of teaching either. The assessment that received his highest TAPL score, a written portfolio, was one that he discontinued because it took too long to grade. The portfolio was not that authentic and "it was mostly just to have them understand that it is important to keep up with stuff in your binder" (Matt, Interview 7). Matt did not consider it as a way to explore concepts in an elaborated way.

During his first year of teaching Matt recognized that there were areas where he could improve his teaching in ways related to authentic intellectual work. He had goals to incorporate technology and described an interest in taking students on a field trip to a baseball park to do math problems related to the construction of the field. However, he resisted change. By the end of his second year, he still had not learned technology that other math teachers used to teach geometry and he stopped utilizing the portfolio assessment because they took too long to grade. At the end of the second year, he still described his goal as to "try and bring more real world example[s]" into his classroom (Matt, Interview 11). However, Matt continued to struggle with making connections with

what students were learning in class to the real world. Content in his assessments remained decontextualized. Instead, he emphasized that students needed to learn the material because they were "going to use it again next year" in school and on assessments like the SATs and the State Assessment (Matt, Interview 8). His rationale for pupil learning was based on school academic achievement and not life-based and pupils were less likely to make connections to see math concepts in real situations.

Despite the shortcomings described here, Matt was considered to be a successful teacher. Interviews with supervisors and mentors indicate that he had strong content knowledge and conducted himself in a professional manner. His pupils passed the State Assessment. However, as this analysis illustrates, he was not a strong teacher when it came to engaging students in authentic intellectual work. He did not encourage pupils to critically evaluate and interpret significant math concepts in an elaborated way, and rarely did he present math in a way that emphasized how concepts might be used in professional disciplines beyond the classroom. To illustrate how his practices compared to a teacher with strong authentic intellectual beliefs, I turn to the case of Mara.

**Mara**

In contrast to Matt, Mara continually considered ways to engage her students in authentic intellectual work. She used a variety of lessons and assessments that challenged students to engage in critically analyzing contemporary and historical issues and sought to make the material relevant to her students. Her assessments included PowerPoints on Renaissance art and World War I, a poster assignment where pupils defined political

ideologies by using words and illustrations, and research projects on genocide and other current issues.

Mara had a positive experience in the teacher education program and integrated many of the things that she learned, especially from the social studies methods course. Some of her assessment activities were adapted directly from the activities her methods instructor used in his class. "I stole it all and made it mine," Mara commented when asked to describe what she took away from the methods course (Mara, Interview 11). Her methods instructor emphasized higher-order historical thinking, which is something that Mara emphasized in her class. She focused assignments on historical thinking skills that were stressed in her methods class. "My Ed[ucation] history professor, he's all about sourcing, contextualization, and corroborating," she said, referring to metacognitive strategies used in historical thinking (Mara, Interview 5). She incorporated those strategies into her instruction.

> I combined them with, and I talked about this in my paper, I took it from my lit review. One of the guys called it, called six of these his most important ones. My Ed history professor, he's all about sourcing, contextualization, and corroborating. And I decided to put in question posing and generalizations because I think those are important skills too. (Mara, Interview 5)

She struggled with integrating some of the things that she learned in methods class and her ideas about teaching. Mara thought about balancing her assessments and enjoyed more authentic projects but found herself choosing less authentic summative assessments. She explained,

> [S]o many [sic] of the stuff that I learned in [my methods] class it's about making these great creative assessments that are you know testing so many things on multiple spiral levels and I'm finding that I'm doing good mini assessments and then the final unit assessment will be like your standard test. Or the other way I'll

be giving like your standard you know section quizzes....You know and then the final assessment will be one of those. I just wish I had and I think this will come with being an experienced teacher and just having more time where I'm not planning everything … (Mara, Interview 8)

Mara did not think she was creating summative assessments that reflected authentic goals and considered ways that she could incorporate more authentic assessments.

Mara's strong authentic practices were supported by her strong beliefs about authentic learning. Having students construct their own knowledge was a central theme in her teaching, and she sought ways to balance higher order thinking with covering the content knowledge she felt was important. During the teacher education program she outlined these beliefs as she described who she considered to be an ideal teacher:

My whole college experience was about learning how to think, whereas in high school I learned what to think. They told you what to think and you were, "Okay." Where in college they taught you how to think. So I'd love to be that teacher who teaches you how to think, but with standards, standardized testing…I want my students to graduate high school and they need to know "x" amount of facts. And I need to teach it to them. I'm hoping to find a balance. (Mara, Interview 1)

During the student teaching period she said "the most important thing" was to get students "to be able to be thinking historically" and focus on the process of thinking about the past (Mara, Interview 5). That included things like having pupils interpret artwork by considering the historical context in which it was created or evaluating which cause of World War I was the most significant. She did not want her pupils to just reproduce information. She asked pupils to make judgments about historical events and "express what you're thinking on paper and are you just memorizing stuff or can you actually say something intelligent about it" (Mara, Interview 8). She emphasized interpreting information and active learning.

As the example above suggests, Mara's assignments asked pupils to apply information. For one assessment she asked pupils to define a political ideology by reading competing definitions and "illustrate a pictogram to show that they really understood what it meant" (Mara, Interview 10). As the example in Figure 5.7 illustrates, students demonstrated their understanding of a definition (in this case, the definition of "socialism") by drawing an image that best reflected the definition (in this case, using an umbrella to reflect a system that covered all members of society). In this assignment, Mara encouraged students to beyond just memorizing a definition but applying the definition into their own words. This represents extensive conceptual understanding which is part of the standard disciplined inquiry.
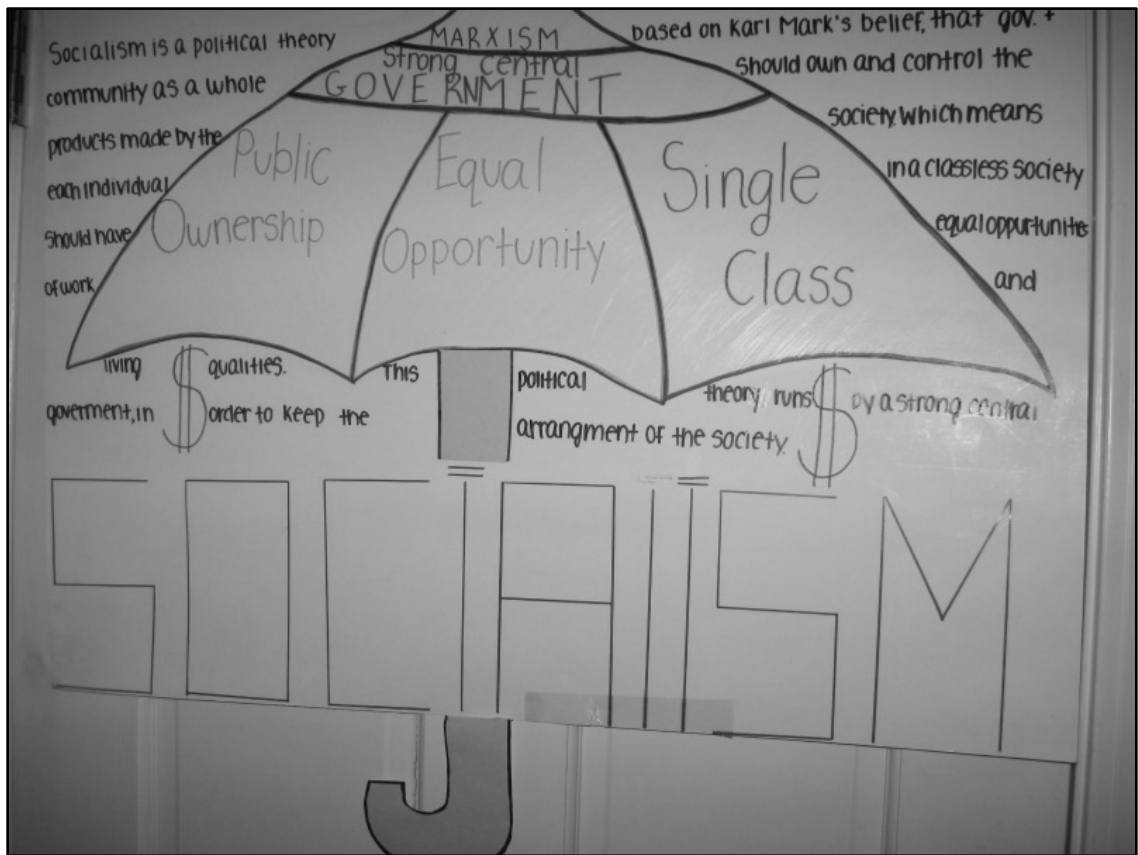


*Figure 5.7.* Pupil work example of Mara's political ideology poster assignment.

Mara used pupils' abilities to construct knowledge as a criterion for evaluating

student learning. For example, on the PowerPoint assessment where pupils interpreted

Renaissance art, Mara considered how well pupils were able to make their own meaning

about the artwork. She described how she graded one group's work:

> I thought they did an excellent job of talking about who the artist is and how she chooses to depict the scene, and then contrasting it to how a man did it and how he chose to depict the scene. So that's what I meant by how was I able to assess what they learned. They clearly were able to look at these and go beyond, all right, what's literally going on. (Mara, Interview 5)

She wanted her pupils to have a deeper understanding of the work and the historical

context than just having students repeat the artist's name and the date when the work was

created. Students had to connect the work to things that were going on in the artist's

lifetime and apply that information to their understanding of the artwork. Using another

example of pupils' work she drew specific attention to her desire to emphasize critical

thinking:

> This one is my "high" because I felt that all three aspects were great. They did excellent research because this is, this was a research project. They found out information that you can't just Google and it's the number one thing. The artwork that they, visually it's an excellent PowerPoint. During their presentation they spoke well. They said the right things. It was interesting. They were funny, and they were able to do those higher-order historical thinking skills that I so wanted them to do. (Mara, Interview 5)

She praised students for using their own ideas and not "regurgitating" Mara's words "but

putting [their] own spin on it" (Mara, Interview 5). Mara wanted her students to interpret

the information for themselves. She was that pupils were able to do good work and

respected them as well.

Mara also stressed that pupils support their opinions with evidence, which is consistent with the authentic intellectual work criterion elaborated written communication. Mara engaged pupils in class debates where she reminded pupils to "[B]ack up your positions. Do it respectfully…Okay, one at a time. Comment on comment. Back up your positions" (Mara, Observation). On her assignment on American foreign policy related to genocides in other parts of the world, she asked pupils to consider different possible options that the United States could pursue. As the description of the assignment in Figure 5.8 demonstrates, pupils were asked to use "specific examples" to support their opinion. In a portion of the assessment that is not shown here, pupils were also asked to consider arguments "for" and "against" each option so that they could better develop their own argument. Pupils expressed their ideas in writing and orally in a classroom debate. They were developing skills to think rationally and support and evaluate arguments.

---

Write a persuasive 1-page paper why your option should serve as the basis for our country's foreign policy concerning genocide. Use specific examples from the 5 case studies to support your argument and feel free to draw on other historical evidence for further support. There is no need to do further research.

*Be prepared to share with the class during a mini-debate

---

*Figure 5.8*. Assessment task description for Mara's assignment on American foreign policy and genocide.

As a result of teaching in a strong academic school, Mara had considerable freedom in constructing learning opportunities and did not face excessive pressure to prepare pupils for the State Assessment. However, she did alter her practices during her

second year to incorporate aspects of the State Assessment to help her pupils for the next

year. She explained the changes she made:

> I'm trying to do more [State Assessment]-based stuff on my tests because my kids are going to have to take it their junior year. So we're looking at like a lot of document-based questions on the test which is a new component. So I still have my multiple choice, my short answer and my chronology. Well, now I'll throw in like you know either a document or like a political cartoon or a picture or a painting. (Mara, Interview 10)

Although she added more formal summative assessment in her second year of teaching,

she continued to assign more authentic projects, such as the current events social action

project that she submitted in her last TAPL interview, a highly authentic assessment that

received a score of 9.5 out of 10. In that project pupils researched issues that were

important to them and the community and created a project. Mara preferred the projects

because she felt they were a better indication of what pupils really knew and found ways

to use those types of assessments.

Mara's commitment to authentic work prompted her to keep making connections

for pupils. She was committed to making connections between what her students were

doing in school and the world outside of the classroom:

> I would say the big picture is, would be content, there's just so much to go through so, and a lot of the kids don't necessarily care about the specific events or the specific people. I think if you can do the themes and the big questions and then kind of keep making those connections all throughout the year, that's the most important thing and kind of showing them why history is important and, you know, how it, not necessarily relates to their life but just how they can use those skills, no matter what they're doing. (Mara, Interview 9)

She believed that these authentic projects that were related to pupils' lives had a greater

impact on student learning. She described one pupil who was usually "a low student"

because she "doesn't get memorizing facts or why it's important." (Mara, Interview 11).

However, on a project where she was allowed to choose a topic to explore, an assessment task that had a score of 9.5 and a class mean of 9.0 (the highest pupil work score in the entire study), this student did "better than some of the other writing that she's done because I think she was making the connection" to current events and history and "she was saying as learned in such and such, and I was like, you paid attention" (Mara, Interview 11). Making the material relevant to pupils' lives influenced pupil learning.

For Mara, connecting material to pupils' lives was also related to preparing pupils for democratic living. She conceptualized the purpose of schooling as connected to the broader social world and stressed these concepts in class discussions. She explained:

> I never really thought about it until even teaching this course, how much of it is preparing them to be citizens in this country, never really thought about that or thought how I was going to address it, but every day we talk about, even current events or even huge things was, you know, the war in Israel.  We talked about it so many parallels with even the American Revolution, Viet Nam, the Civil War, so many of these historical themes we can see in current events.  And also we were just, we had a discussion today about *Roe v. Wade*, and how, you know, [President] Bush was appointing the new Supreme Court, how one person can change so many legal precedents.  …I kind of brought it back and I was like, "Let's just think about this.  Why is it an issue?  Why should you all go out and register to vote?  And I was, like, you know, it doesn't matter what you believe it, it just matters that whether or not you have that choice anymore." And, so, so that's kind of something I never really thought about.  (Mara, Interview 6)

Mara focused on making what they were learning about in school relevant to students' lives outside of school.

Mara's authentic intellectual work came at a price. Mara had difficulty with the work/life balance; she worked extremely long hours and weekends. During her first year she said, "And it's the hardest I've ever worked in my entire life but I'm also the happiest I've ever been" (Mara, Interview 8). She continued to work more than 12 hour days in

her second year as well. During her second year of teaching she realized that she had taken on too much. She spent her weekends "grading and planning" and was concerned that she was "going to burn out." She commented about not seeing any friends and decided to stop coaching the swim team as she moved into the third year of teaching. It was a challenge to do all of the high quality work. However, Mara remained committed to engaging pupils in authentic intellectual work.

**Matt and Mara**

The cases of Matt and Mara illustrate that the degree to which teachers consider assessment and pupil learning that align with the concepts of authentic intellectual work makes for very different classrooms. Here both Matt and Mara were considered by their schools as successful teachers in preparing their pupils academically, but in terms of engaging pupils in authentic intellectual work, they were very different. Matt's pupils recalled information and were not asked to see concepts related to the world or elaborate on their ideas. Mara's pupils, on the other hand, learned to evaluate and interpret information. Matt did not successfully engage pupils in authentic work, whereas Mara did. These two cases illustrate that engaging pupils in authentic intellectual work is more than just preparing pupils for achievement tests. Even teachers who are successful in that regard may not be engaging pupils in authentic work. Measuring teacher performance based on pupil test scores does not get at all aspects of the quality of instruction inside the classroom.

Their cases also demonstrate that beliefs about teaching and learning did inform practices. Mara consistently referred to critical thinking, conceptual understanding, and

relevance to the world, concepts that are part of authentic intellectual work. Similarly, her assessment tasks and her pupils' work received high scores when evaluated according to the RISER rubric which suggests they attended to the concepts of authentic learning. In contrast, Matt did not consider the elements that Mara considered as paramount. His assessment tasks received low scores on the RISER rubric. This suggests that beliefs play a role in how teachers develop their assessments and evaluate pupil learning. Teachers need to be committed to authentic intellectual work to do authentic intellectual work.

In addition, these two cases also illustrate the role that time plays in beginning teachers' abilities to implement authentic intellectual work. Mara spent a significant amount of time planning and evaluating pupil learning, at the expense of her personal life, to engage pupils in authentic work. Matt, on the other hand, placed a stronger emphasis on his personal life and chose assessments based on the amount of time it would take to implement and grade. This has implication for expectations for beginning teachers and authentic work and raises certain questions. Is it reasonable to expect teachers to give up their lives? Do teachers have enough time in school to develop and implement authentic intellectual activities? Chapter 6 explores some of these questions in more detail.

Overall, when the beginning teachers in this study talked about their goals for and understandings of assessment and pupil learning, their responses varied in the ways they attended to concepts related to authentic intellectual work. Some teachers consistently mentioned these ideas, some teachers occasionally mentioned these ideas, and some teachers rarely mentioned these ideas. For the most part, teachers' thoughts were aligned

with the way their assessments and pupil work were evaluated according to the rubric for authentic intellectual work. Teachers in this study utilized assessments that were in line with their goals for assessment and pupil learning. Factors such as time, accountability, student ability/motivation, classroom management, and content are influenced teachers' implementation of authentic intellectual work. In the next chapter I expand upon implications for research, policy, and practice based on these findings.

## CHAPTER SIX: CONCLUSION AND IMPLICATIONS

Preparing students to engage in democratic society requires cultivating citizens who are capable of making informed, rational decisions about complex issues related to the common good (Gutmann, 1987). Research suggests that teachers and pupils do not routinely engage in this type of thinking (Bryk et al., 2000; Gore et al., 2004; Newmann et al., 1996). To ensure that teachers are capable of providing learning opportunities that develop these skills, teacher education needs to promote critical thinking and meaningful instruction. Few studies examine the ways in which teacher education prepares new teachers to consider skills, such as critical thinking, that are vital to participation in democratic society (Gore et al., 2004; Michelli & Keiser, 2005b). One possible way to evaluate the learning opportunities that teachers provide in a manner consistent with democratic education, is through the lens of Fred Newmann's framework of "authentic intellectual work" (Newmann & Associates, 1996). With an emphasis on the construction of knowledge, disciplined inquiry, and value beyond school, the framework provides criteria to measure teacher candidate and pupil learning in a manner that takes into account the type of complex thinking that citizens need.

This dissertation presents the results of the Teacher Assessment/Pupil Learning (TAPL) study, which examined both how teacher candidates engaged pupils in authentic intellectual work and their pupils' responses to those opportunities. My analysis suggests that teachers who were prepared in a teacher education program whose mission was consistent with authentic intellectual work were able to provide moderate levels of authentic learning, although some teachers were able to attain higher levels than others.

Teachers' abilities to promote authentic intellectual work were influenced by their beliefs about assessment and pupil learning, as well as certain contextual factors in their classrooms and schools. Drawing on longitudinal quantitative and qualitative data, this study illustrates the importance of using multiple methods to assess teachers' practices and their pupils' learning to better understand the type of learning that occurs in beginning teachers' classrooms. In this chapter I review the overall findings of this study and explore implications for research, practice, and policy.

<div align="center">

**Beginning Teachers and Authentic Intellectual Work**

</div>

This study shows that the beginning teachers who were prepared in a teacher education program that emphasized pupil learning and the construction of knowledge, were, on the whole, able to provide learning opportunities for their pupils that were considered to be at a moderate level of authentic intellectual work. These teachers created assessments that were of similar quality to veteran teachers (Bryk et al., 2000; King et al., 2001). There were some teachers who engaged their pupils at higher levels of authentic intellectual work than others. Their pupils' responses to these learning opportunities varied, although, generally speaking, were also at a moderate level and were related to the quality of the assessment. Pupils who were given assessments of higher authentic intellectual quality were more likely to produce work of higher quality. Conversely, pupils who were given assessments that consisted of low levels of authentic intellectual work were more likely to produce low levels of authentic intellectual work. This dissertation confirms other studies (Bryk et al., 2000; King et al., 2001) that show the importance of assessments to pupil performance. This is significant because it suggests

that pupil achievement is directly influenced by the quality of assessments and assignments teachers use. These beginning teachers were able to engage pupils in authentic intellectual work. The teachers from this teacher preparation program attended to some of the necessary components of authentic intellectual work as they constructed learning opportunities.

Embedded within a larger qualitative case studies project, this concurrent mixed methods study provides insight into the way teachers' beliefs and classroom contexts influence their abilities to implement authentic intellectual work. Teachers whose goals and thoughts about assessment and learning aligned with authentic intellectual work were more likely to use assessment tasks that were considered as high levels of authentic intellectual work. Teachers who did not talk about authentic assessment, and focused instead on lower-order skills such as factual recall and decontexutalized content knowledge, were more likely to utilize assessments with very low levels of authentic intellectual work. This suggests that teachers have to think that the underlying principles of authentic intellectual work are important learning goals in order to utilize assessments that engage pupils in these concepts. Teachers' beliefs about teaching and learning influenced their teaching practices and the decisions they make related to assessment.

Certain classroom factors also influenced teachers' use of authentic intellectual work. One significant factor was the amount of time teachers had to construct and implement assessments. Focusing on authentic intellectual work is time consuming, both in the creation of authentic assessments and in evaluating pupil performance. Teachers often made decisions related to assessment based on the time it would take to use or

grade assessments. According to interviews, the beginning teachers in this study who scored high on authentic assessments were teachers who worked 10- to 15-hours a day and throughout the weekend. These teachers devoted tremendous amounts of time to their teaching. Other teachers chose not to utilize authentic assessments at times because the assessments took too long to create or evaluate. These teachers typically spent less time during non-school hours on work. This raises serious questions about the demands on new teachers' time and suggests that new teachers need more time during the school day that allows them to do this type of work.

This study also found that one of the reasons teachers did not implement authentic intellectual work was because there was no space in the curriculum to include "additional" material. Teachers did not always have autonomy to choose what assessments and assignments they used in their classrooms. Many teachers in this study incorporated aspects of inauthentic assessment into their assignments in order to prepare pupils for standardized tests or, in other cases, were mandated to use particular curricular materials which often emphasized lower-order skills. It is also interesting to note that when the beginning teachers in this study were able to create their own assessments, the assessments were more likely to have higher levels of authentic intellectual work than those assessments that were created by others, whether the assessments came from a textbook, scripted curriculum, or school- or district-wide assessments. This is important for two reasons. First, it illustrates that some of the beginning teachers in this study were able to produce high quality assessments when given the opportunity. Second, this

finding raises significant questions about the prescriptive assessments that some schools currently use. These questions are addressed below in the implications section.

My analysis also suggests that the level of authentic intellectual work may be influenced by grade level and content area, although it is evident from this study that more research is needed to develop a clearer understanding of this relationship. The external TAPL scores (researchers' ratings using the authentic intellectual work rubrics) clearly show a difference in quality between assessments and pupil work in writing and social studies when compared to assessments and pupil work in math and science. Writing and social studies assessments were generally of higher quality and enabled pupils to perform at higher levels. Although this study's small sample size warrants further investigation into the relationship between academic discipline and authentic intellectual work, the findings here are consistent with large-scale projects that concluded that writing and social studies assessments had higher levels of authentic intellectual work than math assessments (Bryk et al., 2000). The findings related to grade level are less clear. There were no statistical differences between the elementary and secondary teachers' assessment tasks, although secondary pupils were more likely to produce slightly higher levels of authentic intellectual work than elementary pupils. My analysis of the interview data suggests that elementary teachers thought more about balancing lower- and higher-order thinking for their pupils than the secondary teachers, who talked more about pushing their pupils to think critically. However, it was not clear from the data in this study what the nature of this relationship is and further investigation is warranted. This does seem to suggest that secondary English and social studies

classrooms engage in higher levels of authentic intellectual work than in elementary or math and science classrooms.

<p style="text-align:center"><strong>Implications</strong></p>

This study has several implications for research, practice, and policy. The findings shed light on issues related to how educational researchers and teacher educators evaluate teacher candidate and pupil learning in complex ways and attempts to utilize multiple measures for evaluation. In addition, this study raises important questions about conditions in schools that limit teachers' abilities to provide authentic learning opportunities for their pupils.

### Implications for Research

First, this study has several implications for educational researchers. This study presents one attempt at measuring teacher candidate and pupil learning based on multiple measures of assessment and defines learning by the actual work that teachers and pupils do in their classrooms on a daily basis. Here, pupil learning is not defined simply by teacher or pupil achievement on standardized tests. However, to follow teachers longitudinally through the teacher preparation program and first two years of teaching, as this study did, is cost-, labor-, and time-intensive. The TAPL study was part of a larger Qualitative Case Studies (QCS) project which involved more than 15 researchers, had significant funding, and took place over seven years. Specifically, this dissertation required multiple researchers to conduct the 54 TAPL interviews and score the 53 assessment tasks and 481 samples of pupil work. As described in Chapter 2, the majority of other studies that have used the framework of authentic intellectual work to evaluate

teachers' assessments and pupil work were all part of large-scale research projects, such as the Center on Organization and Restructuring of Schools (Newmann & Associates, 1996), the Consortium on Chicago School Research (Newmann, Lopez et al., 1998), the Queensland School Reform Longitudinal Study (2001), and the Centre for Research in Pedagogy and Practice (Luke et al., 2005). In addition, the TAPL study went beyond the work of these research groups in its attempt to capture teachers' perspectives through intensive, in-depth interviews, all of which took significant time and labor to conduct. This study demonstrates that extensive resources are needed to conduct this type of research which is something that should be considered when choosing to do this type of work in the future.

Participants were also asked to invest significant time in interviews and in their selection of assessments and pupil work samples. Some teachers had difficulty submitting whole class sets of pupil work, often because they did not have the time, or in some cases access to a photocopier, to make copies of the work. Sometimes they did not have pupil work available or needed to return materials to pupils before they were able to submit the work for the TAPL analysis. At other times, as described in Chapter 4, teachers submitted whatever assessments were available instead of submitting a "culminating" assessment. Since formative assessments were less likely to be as authentic as culminating assessments (Gleeson et al., 2008), these assessments may not have been an accurate representation of teachers' use of authentic intellectual work in the more significant summative assessments.

This study differed from previous large-scale authentic intellectual work studies (Newmann, Lopez et al., 1998; QSRLS, 2001) by incorporating extensive qualitative data as a way to provide context for how teachers thought about assessment and pupil learning and how they implemented assessments in their classrooms. The TAPL protocol created by the QCS research team combines the quantitative evaluation of assessments and pupil work with an interview component. The addition of the interview contributes to the research by adding a new layer of understanding of how teachers think about concepts related to authentic intellectual work; how they think about assessment; and the factors that influenced their assessments, practice, and pupils' learning. The interview component underscores the importance – and challenges – of using mixed methods to analyze teacher and pupil learning. The interviews provided insight into the contextual factors that constrained teachers' implementation of authentic intellectual work, such as the time required to construct and evaluate authentic work and the pressures they faced to prepare pupils strictly in terms of the standardized state or district assessments.

Classroom management and pupils' academic abilities were also important. This shows that some teachers thought about aspects of authentic intellectual work that were not reflected in their external TAPL scores because of contextual factors. There were also instances where high external scores hid troubling classroom practices which were revealed through the internal interviews. Scoring assessments and pupil work according to a rubric provided a more objective measurement to compare teachers' assessments and pupil work but it was also important to examine how contexts and teachers' beliefs about assessment and learning influenced authentic intellectual work.

One challenge for this type of mixed methods approach was sample size. This study was qualitatively-driven and embedded within a larger qualitative case studies project, wherein a small sample size had been chosen as a way to delve deeply into teachers' experiences. However, the small sample size was a limitation when it came to the quantitative part of the work. For example, as the power analysis on the two repeated measures ANOVAs conducted in this study demonstrated, the sample size was not large enough to provide the necessary power to rule out statistical error. A larger sample is required to reach a more accurate conclusion about the change in teachers' assessment and pupil work external TAPL scores over time. Researchers might consider selecting a larger sample size or extending the study beyond the preservice period and first two years of teaching to incorporate additional data points in future research.

Another limitation of the TAPL protocol is that the protocol was not specific enough in its description of the types of assessments teachers were asked to submit. The protocol only stated that teachers should submit a "culminating" assessment that they used with their pupils in their classroom. As a result, teachers sometimes submitted assessments that they created themselves, while at other times, submitted assessments that they had no role in creating. When teachers submitted an assessment that they did not create themselves, it was difficult to ascertain the teachers' intentions for authentic intellectual work, particularly when the assessment was mandated for use in their classrooms by the school. If the purpose of evaluating teachers' assessments is to see the degree to which the assessments teachers create reflect authentic intellectual work, then having teachers submit an assessment that they did not create provides little to no

information about their own assessment practices. The internal TAPL interview sought to rectify this by allowing teachers to talk about their goals for the assessment, explain pupil performance, and provided essential information to illuminate teachers' understandings. On the other hand, because teachers were not told what type of assessment they should submit, the work they did submit reflects the type of work that they actually used in their classroom. This provided insight into the variety of assessments that teachers used on a daily basis. These assessment task samples revealed that standardized curriculum and assessments were used regularly in some subject areas, grade levels, and school contexts. When using the TAPL protocol in the future, researchers should consider the purpose of their research and modify the assessment submission guidelines accordingly.

Finally, this study raises questions about the framework of authentic intellectual work and the RISER (2001) rubric, which accompanies it. Although the framework and rubric were designed to apply to the four major academic disciplines – writing, social studies, math, and science – the criterion, elaborated written communication, appears to privilege subjects that traditionally engage pupils in written communication. The rubric does provide explicit definitions for elaborated written communication in science and math, such as the use of graphs and charts, but there were very few instances in this study where math and science assessments fit within the parameters of that definition. Since this study confirmed previous studies that also found higher levels of authentic intellectual work in writing and social studies than in math (Newmann & Associates, 1996; Newmann, Lopez et al., 1998), it is possible that writing and social studies as subjects are more conducive to authentic intellectual work based on the definition used.

There appears to be an implicit privileging of writing. However, it is also possible that math and science teachers are more likely to utilize traditional inauthentic assessments and to implement assessments that do not encourage pupils to think more conceptually and critically about significant material. The data collected in this study did not provide enough information to explore this issue in greater detail, nor was the sample representative of all math and science teachers. This topic is worth pursuing because if the rubric fails to account for work in the disciplines of math and science, then the validity of the instrument is compromised. If, however, math and science teachers are not truly engaging pupils in authentic learning, then teacher educators need to reconsider the ways in which math and science teachers are prepared to teach. Educators need to conceptualize ways to construct more authentic assessments in these disciplines.

Another limitation of the rubric is the way the guidelines account for English Language Learners and students with disabilities. In the guidelines for scoring student writing, the RISER (2001) rubric specifically states that "[s]corers should assess the quality of the actual written work and not take into consideration possible effects of a student's linguistic background or learning disability" (p. 7). This means that ELL students and students with learning disabilities are penalized and their scores may not accurately reflect their learning; these students could have a conceptual understanding of material that exceeds what they are able to articulate in written form. Since many pupils in this study were English Language Learners and/or had IEPs, additional research should be conducted to see how their external TAPL scores were influenced by their language

and learning disabilities. Researchers might also consider modifying the rubric to take pupil ability into account.

In addition, the RISER (2001) rubric should be updated to incorporate new forms of assessment. New technologies have changed the format of how professionals engage in their field, which expands the concepts behind the criterion, value beyond school. Technology has altered the format of assessments as well. Researchers in this study found it difficult to evaluate assessments that utilized new media. For example, two of the assessments used by one teacher in the study were in the form of PowerPoint presentations. Typically, the level of written detail in a PowerPoint is not as extensive as a written essay; presenters articulate their points in greater detail during their presentation. However, since the researchers did not always observe the actual Power-Point presentations, they were not able to hear pupils articulate their understandings. Thus, pupils were evaluated solely on what was written on the PowerPoint slides. Pupils might have received different scores if their actual presentations were scored. Although this limitation suggests that future research designs should incorporate classroom observations to coordinate with the pupil work addressed in the TAPL interviews, this example also suggests that it is necessary to reconsider how to better evaluate innovative assessments.

**Implications for Practice**

This study also has implications for practice in teacher education and teaching in schools. The study presents an example of how teacher education programs can attempt to assess teacher candidate and pupil learning in a way that accounts for the work that

teachers and pupils actually do in their classroom and uses multiple measures of evaluation to construct a deeper understanding of teachers' practices. In addition, this close examination of teachers' practices raises several issues for educators to consider concerning teaching and learning in schools.

**Implications for practice in teacher education**. Current policy calls for teacher education to measure the impact of teacher preparation programs on pupil learning. As described in Chapter 1, we categorized attempts to connect teacher education to pupil learning into four types: (1) correlating preparation programs and entry pathways with pupils' achievement scores; (2) evaluating programs in terms of candidates' demonstration of classroom behaviors correlated with pupils' test scores; (3) assessing the learning opportunities teacher candidates create along with pupils' performance; and (4) assessing the learning opportunities program graduates create along with their pupils' performance (Cochran-Smith et al., 2010). This study combines the last two categories through its examination of learning opportunities that teacher candidates/ teachers create during the student teaching period and first two years of teaching. The TAPL protocol was designed to assess teacher candidate and pupil learning in a manner that reflected the teacher preparation program's commitment to deep learning and understanding; appreciation of the complexities of teaching and learning; and commitment to utilize multiple measures of assessment. This study assessed teacher and pupil learning through an evaluation of actual teacher practices, thus achieving close proximity to what teachers do on a daily basis in their classrooms. The assessment was also in close proximity to the teacher education program by following teacher candidates during the preservice period

and immediately following the teacher preparation program. In addition, through the framework of authentic intellectual work, learning was defined in the TAPL protocol as in-depth, conceptual understanding of a topic and was measured in a complex way that went beyond whether or not pupils could memorize isolated information.

One limitation of the use of the TAPL protocol in teacher education is that the protocol is a time- and labor-intensive measurement. This particular study evaluated 11 teacher candidates over a 3-year period of time. For a teacher education program that prepares large numbers of undergraduate and graduate-level teacher candidates per year, this type of extensive longitudinal study is not feasible to conduct with every teacher candidate. However, the TAPL protocol – interviews and evaluation of assessments/pupil work – could be used on a much smaller scale. Teacher educators could use the protocol during the student teaching period and evaluate one assessment task and a few samples of pupil work that teacher candidates used with their pupils. A modified TAPL protocol could also be used after teacher candidates exit the program. Teachers could submit, either electronically or by mail, one assessment task that they have used and a few samples of pupil work to the teacher education institution and record, in written or oral form, their answers to the TAPL interview questions. The scenario described here is also time-consuming but it provides deeper understandings of teacher practices and their pupils' learning that cannot be found in achievement test scores.

The use of both quantitative and qualitative data in the TAPL protocol is essential in constructing a nuanced, and more accurate, understanding of teacher and pupil learning. There were times in this study when the external TAPL scores did not

adequately represent teachers' beliefs or practices related to authentic intellectual work. In cases where teachers submitted an assignment that they did not create, the qualitative internal TAPL data provided insights into how teachers understood assessment and their pupils' learning, something that was not apparent based on the rubric scores alone. The interview data helped explain the external scores. In some cases, teachers had pupils who did not attend class or refused to do work. Most likely, those students received low external scores, which may not be the most accurate measurement of teacher practice because those students were not present during instruction. The interview data illustrated that high absenteeism was a bigger, systemic issue that negatively impacted pupil learning and was something that the teachers in this study had little control over. High absenteeism may have negatively impacted pupil work scores on the external analysis. Since the quantitative data alone might suggest the teacher did not impact pupil learning, it is imperative to examine additional data. Pupil learning is influenced by a variety of factors, factors which may have gone unnoticed without the qualitative data.

In addition, the integration of qualitative and quantitative data helped illustrate the relationship between teachers' beliefs about assessment and pupil learning and the authentic intellectual quality of their assessments. This suggests that teacher educators need to consider how their programs shape teachers' beliefs and emphasize core values related to teaching and learning within the program. Gore (2001) calls for teacher educators to explicitly address and embed concepts related to authentic intellectual work within the teacher education program as a way to help teacher candidates implement these ideas in their own practice. Although the teacher education program in this study

did not explicitly address Newmann's framework in particular, concepts of authentic

intellectual work were fundamental themes in the program. The results of this study

suggest that these ideas were either not routinely emphasized in all classes or that not all

of the teacher candidates internalized these understandings into their own practice.

Teacher education programs need to consider the ways in which programmatic themes

and values are integrated throughout the program and the ways in which these ideas are

incorporated into their graduates' teaching practices.

**Implications for practice in schools**. This study also has implications for

teaching in schools. As mentioned previously, one finding of this study was that teachers'

use of authentic intellectual work was influenced by the amount of teachers' time that it

took to implement and evaluate this type of work. Some teachers decided not to use more

authentic assessments because of the amount of time required to create and grade

authentic assessments and those who did spent long hours planning and grading. This

suggests that in order to foster authentic learning, school administrators should consider

ways to support teachers, especially beginning teachers, in this task by providing more

time for teachers to do this. Possible ways to do this include reducing class size, teaching

responsibilities, or adding more preparation time to teachers' schedules. Providing

opportunities for teachers to collaborate on constructing authentic assessments that could

then be used across classes is also one way to reduce individual teachers' workloads. If

authentic learning is a goal, it should be a goal that all teachers can attain without having

to spend all of their time outside of the classroom working.

A second finding of this study was that standardized curriculum assessments provided fewer opportunities for authentic intellectual work than the assessments that the beginning teachers in this study created themselves. Due to the increased use of standardized curriculum and statewide assessments, it is imperative that administrators, teachers, and those who create tests at the statewide level, consider the type of learning reflected in these assessments. If schools continue to use these assessments, then administrators and teachers need to select assessments that foster students' abilities to cultivate and demonstrate critical thinking skills and their own construction of knowledge. One way to do this is to consider new assessments that attempt to measure some of these skills in a standardized way (Baldi, Jin, Skemer, Green, & Herget, 2007; Council for Aid to Education, 2007; Tannenbaum & Katz, 2008). In addition, schools should consider ways to work with teachers in developing more authentic assessments and keep teachers central to curriculum development.

**Preparing Teachers and Students for Democracy**

Teaching for democratic education and preparing students to become citizens in a democracy requires fostering a learning environment that emphasizes a commitment to deep conceptual understandings, the ability to make reasoned judgments, and communicate effectively so that students, as citizens, are able to deliberate on important issues that impact the collective good. Although people may agree on these goals in theory, the reality is that competing notions of the purposes of schooling and what teachers should know and be able to do have limited opportunities for students to construct their own knowledge and apply their understandings to meaningful situations

317

and content. In an accountability climate where teacher and student learning are largely defined by single measure achievement test scores, it is troubling that efforts to prepare students to pass these tests often sacrifice the knowledge and skills necessary to maintain and extend American democratic society and government. More troubling is that the findings of this study confirm previous work suggesting that accountability frameworks constrain opportunities for democratic and critical teaching and learning (Carlson, 2008; Kahne & Westheimer, 2003; Michelli & Keiser, 2005b). This further supports that democratic education is not pervasive in American schools and illuminates the difficulties teachers who want to promote democratic education encounter.

Preparing teachers to consider teaching and learning in meaningful ways is challenging because it involves cultivating teacher candidates' abilities to think critically themselves and examine their own beliefs. However difficult, it is imperative that we meet these challenges in order to serve the best interests of a democratic society. This rests upon a commitment for teacher education to foster democratic education and prepare teachers and students to do the same.

# References

Achieve. (2008). *Out of many, one: Toward rigorous common core standards from the ground up*. Washington, DC: Author.

Alleman, J., & Brophy, J. (1997). Elementary social studies: Instruments, activities, and standards. In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 321-357). San Diego, CA: Academic Press.

Amosa, W., Ladwig, J. G., Griffiths, T., & Gore, J. M. (2007). *Equity effects of quality teaching: Closing the gap.* Paper presented at the Australian Association for Research in Education Annual Meeting, Fremantle.

Apple, M. W. (2002). Patriotism, pedagogy, and freedom: On the educational meanings of September 11th. *Teachers College Record, 104*(8), 1760-1772.

Archbald, D. A., & Newmann, F. M. (1988). *Beyond standardized testing: Assessing authentic academic achievement in the secondary school*. Washington, D.C.: National Association of Secondary School Principals, Reston, VA; National Center on Effective Secondary Schools, Madison, WI.

Arter, J. A., & Bond, L. (1996). Why is assessment changing? In R. E. Blum & J. A. Arter (Eds.), *Handbook for student performance assessment in an era of restructuring* (pp. 1-3:1- 1-3:4). Alexandra, VA: Association for Supervision and Curriculum Development.

Avery, P. G. (1999). Authentic assessment and instruction. *Social Education, 65*(6), 368-373.

Avery, P. G., Kouneski, N. P., & Odendahl, T. (2001). Authentic pedagogy seminars: Renewing our commitment to teaching and learning. *The Social Studies, 92*(3), 97-101.

Ayres, L., Kavanaugh, K., & Knafl, K. A. (2003). Within-case and across-case approaches to qualitative data analysis. *Qualitative Health Research, 13*(6), 871-883.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies, 31*(3), 285-302.

Baldi, S., Jin, Y., Skemer, M., Green, P. J., & Herget, D. (2007). *Highlights from PISA 2006: Performance of U.S. 15-year-old students in science and mathematics literacy in an international context (NCES 2008-016)*. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Banks, J. A. (2001). Citizenship education and diversity: Implications for teacher education. *Journal of Teacher Education, 52*(1), 5-16.

Banks, J. A. (2007). *Educating citizens in a multicultural society* (2nd ed.). New York: Teachers College.

Barber, B. R. (2001). Public education and democracy. *The Kettering Review, 19*(2), 18-26.

Barton, K. C., & Levstik, L. S. (2004). *Teaching history for the common good*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Benner, P. (Ed.). (1994). *Interpretive phenomenology*. Thousand Oaks, CA: Sage.

Biggs, J., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.

Blum, R. E., & Arter, J. A. (Eds.). (1996). *A handbook for student performance assessment in an era of restructuring*. Alexandria, VA: Association for Supervision and Curriculum Development.

Boone, M. (2008). A democratic education: Three interpretations. *Journal of Philosophy and History of Education, 58*, 16-22.

Boston College Lynch School of Education. (2010). Teacher education/special education, curriculum & instruction. Retrieved January 24, 2011, from http://www.bc.edu/schools/lsoe/academics/departments/teseci.html

Braden, J., Schroeder, J., & Buckley, J. A. (2001). *Secondary school reform, inclusion, & authentic assessment* (Brief no. 3). Madison, WI: Research Institute on Secondary Education Reform.

Britzman, D. P. (1991). *Practice makes practice*. Albany, NY: State University of New York Press.

Bryk, A. S., Nagaoka, J. K., & Newmann, F. M. (2000). *Chicago classroom demand for authentic intellectual work: Trends from 1997-1999*. IL: Consortium on Chicago School Research.

Burbules, N. C., & Berk, R. (1999). Critical thinking and critical pedagogy: Relations, differences, and limits. In T. S. Popkewitz & L. Fendler (Eds.), *Critical theories in education: Changing terrains of knowledge and politics*. New York, NY: Routledge.

Caracelli, V. J., & Greene, J. C. (1997). Crafting mixed-method evaluation designs. In J. C. Greene & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 19-32). San Francisco, CA: Jossey-Bass Publishers.

Carcieri, M. D. (1997). Democracy and education in the thought of Jefferson and Madison. *Journal of Law and Education, 26*(1), 1-30.

Carlson, D. (2008). 2007 AESA Presidential Address - Conflict of the Faculties: Democratic progressivism in the age of "No Child Left Behind". *Educational Studies, 43*(2), 94-113.

Carnegie Corporation of New York, & CIRCLE: The Center for Information and Research on Civic Learning and Engagement. (2003). *The civic mission of schools*. New York: Author.

Charmaz, K. (2004). Grounded theory. In S. N. Hesse-Biber & P. Leavy (Eds.), *Approaches to qualitative research: A reader on theory and practice* (pp. 496-521). New York: Oxford University Press.

Cheney, L. V. (2002). Protecting our precious liberty. In Thomas B. Fordham Foundation (Ed.), *September 11th: What our children need to know*. Washington, D.C.: Author.

Chung, R. R. (2008). Beyond assessment: Performance assessments in teacher education. *Teacher Education Quarterly, 35*(1), 7-28.

Clare, L. (2000). *Using teachers' assignments as an indicator of classroom practice. CSE technical report*. Los Angeles, CA: Center for Research on Evaluation, Standards,

and Student Testing, Los Angeles, CA; California Univ., Los Angeles. Center for the Study of Evaluation.

Clare, L., & Aschbacher, P. R. (2001). Exploring the technical quality of using assignments and student work as indicators of classroom practice. *Educational Assessment, 7*(1), 39-59.

Cochran-Smith, M. (2001). Reforming teacher education: Competing agendas. *Journal of Teacher Education, 52*(4), 263-265.

Cochran-Smith, M. (2005). Studying teacher education: What we know and need to know. *Journal of Teacher Education, 56*(4), 301-306.

Cochran-Smith, M., & Fries, K. (2005). Researching teacher education in changing times: Paradigms and politics. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 69-110). Mahwah, NJ: Lawrence Erlbaum Associates.

Cochran-Smith, M., Gleeson, A. M., & Mitchell, K. (2010). Teacher education for social justice: What's pupil learning got to do with it? *Berkeley Review of Education, 1*(1), 35-61.

Cohen, J. (1992). A power primer. *Quantitative Methods in Psychology, 112*(1), 155-159.

Collet-Klingenberg, L. L., Hanley-Maxwell, C., & Stuart, S. (2000). *The relationship of an ecological model of career development to authentic learning*. Madison, WI: Research Institute on Secondary Education Reform, Wisconsin Center for education Research, University of Wisconsin, Madison.

Corbett, H. D., & Blum, R. E. (1996). Placing student performance assessment within the context of school restructuring: Introduction. In R. E. Blum & J. A. Arter (Eds.), *A handbook for student performance assessment in an era of restructuring* (pp. 11-11:11-11-11:13). Alexandria, VA: Association for Supervision and Curriculum Development.

Cornett, J. W., & Gaudelli, W. (2003, April 10-12). *Fostering democratic discussions in the classroom: An examination of tensions in civic education, contemporary schools, and teacher education*. Paper presented at the Conference of the Education in Democracy for Social Studies Teachers: An Institute for Florida Teacher Educators, Gainesville, FL.

Council for Aid to Education. (2007). *College & work readiness assessment institutional report*. New York, New York: Author.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles, CA: SAGE.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2002). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 209-240). Thousand Oaks, CA: SAGE Publications.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 209-240). Thousand Oaks, CA: SAGE Publications.

Darling-Hammond, L. (2006). Assessing teacher education: The usefulness of multiple measures for assessing program outcomes. *Journal of Teacher Education, 57*(2), 120-138.

Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work*. New York: Teachers College Press.

Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and Teacher Education, 16*(5-6), 523-545.

Denzin, N. K. (1994). The art and politics of interpretation. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 500-515). Thousand Oaks, CA: Sage.

Denzin, N. K., & Lincoln, Y. S. (1994). Introduction: Entering the field of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 1-17). Thousand Oaks, CA: SAGE Publications.

Denzin, N. K., & Lincoln, Y. S. (2005). The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (3rd ed., pp. 1-32). Thousand Oaks, CA: Sage.

Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. New York: The Free Press.

Dixson, A. D., & Dingus, J. E. (2007). Tyranny of the majority: Re-enfranchisement of African-American teacher educators teaching for democracy. *International Journal of Qualitative Studies in Education, 20*(6), 639-654.

Dobbert, M. L., & Kurth-Schai, R. (1992). Systematic ethnography: Toward an evolutionary science of education and culture. In M. D. LeCompte, W. L. Millroy & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 93-159). San Diego, CA: Academic Press, Inc.

Duncan, A. (2010, October 27). *The link between standards and innovation: Secretary Arne Duncan's remarks to innovate to education symposium*, Richmond, VA.

Eisenhart, M. (2001). Changing conceptions of culture and ethnographic methodology: Recent thematic shifts and their implications for research on teaching. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 209-225). Washington, D.C.: American Educational Research Association.

Eisenhart, M., & Howe, K. R. (1992). Validity in educational research. In M. D. LeCompte, W. L. Millroy & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 643-680). San Diego, CA: Academic Press, Inc.

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher, 18*(3), 4-10.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Executive summary. "The Delphi report"*. Millbrae, CA: The California Academic Press.

Finn, C. E., Jr. (2002). Introduction. In Thomas B. Fordham Foundation (Ed.), *September 11th: What our children need to know* (pp. 4-11). Washington, D.C.: Author.

Finn, C. E., Jr. (2004). Faulty engineering. *Education Next, 4*(2), 16-21.

Fischer, C. F., & King, R. M. (1995). *Authentic assessment: A guide to implementation*. Thousand Oaks, CA: Corwin Press.

Gee, J. P. (1996). *Social linguistics and literacies: Ideology in discourse* (2nd ed.). New York: Routledge Falmer.

Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.

Gibson, C. (1995). Critical thinking: Implications for instruction. *RQ, 35*(1), 27-35.

Girden, E. R. (1992). *ANOVA: Repeated measures*. Newbury Park, CA: Sage Publications.

Girod, G. R. (2002). *Connecting teaching and learning: A handbook for teacher educators on teacher work sample methodology*. Washington, DC: AACTE Publications.

Giroux, H. A. (1988). *Schooling and the struggle for public life*. Minneapolis, MN: University of Minnesota Press.

Giroux, H. A. (2002). Democracy, freedom, and justice after September 11th: Rethinking the role of educators and the politics of schooling. *Teachers College Record, 104*(6), 1138-1162.

Glaser, B. (1992). *Basics of grounded theory analysis*. Mil Valley, CA: Sociology Press.

Glaser, B., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Boston, MA: Allyn and Bacon.

Gleeson, A. M., Mitchell, K., Baroz, R., & Cochran-Smith, M. (2008, March). *Teacher education for social justice: What's pupil learning got to do with it?* Paper presented at the American Educational Research Association Annual Meeting, New York, NY.

Gore, J. M. (2001). Beyond our differences: A reassembling of what matters in teacher education. *Journal of Teacher Education, 52*(2), 124-135.

Gore, J. M., Griffiths, T., & Ladwig, J. G. (2002, December). *Exploring "Productive pedagogy" As a framework for teacher learning.* Paper presented at the Annual Meeting of the Australian Association for Research in Education, Brisbane, Australia.

Gore, J. M., Griffiths, T., & Ladwig, J. G. (2004). Towards better teaching: Productive pedagogy as a framework for teacher education. *Teaching and Teacher Education, 20*(4), 375-387.

Gore, J. M., Ladwig, J. G., Amosa, W., & Griffiths, T. (2008, November/December). *On the status and quality of physical education as school subject: An empirical study of classroom and assessment practice.* Paper presented at the Australian Association for Research in Education Annual Meeting, Brisbane.

Gore, J. M., Ladwig, J. G., Amosa, W., Griffiths, T., Parkes, R., & Ellis, H. (2008, November/December). *Quality assessment in university social science courses.* Paper presented at the Australian Association for Research in Education Annual Meeting, Brisbane.

Gore, J. M., Ladwig, J. G., Griffiths, T., & Amosa, W. (2007, November). *Data-driven guidelines for high quality teacher education.* Paper presented at the Australian Association for Research in Education Conference, Fremantle.

Gore, J. M., Williams, C., & Ladwig, J. G. (2006, November). *On the place of pedagogy in the induction of early career teachers.* Paper presented at the Australian Association for Research in Education Annual Conference, Adelaide.

Grant, S. G., Gradwell, J. M., & Cimbricz, S. K. (2004). A question of authenticity: The document-based question as an assessment of students' knowledge of history. *Journal of Curriculum and Supervision, 19*(4), 309-337.

Greene, J. C., & Caracelli, V. J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. In J. C. Greene & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 5-17). San Francisco, CA: Jossey-Bass Publishers.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*, 255-274.

Griffiths, T., Amosa, W., Ladwig, J. G., & Gore, J. M. (2007). *Equity and pedagogy: Familiar patterns and qt based possibilities.* Paper presented at the Australian Association for Research in Education Conference, Fremantle.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: SAGE Publications.

Gutmann, A. (1987). *Democratic education.* Princeton, New Jersey: Princeton University Press.

Halpern, D. F. (1997). *Critical thinking across the curriculum: A brief edition of thought and knowledge.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist, 53*(4), 449-455.

Hanley-Maxwell, C., Phelps, L. A., Braden, J., & Warren, V. (1999). *Schools of authentic and inclusive learning (brief no. 1).* Madison, WI: Research Institute on Secondary Education Reform, Wisconsin Center for Education Research, University of Wisconsin, Madison.

Hanley-Maxwell, C., Phelps, L. A., Braden, J., & Warren, V. (2000). *Schools of authentic and inclusive learning: Research institute on secondary education reform (RISER) for youth with disabilities brief no. 1.* Madison, WI: Wisconsin Center for education Research.

Hayes, D., Mills, M., Christie, P., & Lingard, B. (2006). *Teachers and schooling making a difference: Productive pedagogies, assessment and performance.* Australia: Allen & Unwin.

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice.* New York: The Guilford Press.

Hill, C. E., Thompson, B. J., & Williams, E. N. (1997). A guide to conducting consensual qualitative research. *The Counseling Psychologist, 25*(4), 517-572.

Hodder, I. (1994). The interpretation of documents and material culture. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 393-402). Thousand Oaks, CA: SAGE Publications.

Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher, 17*(8), 10-16.

Huberman, M., & Miles, M. (1994). Data management and analysis methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 428-444). Thousand Oaks, CA: SAGE Publications.

Jefferson, T. (1778/1984). A bill for the more general diffusion of knowledge. In M. D. Peterson (Ed.), *Writings: Thomas Jefferson* (pp. 365-373). New York, NY: Literary Classics of the United States.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133.

Kahne, J., & Westheimer, J. (2003). Democracy and civic engagement: What schools need to do. *Phi Delta Kappan, 85*(1), 34.

Kay, K., & Honey, M. (2006). Establishing the R&D agenda for twenty-first century learning. *New Directions for Youth Development, Summer* (110), 68-80.

Kennedy, M. (1999). The problem of evidence in teacher education. In R. Roth (Ed.), *The role of the university in the preparation of teachers* (pp. 87-107). Philadelphia, PA: Falmer Press.

King, M. B., Schroeder, J., & Chawszczewski. (2001). *Authentic assessment and student performance in inclusive schools, brief no. 5*. Madison, WI: Research Institute on Secondary Education Reform (RISER) for Youth with Disabilities.

Koh, K. H., Lee, A. N., Tan, W., Wong, H. M., Guo, L., Lim, T. M., et al. (2005, May). *Looking collaboratively at the quality of teachers' assessment tasks and student work in Singapore schools.* Paper presented at the Rethinking and reassessing the construct of authentic intellectual quality in the Singapore context, National Institute of education, Nanyang Technological University, Singapore.

Kovacs, P. (2009). Education for democracy: It is not an issue of dare: It is an issue of can. *Teacher Education Quarterly, 36*(1), 9-23.

Kunzman, R. (2006). The civic (and pedagogical) virtue of recognizing reasonable disagreement. *Theory and Research in Social Education, 34*(2), 162-182.

Ladwig, J. G. (1998). Authentic school reform. *Discourse: Studies in the Cultural Politics of Education, 19*(1), 113-119.

Ladwig, J. G., & King, M. B. (2003). *Quality teaching in NSW public schools: An annotated bibliography*. Sydney: State of NSW Department of Education and Training.

Ladwig, J. G., Smith, M., Gore, J. M., Amosa, W., & Griffiths, T. (2007, November 25-29). *Quality of pedagogy and student achievement: Multi-level replication of authentic pedagogy.* Paper presented at the Australian Association for Research in Education Conference, Fremantle.

Lee, V. E., & Smith, J. B. (1994). *High school restructuring and student achievement: A new study finds strong links*. Madison, WI: Center on Organization and Restructuring of Schools.

Lee, V. E., Smith, J. B., & Croninger, R. G. (1995). *Another look at high school restructuring: More evidence that it improves student achievement, and more*

*insight into why*. Madison, WI: Center on Organization and Restructuring of Schools.

Lee, V. E., Smith, J. B., & Croninger, R. G. (1997). How high school organization influences the equitable distribution of learning in mathematics and science. *Sociology of Education, 70*(April), 128-150.

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2008). *SPSS for intermediate statistics: Use and interpretation* (3rd ed.). New York, NY: Lawrence Erlbaum Associates.

Lesh, R., & Lamon, S. J. (1992). Assessing authentic mathematical performance. In R. Lesh & S. J. Lamon (Eds.), *Assessment of authentic performance in school mathematics*. Washington, D.C.: American Association for the Advancement of Science.

Lingard, B., Hayes, D., & Mills, M. (2003). Teachers and productive pedagogies: Contextualising, conceptualising, utilising. *Pedagogy, Culture, & Society, 11*(3), 399-424.

Lipman, M. (1988). Critical thinking - what can it be? *Educational Leadership, 46*(1), 38-43.

Luke, A., Freebody, P., Shun, L., & Gopinathan, S. (2005). Towards research-based innovation and reform: Singapore schooling in transition. *Asia Pacific Journal of Education, 25*(1), 5-28.

Mabry, L. (1999). *Portfolio plus: A critical guide to alternative assessment*. Thousand Oaks, CA: Corwin Press, Inc.

Macedo, S. (2005). *Democracy at risk: How political choices undermine citizen participation, and what we can do about it*. Washington, D.C.: Brookings Institution Press.

Macedo, S., Alex-Assensoh, Y., Berry, J. M., Brintnall, M., Campbell, D. E., Fraga, L. R., et al. (2005). *Democracy at risk: How political choices undermine citizen participation, and what we can do about it*. Washington, D.C.: Brookings Institution Press.

Marks, H. M. (1995). *Student engagement in the classrooms of restructuring schools*. Madison, WI: Center on Organization and Restructuring of Schools.

Martin, D. S. (2005). Critical thinking for democracy and social justice. In N. M. Michelli & D. L. Keiser (Eds.), *Teacher education for democracy and social justice* (pp. 209-228). New York: Routledge.

Matsumura, L. C. (2003). *Teachers' assignments and student work: Opening a window on classroom practice. CSE report.* Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA; California Univ., Los Angeles. Center for the Study of Evaluation.

McLaren, P. L. (1994). Foreword. In K. S. Walters (Ed.), *Re-thinking reason: New perspectives in critical thinking* (pp. ix-xviii). Albany, NY: State University of New York Press.

Michelli, N. M. (2005). Education for democracy: What can it be? In N. M. Michelli & D. L. Keiser (Eds.), *Teacher education for democracy and social justice* (pp. 3-31). New York: Routledge.

Michelli, N. M., & Keiser, D. L. (2005a). Introduction: Teacher education for democracy and social justice. In N. M. Michelli & D. L. Keiser (Eds.), *Teacher education for democracy and social justice* (pp. xvii-xxv). New York: Routledge.

Michelli, N. M., & Keiser, D. L. (Eds.). (2005b). *Teacher education for democracy and social justice*. New York: Routledge.

Miles, M., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Mitchell, J. C. (1979). On quantification in social anthropology. In A. L. Epstein (Ed.), *The craft of social anthropology* (pp. 17-45). Oxford: Pergamon Press.

Mooney, M., & Phelps, L. A. (2001). *Life after high school: Which postschool outcomes matter for students with and without disabilities in restructured high schools?* (Brief no. 4). Madison, WI: Research Institute on Secondary Education Reform, Wisconsin Center for Education Research, University of Wisconsin, Madison.

Mooney, M., Phelps, L. A., & Anctil, T. M. (2002). *Using postschool outcome data to improve practices and policies in restructured inclusive high schools* (Brief no. 6). Madison, WI: Research Institute on Secondary Education Reform, Wisconsin Center for Education Research, University of Wisconsin, Madison.

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research, 40*(1), 120-123.

Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 189-208). Thousand Oaks, CA: Sage Publications.

Morse, J. M., & Niehaus, L. (2009). *Mixed method design: Principles and procedures*. Walnut Creek, CA: Left Coast Press.

Nash, M. A. (2005). "How to be thankful for being free": Searching for a convergence of discourses on teaching patriotism, citizenship, and United States history. *Teachers College Record, 107*(1), 214-240.

National Education Goals Panel. (1991). *The national education goals report: Building a nation of learners*. Washington, DC: U.S. Government Printing Office.

Newmann, F. M. (1991). Linking restructuring to authentic student achievement. *Phi Delta Kappan, 72*(6), 458-463.

Newmann, F. M., & Associates. (1996). *Authentic achievement: Restructuring schools for intellectual quality*. San Francisco: Jossey-Bass Publishers.

Newmann, F. M., Brandt, R., & Wiggins, G. (1998). An exchange of views on "Semantics, psychometrics, and assessment reform: A close look at 'authentic' assessments". *Educational Researcher, 27*(6), 19-22.

Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence?* Chicago, IL: Consortium on Chicago School Research.

Newmann, F. M., King, M. B., & Carmichael, D. L. (2007). *Authentic instruction and assessment: Common standards for rigor and relevance in teaching academic subjects*. Des Moines, Iowa: Iowa Department of Education.

Newmann, F. M., Lopez, G., & Bryk, A. S. (1998). *The quality of intellectual work in Chicago schools: A baseline report*. Chicago, IL: Consortium on Chicago School Research.

Newmann, F. M., Marks, H. M., & Gamoran, A. (1995). *Authentic pedagogy: Standards that boost student performance. Issue report no. 8*. Madison, WI: Center on Organization and Restructuring of Schools.

Newmann, F. M., Marks, H. M., & Gamoran, A. (1996). Authentic pedagogy and student performance. *American Journal of Education, 104*(4), 280-312.

Newmann, F. M., Secada, W. G., & Wehlage, G. G. (1995). *A guide to authentic instruction and assessment: Vision, standards, and scoring*. Madison, WI: Wisconsin Center for Education Research.

Newmann, F. M., & Wehlage, G. G. (1993). Five standards of authentic instruction. *Educational Leadership, 50*(7), 8-12.

Newmann, F. M., & Wehlage, G. G. (1995). *Successful school restructuring: A report to the public and educators*. Madison, WI: Center on Organization and Restructuring of Schools.

Onwuegbuzie, A. J. (2003). Expanding the framework of internal and external validity in quantitative research. *Research in the Schools, 10*(1), 71-90.

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools, 13*(1), 48-63.

Onwuegbuzie, A. J., & Leech, N. L. (2004). Post hoc power: A concept whose time has come. *Understanding Statistics, 3*(4), 201-230.

Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 351-383). Thousand Oaks, CA: SAGE Publications.

Papastephanou, M., & Angeli, C. (2007). Critical thinking beyond skill. *Educational Philosophy and Theory, 39*(6), 604-621.

Partnership for 21st Century Skills. (2006). *Results that matter: 21st century skills and high school reform*. Tuscan, AZ: Author.

Patrick, J. J., & Vontz, T. S. (2001). Components of education for democratic citizenship in the preparation of social studies teachers. In J. J. Patrick & R. S. Leming (Eds.), *Principles and practices of democracy in the education of social studies teachers: Civics learning in teacher education* (Vol. 1, pp. 39-63). Bloomington, IN: ERIC Clearinghouse for Social Studies/Social Science Education.

Paul, R. W., & Elder, L. (2006). *The miniature guide to critical thinking concepts and tools* (4th ed.). Dillon Beach, CA: The Foundation for Critical Thinking.

Pecheone, R. L., & Chung, R. R. (2006). Evidence in teacher education: The Performance Assessment for California Teachers (PACT). *Journal of Teacher Education, 57*(1), 22-36.

Perkins, J. (2004). *Confessions of an economic hit man*. San Francisco, CA: Berrett Koehler Publishers.

Queensland School Reform Longitudinal Study (QSRLS). (2001). *School reform longitudinal study*. Brisbane: The State of Queensland (Department of Education).

Ravitch, D. (2001). Now is the time to teach democracy. *Editorial Projects in Education, 21*(7).

Ravitch, D. (2008). Education and democracy: The United States of America as a historical case study. *Yearbook of the National Society for the Study of Education, 107*(1), 42-57.

Rawls, J. (1993). *Political liberalism*. New York, NY: Columbia University Press.

Research Institute on Secondary Education Reform (RISER). (2001). *Standards and scoring criteria for assessment tasks and student performance*. Madison, WI: Author.

Resnick, L. B. (1987). The 1987 presidential address: Learning in school and out. *Educational Researcher, 16*(9), 13-20, 54.

Rossman, G. B., & Rallis, S. F. (2003). *Learning the in the field: An introduction to qualitative research*. Thousand Oaks, CA: SAGE Publications.

Rubin, B. C., & Justice, B. (2005). Preparing social studies teachers to be just and democratic: Problems and possibilities. In N. M. Michelli & D. L. Keiser (Eds.), *Teacher education for democracy and social justice* (pp. 79-103). New York: Routledge.

Rugg, H. O. (1921/1996). Reconstructing the curriculum: An open letter to professor Henry Johnson commenting on committee procedure as illustrated by the report of the joint committee on history and education for citizenship. In W. Parker (Ed.), *Educating the democratic mind* (pp. 45-60). Albany, NY: State University of New York Press.

Sandelowski, M. (2001). Real qualitative researchers do not count: The use of numbers in qualitative research. *Research in Nursing & Health, 24*(3), 230-240.

Sandelowski, M., Voils, C. I., & Knafl, G. (2009). On quantitizing. *Journal of Mixed Methods Research, 3*(3), 208-222.

Sarason, S. (1971). *The culture of the school and the problem of change*. Boston, MA: Allyn and Bacon.

Schalock, H. D., Schalock, M., & Girod, G. (1997). Teacherwork sample methodology as used at Western Oregon State College. In J. Millman (Ed.), *Grading teachers, grading schools* (pp. 15-45). Thousand Oaks, CA: Corwin Press.

Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 118-137). Thousand Oaks, CA: SAGE Publications.

Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 189-213). Thousand Oaks, CA: Sage.

Sears, A., & Parsons, J. (1991). Towards critical thinking as an ethic. *Theory and Research in Social Education, 19*(1), 45-68.

Shepard, L. (1989). Why we need better assessments. *Educational Leadership, 46*(7), 4-9.

Silva, E. (2009). Measuring skills for 21st-century learning. *Phi Delta Kappan, 90*(9), 630-634.

Sleeter, C. E. (2008). Teaching for democracy in an age of corporatocracy. *Teachers College Record, 110*(1), 139-159.

Smith, G. (2002). Are there domain-specific thinking skills? *Journal of Philosophy of Education, 36*(2), 207-227.

Splitter, L. J. (2009). Authenticity and constructivism in education. *Studies in Philosophy and Education, 28*(2), 135-151.

Stake, R. E. (1994). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236-247). Thousand Oaks, CA: SAGE Publications.

Stemler, S. E., & Tsai, J. (2008). Best practices in interrater reliability: Three common approaches. In J. W. Osborne (Ed.), *Best practices in quantitative methods* (pp. 29-49). Los Angeles: Sage Publications.

Strauss, A. L., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273-285). Thousand Oaks, CA: SAGE Publications.

Tannenbaum, R. J., & Katz, I. R. (2008). *Setting standards on the core and advanced iSkills assessments*. Princeton, NJ: ETS.

Tanner, D. E. (2001). Authentic assessment: A solution or part of the problem? *High School Journal, 85*(1), 24-29.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches* (Vol. 46). Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social & behavioral research*. Thousand Oaks, CA: Sage Publications.

Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 3-50). Thousand Oaks, CA: SAGE Publications.

Ten Dam, G., & Volman, M. (2004). Critical thinking as a citizenship competence: Teaching strategies. *Learning and Instruction, 14*(4), 359-379.

Terwilliger, J. S. (1997). Semantics, psychometrics, and assessment reform: A close look at "authentic" assessments. *Educational Researcher, 26*(8), 24-27.

The Academy for Educational Development. (2002). *National writing project: Final evaluation report*. New York, NY: Author.

The National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, D.C.: Author.

The State of Queensland Department of Education and the Arts. (2004). *New Basics research report*. Brisbane: Author.

Tye, B. B. (1987). The deep structure of schooling. *Phi Delta Kappan, 69*(4), 281-284.

van't Hooft, M. (2005). The effect of the "Ohio schools going solar" Project on student perceptions of the quality of learning in middle school science. *Journal of Research on Technology in Education, 37*(3), 221-243.

Van de Walle, J. (2007). *Elementary and middle school mathematics: Teaching developmentally* (6th ed.). Boston: Pearson Education, Inc.

VanWynsberge, R., & Khan, S. (2007). Redefining case study. *International Journal of Qualitative Methods, 6*(2), 80-94.

Wagner, T. (2008). *The global achievement gap: Why even our best schools don't teach the new survival skills our children need - and what we can do about it*. New York: Basic Books.

Walser, N. (2008, September/October). Teaching 21st century skills. *Harvard Education Letter, 24*.

Walters, K. S. (1994). Introduction: Beyond logicism in critical thinking. In K. S. Walters (Ed.), *Re-thinking reason: New perspectives in critical thinking* (pp. 1-22). Albany, NY: State University of New York Press.

Westheimer, J., & Kahne, J. (2003). Reconnecting education to democracy. *Phi Delta Kappan, 85*(1), 9-14.

Wiggins, G. (1989a). Teaching to the (authentic) test. *Educational Leadership, 46*(7), 41-47.

Wiggins, G. (1989b). A true test: Toward more authentic and equitable performance. *Phi Delta Kappan, 70*(9), 703-713.

Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Upper Saddle River, NJ: Pearson.

Williams, R. L. (2005). Targeting critical thinking within teacher education: The potential impact on society. *The Teacher Educator, 40*(3), 163-187.

Winn, S. S. (2004). The high cost of uncritical teaching. *Phi Delta Kappan, 85*(7), 496-497.

Wolcott, H. F. (1992). Posturing in qualitative inquiry. In M. D. LeCompte, W. L. Millroy & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 3-52). San Diego, CA: Academic Press, Inc.

Wolcott, H. F. (1999). *Ethnography: A way of seeing*. Walnut Creek, CA: AltaMira.

Woodruff, P. (2005). *First democracy: The challenge of an ancient idea*. New York: Oxford University Press.

# Appendix A

## QCS Interview Protocols

*Note:* All interview protocols were developed by the Boston College QCS research team.

### Interview 1 - Personal History and Education Experience

**Background: Educational experience**
Let's begin our conversation by talking about what brings you here to BC.

1. Why did you choose BC for graduate school? What do you hope to learn about teaching while you are here?

**Probe:** What are your expectations for the program and learning environment at BC? What do you think the program will offer?
**Probe:** How long has it been since you graduated from undergraduate college? What have you been doing since graduating?

2. Describe your college education? Where did you go? Why? What was your major in college? Why?

**Probe:** What incidents or experiences stand out during your college years? For example, were you active on student organizations or political activities on campus?
**Probe:** Did you work through college and/or did you have financial aid?

3. Describe your past school experiences.
    A.  Let's start with your secondary school experience.

**Probe** for context—was it a small or large school; an urban or suburban, parochial–single sex? Would you say it was diverse? If so, how?
**Probe**: What was the school like at the time you were there? For example, some people were in school during times of major change, such as during school integration, the merging of two high schools, or witnessing a shift in population in community, leading to increased diversity in the school, OR there were also some local changes such as a new teacher or administrator, a different tracking or grouping system, or a change in courses.

    B.  Now tell me about your elementary school experience.

**Probe** for context—was it a small or large school; an urban or suburban, parochial–single sex? Would you say it was diverse? If so, how?
**Probe**: Again, what was the school like at the time you were there?

 4. How did you experience school as a student?

**Probe** for their experiences as learners-- So if an individual responds about the social aspects of schooling, ask them how they experienced school as learners?
**Probe**: What was your most memorable experience? Were you involved in extracurricular activities? If so, what type of activities were you involved in?

5. Now, I want to switch topics a bit to talk about what brings you to teaching. When did you first start thinking you might want to teach? Why are you interested in teaching?

**Probe**: Did you consider becoming a teacher while you were an undergrad? Why or why not?
**Probe:** for their intellectual interests and the perspective they hold as a student. For instance, many of the elementary candidates mention their love of reading and children. Try also to discover what the person especially enjoys about school or about learning.

6. You're planning to teach _____ (elementary or high school) is that right? When you think back to your own experience in _____(elementary or high school), what stands out to you?

**Probe:** for specificity: What do you mean? Can you give me an example of that? Is there anything else you remember?

If the teacher candidate does not mention one of the following: You haven't mentioned (much about)_____.  Do you remember anything in particular about that?"
- what you learned
- your teachers
- how you felt about different subjects

**Probe (Elementary folks):** How do you think an individual best learns to read or to write?
**Probe (Secondary folks):** How do you think an individual best learns _____ (history, English, science, math)?
**Probe:** Do you think you received a good education?  Why or why not?

**Background:  Beliefs:**
7. A part of our research focuses on individuals' ideas, beliefs and experience as they relate to teaching and learning. At BC, one of the stated purposes is to prepare individuals to teach for social justice.  What does that mean to you?

**Probe A:** If teacher candidate says that he/she does not know what teaching for social justice is, move on to question 9.
**Probe B**. If teacher candidate gives an answer to the social justice question, ask: So, how do you think that plays out in _____ (reading or math: elementary folks) or (history, English, or science: high school folks)?

8. As you think about your future profession, what do you believe is/are the role(s) of the teacher?

**Probe:** Think of a teacher you have known.  Are there things you admired about this teacher? Things you would like to have changed?
**Probe:**  From your perspective, what are the top two or three challenges that teachers face today?

**Background: Knowledge**
9. Now, think about the content areas you will be teaching as an elementary or high school teacher.  What do you think are your strengths and weaknesses in the content area(s) you might have to teach?

**Probe:** What are you hoping the BC program will provide in terms of your preparation?
(**Note:** This can focus on fears and concerns if it hasn't been covered OR it can be skipped if it was thoroughly discussed.)

**Probe:** Now think about the range of things a teacher does. What might be your strengths? What areas might you need support?

**Background: Practice (Future plans)**
10. What are you looking forward to in your Student Teaching Practicum? Is there anything you are concerned about? What challenges do you think you will face?

**Probe:** How will you prepare yourself for these challenges?

11. When you think about next year, where do you see yourself working? Where would you like to teach?

**Probe:** Talk to me about what you hope your classroom will be like? How will you teach? What will your relationships with students, faculty, and parents look like?

12. In conclusion, we'd like to get some information about your background, especially your demographics. (**Note**: Make references to prior responses to pull pieces together. Continue probing so we don't receive a mere list.)

**Probe:** For example: your age, race, ethnicity, cultural background, language, religion and political orientation?

Closing Remarks:
    Is there anything else you'd like to share that we didn't cover?

(Thank the participant!)

---

### Interview 2: Pre-practicum Experience

The focus of this interview is on your pre-practicum experience. We will meet again in January to talk more about your coursework at BC in the first semester. For this interview, I would like to learn about how your pre-practicum went, what you learned, what you struggled with, what impact the experience has had on your ideas about teaching, etc.

**Practicum Experiences**
1. Let's talk about your practicum. Describe a typical day at your practicum.

**Probe**: How have you found the structure of the pre-practicum?
**Probe:** What is your role in the classroom?
**Probe:** What is the school environment and community like?
**Probe:** Is the environment different from other places where you've been a student or volunteer/aide?
**Probe**: Do you observe teachers teaching in all subject areas (for elementary)?

2. Tell me about you Cooperating Teacher? (Age, Race, Ethnicity, years teaching, teaching style, etc.) What is the role of the cooperating teacher in shaping your practice and philosophy?

**Probe:** Would you describe a particular lesson you observed that was note worthy? Why?
**Probe:** How do you think your CT knows what to do next?

**Probe:** How do you think your CT knows if the kids are learning?
**Probe:** What types of classroom assessments does your CT use? Formative/summative? In what ways do assessments reflect the instruction?
**Probe**: Every teacher has strengths and weaknesses; can you tell me about those with regard to your Cooperating Teacher? Are there things you have observed and would do/wouldn't do? (specific content areas)
**Probe**: Do you and your Cooperating Teacher have similar teaching philosophies? Explain. (N.B. You want to understand what the teacher candidate's teaching philosophy is—skip if you have gotten at this in Question 2)
**Probe**: Do you think your Cooperating Teacher has the same ideas about teaching and learning as your BC Professors? Why or why not? Do you consider this a problem?
**Probe**: What advice have you gotten from your Cooperating Teacher? How has your Cooperating Teacher helped you in understanding teaching? How has he/she helped your understanding of pupil learning?

3. OK, let's move from your CT to your Supervisor; tell me about your Supervisor? (Age, Race, Ethnicity, years teaching, teaching style, etc.) What is the role of the Supervisor in shaping your practice and philosophy?

**Probe**: What advice have you gotten from your Supervisor? How has he/she helped you in understanding teaching? How has he/she helped your understanding of pupil learning?
**Probe**: What would you say are your Supervisor's strengths and weaknesses?
**Probe**: Do you and your Supervisor have similar teaching philosophies? Explain.
**Probe**: Do you think your Supervisor has similar ideas about teaching and learning as your BC Professors? Why or why not? Do you consider this a problem?
**Probe**: So, I understand that all of the pre-pracs in this school meet together with the supervisor at the school once a week? How's that been?

4. So we've talked about all the grown-ups…the other important people here are the kids. Tell me about the Students in the classroom?

**Probe:** What is their role in shaping your practice and philosophy? (Ask about the child study pupil if relevant)
**Probe:** Diversity (ELLs, SPED, SES, Ethnicity)? How would you describe their experience in school? Do they enjoy it? Why or why not?
If elementary: How is the weekly read aloud going with your ELL pupil?
**Probe**: Tell me about the lessons you taught. How did they go? What did you learn? (Insert here a question about something you observed in a classroom. For example, a unique method, approach, visual aide).
**Probe:** Some people say the most important thing about any lesson is whether the kids are learning. What do you think they learned? How do you know?
**Probe**: What are you learning about how children learn? How does this influence your perspective on the role of a teacher?
**Probe:** Can you describe a particular learning moment you observed that was note worthy? Why?
**Probe**: What advice have you gotten from your pupils? How have the pupils helped you in understanding teaching? How have they helped your understanding of pupil learning?

## Overall Questions
5. Have you observed examples of teaching for social justice in your pre-practicum experience? Please describe them.

6. Are you making connections between what you're learning at BC and what you're experiencing in your practicum?

7. Based on your pre-prac experience, what would you say are the most important skills and knowledge for teaching?

8. How have your practicum experiences thus far influenced your ideas about teaching?

**Probe:**  Based on the practicum, have you changed your plans on where and how you'd like to teach? Explain.

---

**Interview 3**

**2005 Summer & Fall Courses**

**Please fill table before interview.**

| Methods Courses | Foundations Courses | Content Courses |
|---|---|---|
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Last time we met we focused on your pre-practicum experience.  Today's topic is your coursework so far at BC.

**General Course Experiences**
1. Generally, how have your courses gone so far?

**Probe:**  What have you enjoyed about these courses so far?
Have there been any surprises?"
**Probe:**  Can you give me some examples of anything that has been particularly interesting or helpful?

2.  Foundations courses are generally used to give people the broad overviews of learning and schooling: broader contexts of children, schooling, and curriculum. Did you find the courses to be valuable in terms of providing that?  In what ways? **(Specify what courses we are referring to)**

**Probe:** Do you think the foundations courses helped you understand the realities of schools today?

3.  Methods courses are intended to prepare you to gain strategies to teach specific subjects.  What skills and knowledge did you acquire from your methods courses?  (Examples?)

**Probe:**  Did they meet your expectations?  If not, how might they have better met your expectations?
**Probe:**  Some people say the most important thing to learn is classroom management.  Do you agree?
**Probe:**  How did the methods courses help your knowledge of the content?
**Probe**:  Often a lesson in a methods class will demonstrate a teaching strategy which also includes content material.  Did these "model lessons" increase your understanding about the content (e.g., looked at content from new perspective, etc)?  Were they equally helpful for both strategy and content?

**Elementary**—How did the methods courses relate to each other?

(e.g. math, science, literacy, and social studies)

**Secondary**—Have you taken any courses in Arts & Science?
Was the course valuable to you in terms of pedagogy, broadening content knowledge, curriculum, and assessment?

**Probe:** What have you learned about bilingual students? Students with special needs?

4. Now let's talk about the teaching in the methods course? How would you characterize your methods professors' approaches to teaching?

**Probe** Do you think they modeled the kind of teaching they advocated (practiced what they preached)?
**Probe**: Do you think the faculty structured their courses around the realities of schools today?
**Probe**: Did the methods faculty explicitly address issues of social justice? If so, how?
**Probe**: What did you learn about pupil learning? (ways of learning, etc…)
**Probe**: What did you learn about assessment? (ongoing/formative & high-stakes; pupil learning)

5. You said you were hoping to learn about_____, has that been the case? Are there any gaps that remain in your coursework?

Overall Questions
6. Are you making connections between what you're learning at BC (methods, & foundation courses) and what you experienced in your pre-practicum? How? Examples?

7. When we first talked in the summer, I asked you a question about your definition of teaching for social justice. How do you see it now?
Has your definition changed? If so, why?

---

**Interview 4 with Participants: Full-Practicum Experience**

**1. Let's talk about your practicum.**

**Probe**: What's the school environment and community like?
**Probe**: What pressures and issues do teachers face in the school? What pressures do students face? (e.g. test scores, safety, race issues, etc.)
**Probe**: How are student teachers viewed? What's your relationship to other colleagues in the school?
**Probe**: How have things changed from your pre-practicum? (if relevant)

**2. What's your role in the classroom?**

**Probe**: How much teaching have you done so far? What have you been teaching? What haven't you been teaching?
**Probe**: Do you have any other responsibilities? How much freedom have you had in what and how you teach?
**Probe**: How are you approaching planning? Are you co-planning?

**Only if the participant has a new CT:**
**3. Tell me about your cooperating teacher? (race, age, ethnicity, years teaching, teaching style, etc.)**

**Probe**: What are you learning from her/him?
**Probe**: How do you think your cooperating teacher knows students are learning?
**Probe**: What types of assessments does your cooperating teacher use (formative, summative?)?
**Probe**: In what ways do assessments reflect the instruction?
**Probe**: Do you and your CT have similar teaching philosophies?
**Probe**: Do you think your CT has the same ideas about teaching and learning as your BC professors? Why/why not?  Do you consider this a problem?
**Probe:** Has your CT helped you improve social justice and/or equity in your teaching?

**4.  Tell me about your clinical faculty supervisor?  Is s/he different from the person you had for your pre-practicum (race, age, ethnicity, years teaching, teaching style, etc.)?**

**Probe**: What role is your supervisor playing in your practicum experience? (mediator, moral support, academic advice and content support)
**Probe**: What does your supervisor focus on in her observations and feedback? *(if nothing, remember to ask about classroom management?)*
**Probe**: Has s/he helped you provide strong academic content?
**Probe**: How has s/he helped you help pupils to learn?
**Probe:** Has your supervisor helped you improve social justice and/or equity in your teaching?
**Probe:** Do you and your supervisor have the same approach to teaching practices?
**Probe:** Do you think your supervisor has the same ideas about teaching and learning as your BC professors?  Why/why not?  Do you consider this a problem?
**Probe:**  I understand that the BC full practicum students in this school meet as a group with the supervisor once a week.  How has that gone?  What kinds of issues have you discussed?
**Probe:**  What are the other ways that you and your supervisor communicate about the classroom teaching experience? *(ask this if it's not touched on earlier in the interview)*

**5.  We've talked about the adults; the other important people are the kids. Tell me about the students in your classroom(s).**

**Probe:**  What are you learning from the students about being a teacher?
**Probe:**  What is the diversity in the classroom?  (ELLs, SpEd, Ethnicity?)  What's that have to do with what and how you teach?
**Probe:**  How do you think the kids in your classroom would describe their experience in the school?
**Probe:**  How has your relationship changed with the kids over the course of the year?
**Probe:**  In general, do you think the kids in the classroom are learning? What evidence do you have that they're learning?
>    **Probe:**  Now, let's talk about your teaching in relation to the students.  I noticed that you…. (Insert something here that you noticed from their classroom:  about a particular student, a group of students, a unique method, etc.)

**6.  In your own classroom and in the school, either in what you are doing or what the teachers are doing, do you see examples of teaching for social justice?  In your own teaching, how are you addressing issues of equity and justice?**

---

**Interview 5: Pupil Learning**

*NOTE: Teacher Candidate needs to bring three sets of pupil work: a full class set of a cumulative assignment and two examples of tasks that led up to it. TCs also need to pick out one high, one medium, and one low example of pupil performance for the cumulative assignment. Finally, have the teacher candidate bring any rubrics she or he used to score these assignments, as well as any assignment description that the TC gave to the pupils.*

**The purpose of this interview is to see what you are thinking about pupil learning and how it relates to your own instruction. First, I will ask you a series of general questions about the assignments you brought, then we'll get into the specific student examples you have selected as high, medium, and low. Finally, I'll ask you talk about your inquiry project.**

1. First, let's take a look at the assignments you brought. As a way to walk me through this work, it might be helpful for you to start at the end with the cumulative project and work backwards. Or you might want to start with the first task and move chronologically to the end, the cumulative task.

Probe: How does it fit into a larger unit?

Probe: Was this something you devised yourself?

Probe: Was any part of this lesson from a preexisting lesson that you adapted?

Probe: Why did you decide this lesson/assignment/assessment would be appropriate? How much autonomy did you have in creating the lesson or assignment?

2. What did you want students to get out of this activity**?** How do you know whether or not students accomplished what you wanted them to get out of this activity/lesson/unit?

Probe: How did you evaluate these assignments (rubric, scoring, etc.)?

3. Is there anything you would change about this lesson or assignment or unit? What? Why?

4. Let's now look at your examples of a high, a medium, and a low-level response. How do these samples compare to the overall class?
Probe: Is this work representative of the class? Is this what you expected?

5. Did the students who completed these examples meet your expectations? Why or why not?
Probe: What might you do differently in the future for each of these students?

6. Why did you choose these three examples?
Probe: Tell me about these three students (SPED, ELL, Bilingual).

### *General Pupil Learning Ideas*

7. What do you do to address the range of abilities in your classroom?

8. How do you know if your pupils are learning? What counts as evidence for learning?

9. Of course, teachers are not just interested in their pupils' academic learning; they are also very interested in their social and emotional development. Do you see your students making progress socially and emotionally? Like what?

Probe: How do you know if pupils are making this kind of progress?

10. Are you able maintain high expectations when the pupils have a variety of learning styles and needs? If so, how? If not, why?

*The Inquiry Project*

11. What was your Inquiry Question? What did you collect as data for your question?

12. What important insights did you get from your inquiry project concerning pupil learning?

Probe: While doing your inquiry project, what surprised you about students' learning?

Probe: How will the results of your inquiry project influence your practice as a teacher?

13. What would you categorize as social justice insights? Why?

Probe: How will you incorporate these insights into your own teaching?

14. While it is unlikely you would jump right into an inquiry project as you start your first year of teaching, what inquiry skills do you imagine using in your classroom practice?

Probe: Do you see yourself doing a formal inquiry project again in the future?

---

## Interview 6 – End of Teacher Education

This is our last interview for the year, so it will include an overview of what you have learned through the year and the influences that have been most significant. We will also talk about your future plans and then, at the end of the interview, give you an opportunity to provide us with some feedback about the program.

First, we'll talk about the learning overview: Specifically, we'll be looking for information about how you may have changed personally and professionally, your understanding of the role of a teacher, about teaching and learning, and social justice – and the most important influences that have shaped this experience.

I. Learning

I'd like to start with a set of questions about what you learned during this year in your teacher education program…

1. You've been in schools for almost a year and have finished your full-time student teaching, Some people say they ended up learning as much about themselves as they did about students or teaching methods teaching during this period. What would you say you have you learned about yourself?
- As a Teacher?
- As a Learner?

2. What did you learn about teaching/the activity of teaching?  What's the hardest thing?  What's the easiest?  What most surprised you?


3. What has had the greatest impact on this learning?
(Probe: What about—depending on their answer—your practicum experience, teacher education courses, A&S courses, your peers?

We're going to shift the focus a bit here and talk about some of the themes and concepts that pervade the program:

Let's start with the idea of pupil learning.

4. What's the most important thing you'd say you've learned about teaching reading/mathematics (for elementary)? _____ (specific subject) for secondary)(be specific for secondary)?
- How/Where/From whom did you learn that? What was the biggest influence on your learning? Who or what played the biggest role? What role did the courses play?
- What have you learned about teaching about literacy in the elementary school?  Math?
- What have you learned about teaching bilingual students/ELLs? How/Where/From whom did you learn that?
- Which content areas do you feel the most/least prepared to teach?


All through BC's teacher education program, there's been a lot of talk about social justice.  We asked you about this in the first interview, as you might remember…

5. As you complete your teacher education experience, what do you make of this idea of Teaching for Social Justice?
- Has your definition changed?
- What impact did your practicum experiences have on your understanding of TSJ?

6. Did you have any strong models of teachers for social justice (either at BC or at your school site)?
- What made them good models?

7. How do you see yourself teaching for social justice in your own classroom?

8. Can you talk a bit about what you understand is the purpose of schooling?  Where has that been highlighted in your program?

II. Moving Forward/Your future:

Okay, let's look ahead, now.  In this section we'd like to talk about your future…
➢ What are you planning on doing next year (for benefit of the interview transcript)?
➢ Do you plan on teaching in the future?
➢ How has your experience in the past year impacted your career choice?

9. First, how is your job search going?
       Will you be around this summer?  Do I need to update contact information?
       Are you planning on taking part in BC's mentoring program?

10. When you imagine yourself teaching next year, what do you see?
- What will your classroom be like?

- What will be the biggest challenges?
- What do you expect to be most prepared for?
- How do you think MCAS and NCLB will influence your teaching?
- Professional goals as a teacher?

11. Do you think about teaching as a career? What do you see yourself doing in the next five years?
- Ten years?

III. Program Feedback

Finally, we'll give you the opportunity to tell us more specifically what you think about the BC program….

12. If you could change three things about the program, what would they be?
   Was there anything irrelevant in the program?

13. What three things would you keep, that you found especially valuable in the program?

---

**Interview 7 – November of first-year of teaching**

**Introduction:**
**Now that you've been in the classroom for a few months we're going to ask you some questions that brings us up to date on your school setting and students, how you've settled into teaching, return to a few familiar themes in our research, and then ask just a bit about the future.**
**We'll start with some general questions about your school and schedule.**

**Let's start with a look at the school itself, your students, and the people you work with:**

**1. Tell me about your school…how would you describe it?**
   Probes:
- What kind of resources do they have? Or lack?
- What are the population demographics?
- Are parents involved in the school?
- What kind of goals does the school promote? Is there a mission statement? If so, do both faculty and students buy into it?
- Is there anything major that has happened at the school (AYP problems, new principal, new curriculum they have to use, construction)
- Is this a very different setting from your prac experience(s)?

**2. Let's shift to your students for a bit. I'd like you to describe them to me. Can you start with some general demographics that describe the pupils in your class(es)?**
   Probes:
- Age, ethnicity, language backgrounds, SES
- SPEd
- ELL
- Range of abilities across the group(s)
- Did you get some of this information from teachers who had these students    previously? Did you have prior experience with any of these pupils?

- How would you describe classroom dynamics? Do you have difficulty with certain students or a particular class?
- What is the biggest challenge you have faced so far this year?

**3. "At this point in the school year, are you able to identify goals for your students?"**
    Probes:
- What do you want them to learn? (consider academic, social, and emotional possibilities, here)

**I'd like to return to a question that has been a theme throughout the interviews:**

**4. We talked about learning to teach for social justice many times last year. We are interested in the realities of how this plays out in practice.**

    Probes:
- Do you think about issues of social justice in your classroom?
- In your planning?
- Do feel that teaching for social justice is an explicit part of your classroom        experience at the moment?
- How might this be particular to the context of your school? Classroom?
- How practical is the BC emphasis on social justice for a novice teacher?
- Has your view on teaching for social justice changed over the first few months of        fulltime teaching? If so, how and why?

**5. We've talked about this before, but now that you're fully responsible for classes, I'd like to have you think about it again: How do you know your pupils are learning? Be specific about the way you get this kind of information …**
    Probe:
- Has this changed in anyway since your prac? If so, why?
- Has the inquiry played a role in how you look at your classes?

**6. How about the other adults in the school. What kind of relationships have you been able to develop with school faculty & staff?**
    Probes:
- Principal, department head, fellow teachers
- Is there a lot of interaction among faculty?
- Do you have the opportunity to co-plan or co-teach?

**7. Do you have an assigned mentor or participate in an induction program? If so, has this been a successful match?**
    Probes:
- Are there other people that might be seen as informal mentors or part of your      network of support – including friends and family outside of school?
- Did you attend Summer Start? Why or Why not? Describe your experience. Was        it valuable? How would you change the program?

**Let's spend a few minutes talking about your immersion into fulltime teaching.**

**8. In general, how do you feel things have gone in the past few months?**

**9. What is your workload like?**
    Probes:
- What is your schedule? When do you get in to school? What time do you leave?

- For secondary – number of preps?
- For elementary – breaks?
- Additional school duties (ex: study hall, cafeteria duty, extra-curriculuar activities?)

**10. Tell me about planning…when do you get to do this? How do you decide what to use? What to teach?**

Probes:
- What resources do you have? Use? Where are they from?
- Are you focusing on day-to-day planning or do you have a long-term plan to work       from?
- What strategies/resources have you utilized from your master's program?

**11. How did you plan for this topic that you assessed here (look at the pupil work that the teacher brings to the interview)?**

- Why did you choose to assess your students using this assignment?
- How would you change it if you were to do it again?

**12. Do you see yourself as having a great deal of autonomy in your classroom?**
> **(If teacher asks what you mean by 'autonomy' can say 'when some people talk about autonomy they refer to the role of standards, district mandated curriculum or exams, whether you feel you have a voice in deciding what is taught in your classroom)**

Probes:
Why/why not?
In what area do you have most/least autonomy?
Who or what influences your decisions in the classroom?
Is MCAS a driving force in what you do?

**Let's look at how well prepared you feel and what you attribute to the BC experience:**

**13. What did you feel prepared for? Not prepared for?**

Probes:
- Is there anything that you feel BC did not prepare you for?
- Is there any one thing that you feel especially well prepared for by the BC       program?
- Does your school provide support through PD for what you might not feel       prepared for?
- Where might you turn for additional support/knowledge?
- Do you feel prepared to work with the population of students in your classroom? (ELL, SED, etc)

**14. Is teaching what you expected it to be? Have your aspirations for a career in teaching changed?**
- Do you think you'll teach next year?
- In this school? For how long?

**15. Is there anything that we haven't touched on that you feel is especially important to include in this conversation?**

---

**Interview 8 –Cohort 2– February-March of first year of teaching**

*NOTE: Teacher needs to bring three sets of pupil work: a full class set of a cumulative assignment and two examples of tasks that led up to it, all from same student. Teacher also needs to pick out one high, one*

*medium, and one low example of pupil performance for the cumulative assignment. Finally, have the teacher bring any rubrics she or he used to score these assignments, as well as any assignment description that the TC gave to the pupils.*

**The purpose of this interview is to see what you are thinking about pupil learning and how it relates to your own instruction. First, I will ask you a series of general questions about the assignments you brought, then we'll get into the specific student examples you have selected as high, medium, and low. Finally, I'll ask you talk about your inquiry project.**

1. First, last time you were struggling with … (fill in here with something specific to your teacher; e.g. students not completing their homework; the discipline protocol at the school, etc.). How's it going now?

2. OK, let's take a look at the assignments you brought. As a way to walk me through this work, it might be helpful for you to start at the end with the cumulative project and work backwards. Or you might want to start with the first task and move chronologically to the end, the cumulative task.

Probe: How does it fit into a larger unit?

Probe: Was this something you devised yourself?

Probe: Was any part of this lesson from a preexisting lesson that you adapted?

Probe: Why did you decide this lesson/assignment/assessment would be appropriate? How much autonomy did you have in creating the lesson or assignment?

3. What did you want students to get out of this activity**?** How do you know whether or not students accomplished what you wanted them to get out of this activity/lesson/unit?

Probe: How did you evaluate these assignments (rubric, scoring, etc.)?

4. Is there anything you would change about this lesson or assignment or unit? What? Why?
5. Let's now look at your examples of a high, a medium, and a low-level response. How do these samples compare to the overall class?
Probe: Is this work representative of the class? Is this what you expected?

6. Did the students who completed these examples meet your expectations? Why or why not?
Probe: What might you do differently in the future for each of these students?

7. Why did you choose these?
Probe: Tell me about these three students (SPED, ELL, Bilingual).

*General Pupil Learning Ideas*

8. What do you do to address the range of abilities in your classroom?

9. You have already talked about how you looked for pupil learning in your cumulative assignment. How in general do you know if your pupils are learning? What counts as evidence for learning? (Connect to question two or it may sound repetitive)

Probe: Has this changed in anyway since your practicum? If so, why?

Probe: Has the inquiry project played a role in how you look at your classes/students?

10. What kind of grading or evaluating system do you use?  Are you happy with it?

Probe: To what extent do you have autonomy in this?  Are there school or department guidelines about grades?

11. What kind of pupil data does your school district use in developing curriculum & instruction that might impact your class?

Probe:  This might include MCAS scores; other standardized test scores;  testing coming from, or contributing to IEPs and 504s; Student Success Plans (these are required for students w/o IEP or 504 that don't meet standards on other tests); portfolio or exhibit projects, district benchmark/tests, other?

Probe: Do you have access to this data on an individual or aggregate level to make plans for your classes/pupils?

Probe: Would you be part of the data analysis?

Probe: Do you feel BC has prepared you to be able to use pupil data, both formal, informal, standardized and teacher-developed to make decisions in your classroom?  Do you do this?

12. Of course, teachers are not just interested in their pupils' academic learning, they are also very interested in their social and emotional development. Do you see your students making progress socially and emotionally?  Like what? *(Note:  levels of confidence, enjoyment of learning, engagement in learning, independence in learning, cooperative group work, classroom behavior, interpersonal interactions)*

Probe: How do you know if pupils are making this kind of progress? What evidence do you look for to determine social and emotional growth?

13. What kind of expectations do you have for students? Are you able maintain these expectations when the pupils have a variety of learning styles and needs?   If so, how?  If not, why?

14.  How do you help students develop language abilities?  (ELL, SpEd, Writing, Reading)

Probe: Would you call your classroom language-rich?  Why or why not?

***Experience in Classroom/School***
***Now let's touch base on how the year is going, now that you are about half-way through it.***

15. What kinds of changes, if any, have you made based on your experience in the first half of the year?

Probe: For example, grading, classroom management, differentiated instruction?

Probe: Are there disciplinary or management expectations school-wide? In your teaching team?

Probe: Do you find yourself using any techniques gained from BC? From your practicum?

16. How have you handled classroom management so far?

17. How is the larger school context/culture playing a role in your classroom?

Probe:  What contact have you had with the Principal/Dean/Mentor/Coach/etc.?  Are you satisfied with the amount and nature of your interactions?

Probe: Have you been observed and evaluated? By whom?  What kind of feedback have you received?

Probe: What contact have you had with parents?  What role do they play in the school?

18.  Are you participating in mentoring/induction?  If so, what kind?  Is it helping you professionally or personally?

Probe: Are there other people who might be seen as informal mentors or part of your network of support – including friends and family outside of school?

Probe: Are you attending any programs sponsored by BC? Are they valuable?  How would you change them?

19. Some people say the first year of teaching is the hardest and find it difficult to find balance.  How has your "quality of life" as first year teacher been so far? (Do you have a life?)

20. Do you see yourself working at the same school/in the same job next year?

Probe: If not, ask why.  What would it take for you to stay?

Probe:  If yes, ask what it is that is keeping them in the position.

---

**INTERVIEW 9 – COHORT 2 – End of fist-year of teaching**

**This is our last interview, so it will include an overview of what you have learned, the influences that have been most significant, your thoughts on teaching, and your future plans.  We will also talk about pupil work.**
*Remember to print out various charts, etc. before conducting the interview.*

**Pupil Learning**
1. What's the most important thing you'd say you've learned about teaching reading/mathematics (for elementary)? _____ (specific subject for secondary) over the last year?

Probe: How/Where/From whom did you learn that? What was the biggest influence on your learning? Who or what played the biggest role?

Probe: What have you learned about teaching about literacy in the elementary school?  Math?

Probe: Which content areas do you feel the most/least prepared to teach? How does this affect your teaching?

Probe: What's the most important thing you'd say you've learned about teaching diverse populations? (ELL, SPED, SES, etc.) – How/Where/From whom did you learn that?

2. OK, let's take a look at the assignments you brought.  As a way to walk me through this work, it might be helpful for you to start at the end with the cumulative project and work backwards.  Or you might want to start with the first task and move chronologically to the end, the cumulative task.

Probe: How does it fit into a larger unit?

Probe: Was this something you devised yourself?

Probe: Was any part of this lesson from a preexisting lesson that you adapted?

Probe: Why did you decide this lesson/assignment/assessment would be appropriate? How much autonomy did you have in creating the lesson or assignment?

3. What did you want students to get out of this activity**?** How do you know whether or not students accomplished what you wanted them to get out of this activity/lesson/unit?

Probe: How did you evaluate these assignments (rubric, scoring, etc.)?

4. Is there anything you would change about this lesson or assignment or unit? What? Why?

5. How do you feel your pupils did overall? Do you feel like they gained skills over the year? What? Were you satisfied/disappointed?

6. Let's now look at your examples of a high, a medium, and a low-level response. How do these samples compare to the overall class?
Probe: Is this work representative of the class? Is this what you expected?

7. Did the students who completed these examples meet your expectations? Why or why not?
Probe: What might you do differently in the future for each of these students?

8. Why did you choose these?
Probe: Tell me about these three students (SPED, ELL, Bilingual).

9. Our research group looked carefully at responses from last year's interviews that had to do with pupils' work and your assessments of their learning. We came up with graphic to try to explain what we found. The first box is supposed to represent teacher candidates' experiences during coursework, and the second what happened during student teaching. Overall we found that student teachers created great assessments that showed they had high expectations for pupils and focused on higher-order thinking. (refer to figure) We thought about this as "ownership" —student teachers actively changing strategies, questioning practices, and generally looking for better ways to improve learning in the classroom.
Does that sound to you like what was going on for you during student teaching? How about now, during your first year of teaching?

10. Another thing we found during the interviews when we asked teachers to talk about high-, medium-, and low-, pupil performance on the assessments, was that sometimes there was a kind of distancing. For example, if a pupil performed poorly on a test or a project, sometimes the student teacher attribute this to the pupil's lack of effort or his or her failure to pay attention and follow directions. This made us think a lot about how teachers make sense of it when pupils don't meet their expectations. Can you talk about this a little bit?

11. Do you think teachers should expect to meet the learning needs of every pupil in the class?

**Social Justice**
12. All through BC's teacher education program, there's been a lot of talk about social justice. We asked you about this in the first interview, as you might remember…As you are now completing your first year of

teaching, what do you make of this idea of Teaching for Social Justice? Is it important to you in your daily work? Do you consider yourself to be teaching for social justice?

13. Show them the 4 categories/28 codes for Social Justice (see end of interview for chart) and ask: We looked at all the responses of participants from the pre-service year and earlier this year about what it means to teach for social justice. Here is the way we grouped responses. What strikes you from this list? What's missing, if anything?

14. Some of the people who define TSJ say it's teaching that improves students' learning and enhances their life chances. They say that part of this is teachers trying to work with others to actively address inequities in the system. We didn't find much talk about activism or addressing inequities in our interviews. Any thoughts on this?

**School Context/Teacher Roles**
*Now we're going to switch gears and talk about your school.*
15. What opportunities has the school provided you in terms of what and how you teach?

Probe: Have you experienced any constraints? Are there things you've felt you couldn't do this year but wanted to?

Probe: In terms of what you brought with you from the BC program, are there things that were particularly helpful? Were there things that you didn't have an opportunity to implement?

16. What personal factors have made a difference in your teaching (background, education, personal experiences)? (i.e. knowing a second language having an impact on teaching ELLs)?

17. How would you describe the role you played in the school this year (e.g. with pupils, clubs, committees, with other faculty)? Do you see that changing next year?

18. What role have others in the school (colleagues, mentors, etc.) played in your life this year?

**Inquiry**
19. One of the goals at BC is to develop inquiry as stance – a way of thinking about and questioning what happens in your classroom, collecting data – through pupil work – and making decisions about practice based on that information. Can you give me an example of how you see this occurring in your classroom this year? Is this an important element of your practice?

20. Have you used the strategies you used in your BC inquiry project this year? Why? Why not?

**Future Plans**
*Dependent on their plans for next year:*
21. Why did you decide to stay at the school?
OR
Why did you decide to leave? What were you looking for in your new school?
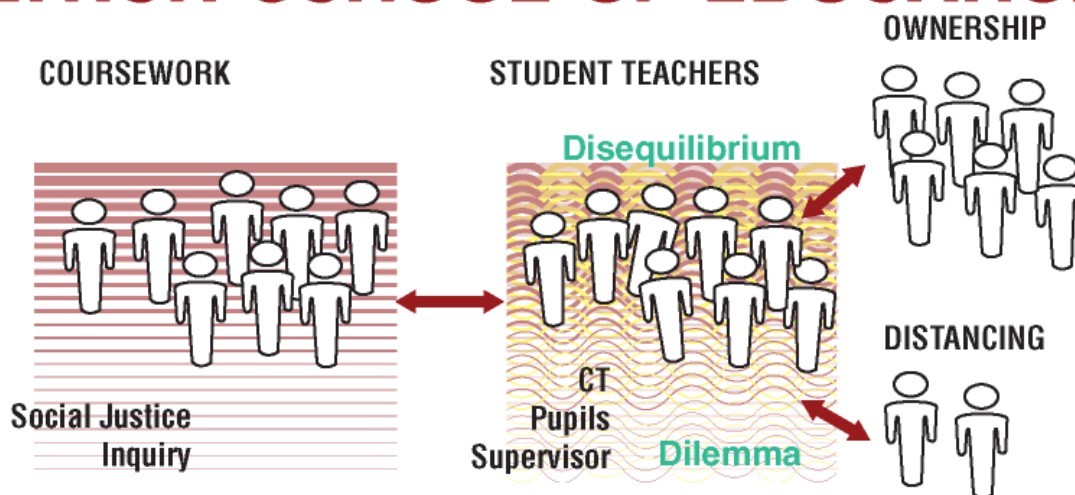AND
What aspects of this first year of teaching encouraged you to stay (or leave)?

22. Do you have any specific goals for next year? Have you thought about what you might keep the same and what you might change in your teaching, your classroom, and in your role in the school?

23. Do you think about teaching as your career? What do you see yourself doing in the next five years? Ten years?

# LYNCH SCHOOL OF EDUCATION

COURSEWORK  STUDENT TEACHERS  OWNERSHIP

Disequilibrium

Social Justice
Inquiry

CT
Pupils
Supervisor  Dilemma

DISTANCING

| Theme | Codes | Description (*Emphasizes…*) |
|---|---|---|
| Pupil Learning | 6 - Curriculum applicable | Teacher as making curriculum relevant and applicable to the pupils |
| | 9 - Accommodate/Differentiate | Idea of accommodating different learners and differentiating instruction |
| | 10 - Everybody learns | Teacher responsible for making sure pupils learn |
| | 11- Promote engagement | Importance of engaging pupils |
| | 13 - Multiple viewpoints | Importance of exposing pupils to multiple viewpoints; encouraging them to consider other perspectives, and expanding ideas and opportunities |
| | 14 - Critical thinking | Critical thinking and deep questioning |
| | 18 - Prepare future | Preparing pupils for a successful future |
| | 19 - Basic skills | Importance of teaching basic skills |
| | 22 - Social/cultural contexts | Knowing and understanding pupils' social and cultural contexts |
| | 23 - High expectations | Holding pupils to high expectations and pushing kids to meet those goals |
| | 24 - Same expectations | Holding same expectations for all pupils |
| Relationships and Respect | 12 - Be Fair | Being fair to all pupils in the classroom; not showing favorites |
| | 20 - Relationships pupils | Building relationships with the pupils |
| | 21 - Parents | Respecting and working with parents |
| | 25 - Culture of respect | Promoting a culture of respect among pupils and between pupil and teacher |
| | 27 - Care | Knowing and caring for pupils |
| Teacher as Activist | 1 - Collaborations/Coalitions | Importance of participating in collaborations/coalitions to support pupils and improve schools |
| | 2 - Advocate for pupils | Role of the teacher in serving as an advocate for pupils |
| | 3 - Activism | Idea that the teacher should participate in activism |
| | 4 - Community work | Role of the teacher in doing community work/volunteering or getting pupils engaged in such activities |

| Recognizing Inequities | 5 - Change agent | Teacher as a change agent, making a difference in society |
| | 7 - Challenge canon | Challenging the canon or altering the standard curriculum |
| | 8 - Gender | The role gender plays in the classroom |
| | 15 - Class/race struggle in Curriculum | How teachers might highlight class/race struggle and social inequities as part of the curriculum |
| | 16 - Connections to oppression | Ways to connect curriculum to real world examples of oppression and exploitation |
| | 17 - Break down barriers | Breaking down racial or SES barriers for pupils |
| | 26 - Challenge stereotypes | Challenging pupils' stereotypes or biases related to race, class, gender, or sexual orientation |

## INTERVIEW 10

**A. Tell me about your school…how would you describe it?**
Probes:
- What kind of resources do they have?  Or lack?
- What are the population demographics?
- Are parents involved in the school?
- What kind of goals does the school promote?  Is there a mission statement?  If so, do both faculty and students buy into it?
- Is there anything major that has happened at the school (AYP problems, new principal, new curriculum they have to use, construction)?
- Is this a very different setting from your last teaching experience?

**B. Let's shift to your students for a bit.  I'd like you to describe them to me.  Can you start with some general demographics that describe the pupils in your class(es)?**
Probes:
- Age, ethnicity, language backgrounds, SES *(How does this compare to last year?)*
- SPEd
- ELL
- Range of abilities across the group(s)
- Did you get some of this information from teachers who had these students      previously? Did you have prior experience with any of these pupils?
- How would you describe classroom dynamics? Do you have difficulty with certain students or a particular class?
- What is the biggest challenge you have faced so far this year?

*C. If the teacher is in the same school start with:*
- *Is there anything major that has happened at the school (AYP problems, new principal, new curriculum they have to use, construction) since last June?*
- *Is there any significant difference in your teaching assignment this year?*

*Then all interviews continue:*

**1.  In general, how do you feel things have gone in the past few months? How are things in comparison to last year?**

**2. What kinds of changes, if any, have you made based on your experience in the first half of the year or from last year?**
Probe: For example, grading, classroom management, differentiated instruction?

**3. At this point in the school year, are you able to identify goals for your students?**
Probes:
What do you want them to learn?  (consider academic, social, and emotional possibilities, here)

**4. How do you know your pupils are learning?**
Probe:
- Has this changed in anyway since last year?  If so, why?
- Has the inquiry played a role in how you look at your classes?

**5. Of course, teachers are not just interested in their pupils' academic learning, they are also very interested in their social and emotional development. Do you see your students making progress socially and emotionally?  Like what?** *(Note:  levels of confidence, enjoyment of learning, engagement in learning, independence in learning, cooperative group work, classroom behavior, interpersonal interactions)*

**6.  What is your workload like?**
Probes:
- What is your schedule?  When do you get in to school?  What time do you leave?
- For secondary – number of preps?
- For elementary – breaks?
- Additional school duties (ex: study hall, cafeteria duty, extra-curricular activities?)

**7.  Tell me about planning…when do you get to do this?  How do you decide what to use?  What to teach? How is it different from last year?**
Probes:
- What resources do you have?  Use?  Where are they from?
- Are you focusing on day-to-day planning or do you have a long-term plan to work         from?
- What strategies/resources have you utilized from your master's program?

**8. Do you see yourself as having a great deal of autonomy in your classroom?**
> **(If teacher asks what you mean by 'autonomy' can say 'when some people talk about autonomy they refer to the role of standards, district mandated curriculum or exams, whether you feel you have a voice in deciding what is taught in your classroom)**
Probes:
Why/why not?
In what area do you have most/least autonomy? *Has this changed since last year?*
Who or what influences your decisions in the classroom?
Is MCAS a driving force in what you do?

**9.  What kind of relationships have you been able to develop with school faculty & staff?**

Probes:
- Principal, department head, fellow teachers?
- Is there a lot of interaction among faculty?
- Do you have the opportunity to co-plan or co-teach?

**Let's look at how well prepared you feel and what you attribute to the BC experience:**

**10. After over a year as a full-time teacher, what do you feel BC best prepared you for? In what ways do you feel least prepared?**
        Probes:
- Pedagogy? Content-knowledge?
- Does your school provide support through PD for what you might not feel        prepared for?
- Where might you turn for additional support/knowledge?
- Do you feel prepared to work with the population of students in your classroom? (ELL, SED, etc)

**Now, I'd like to return to some questions that have been themes throughout the interviews, namely— pupil learning, social justice, and inquiry:**

**11. We've talked about learning to teach for social justice during other interviews. As you know, we're interested in the realities of how teaching for social justice is playing out in practice.**
        Probes:
- Do you think about issues of social justice in your classroom?
- In your planning?
- Do feel that teaching for social justice is an explicit part of your classroom        experience at the moment?
- How might this be particular to the context of your school? Classroom?
- How practical is the BC emphasis on social justice for a novice teacher?
- Has your view on teaching for social justice changed *over the last year?*

*Looking at Pupil Work*
**OK, let's take a look at the assignments you brought. As a way to walk me through this work, it might be helpful for you to start at the end with the cumulative project and work backwards. Or you might want to start with the first task and move chronologically to the end, the cumulative task**.

**12. How do these assignments fit into a larger unit?**
        Probe:
- Was this something you devised yourself?
- Was any part of this lesson from a preexisting lesson that you adapted?
- Why did you decide this lesson/assignment/assessment would be appropriate? How much autonomy did you have in creating the lesson or assignment?

**13. What did you want students to get out of this activity? How do you know whether or not students accomplished what you wanted them to get out of this activity/lesson/unit?**
        Probe:
- How did you evaluate these assignments (rubric, scoring, etc.)?

**14. Is there anything you would change about this lesson or assignment or unit? What? Why?**

15. Let's now look at your examples of a high, a medium, and a low-level response. How do these samples compare to the overall class?
        Probe: Is this work representative of the class? Is this what you expected?

16. Did the students who completed these examples meet your expectations? Why or why not?
   Probe: What might you do differently in the future for each of these students?

17. Why did you choose these?
   Probe: Tell me about these three students (SPED, ELL, Bilingual).

*General Pupil Learning Ideas*

**18. Has your grading system changed from last year? If yes, describe how it has changed.**

> *Ask this question if teachers is in new school context* **-** What kind of grading or evaluating system do you use?  Are you happy with it?
>   Probe:
> - To what extent do you have autonomy in this?  Are there school or department guidelines about grades?

**19.  Is your school doing anything differently with pupil data (MCAS, District exam scores) compared with last year?**

> *Ask this question if teachers is in new school context* - What kind of pupil data does your school district use in developing curriculum & instruction that might impact your class?
>   Probe:
> - This might include MCAS scores; other standardized test scores;  testing coming from, or contributing to IEPs and 504s; Student Success Plans (these are required for students w/o IEP or 504 that don't meet standards on other tests); portfolio or exhibit projects, district benchmark/tests, other?

**20. Do you use data for classroom inquiry?**
   Probe:
- Has inquiry played a role in how you look at your classes/students or pupil data?
- Have you used the strategies you used in your BC inquiry project this year?  Why?  Why not?

**21. Some people say the first year of teaching is the hardest and find it difficult to find balance. Would you say your "quality of life" has changed since the first year? (Do you have a life?)**

**22.  Is there anything that we haven't touched on that you feel is especially important to include in this conversation?**

---

**INTERVIEW 11 (End of second year of teaching)**

> **Introduction -** This interview has some familiar pieces, and one new section. There will be three parts: first questions about "big picture" issues in teaching; second, a look at student work; and third we'd like you to show us how you feel you've changed as a teacher over the past few years. So, let's begin with the questions.

**PART I. Big Picture Questions**

**1.  Now that you've been teaching for two years, what would you say are the key characteristics of a very good teacher?**

> **Probe**: In interview one you talked about teachers you admired and specifically mentioned…. (e.g. FOR LOLA, "YOU'RE A.P. BIO TEACHER WHO REALLY SHOWED HER PASSION FOR THE SUBJECT AND MADE THE STUDENTS IN HER CLASS REALLY LOVE IT TOO…)

> **Probe:** Are these still qualities that you would say are important after being in the classroom as a teacher? If not, how and why have your ideas changed?

**2. Massachusetts requires that novice teachers in public schools are provided mentoring/induction, but the reality is that that is very different from school to school.  In your case, you've had...** (e.g. FOR LOLA, LOTS OF SUPPORT IN YOUR FIRST YEAR AND VERY LITTLE MENTORING AND SUPPORT IN YOUR SECOND YEAR)  **How important has this been to you?**

> **Probes:** Was it an effective program of support?
> What elements were most helpful to you?
> Were outside factors (people/resources) more helpful?
> Any suggestions for change?

> **Probe:**  What ongoing support or professional development would be important to you in your third year in the classroom?  At one time you talked about expanding your knowledge of… (e.g. FOR ELSIE, KNOWLEDGE OF AMERICAN LITERATURE)

**3.  CONTEXT– The school you're in, the student population you teach, the larger community in which you work (that this happens in) – are often mentioned as important to learning to teach.  Can you talk about how these different elements (in your context) influence your learning in the profession, and your students' learning?  In the past, for example you've mentioned ….**

> *(Possible suggestions)*
> *Impact of SES*
> *Impact of nature of student population (bilingual pupils, SPED, etc.)*
> *Impact of high-takes testing*
> *Impact of administration*
> *Impact of support*
> *Impact of expectations*
> *Impact of parents*

> **Probe:** What do you think is working in your school? Why?
> **Probe:** What, in your opinion, is keeping the school from being a place that supports teacher and student growth?

**4.  Of course, as we've discussed, it is complex and sometimes challenging, but would you say at this point in your career you are teaching for social justice? If yes, in what ways? If not, in what ways not?**

> **Probe**:  **Early on you mentioned** (e.g. FOR ELSIE, EXPOSING PUPILS TO DIFFERENT POINTS OF VIEW)….**and in later interviews you also mentioned**… (CARING FOR STUDENTS AND SHOWING THEM THAT YOU WERE INTERESTED IN THEIR LIVES OUTSIDE THE CLASSROOM), **some people might add ideas like improving academic learning, focusing on critical thinking, developing social and emotional learning, or enhancing students' life chances (**only list ideas that the teacher did not already talk about in past interviews)- **Do these ideas play a role in your teaching? If so, how? If not, why?**

**5. You've been in the classroom for two years now, and it's clear that you know (the context of) your school. If you were in charge, what would you change?**

> **Probes:** Are there things you have already been working on? Are there things you think you might be able to work on in the future? What things do you think will be most difficult to change? Why?
>
> *(THESE ARE EXAMPLES OF THINGS THAT COULD BE ACTED ON IF THEY NEED A NUDGE – COULD SHOW THE LIST TO PROVIDE TOPICS CHOICES)*
> *Expectations (for teachers and students)*
> *Opportunities*
> *Curriculum*
> *Availability of resources*
> *Tracking*
> *Emphasis on certain outcomes*

**6. As you begin to think about next year, what are your big picture goals for your students?**
> **Probe:** What is it you want your students to know and be able to do in (math, ELA, history, science, etc.)
> **Probe:** Is this different from last year, or the year before? (this also relates to whether they're teaching the same kids…)
> **Probe:** Will you adjust practice to achieve these goals? How? Why?

**7. Some, but of course not all, of the big challenges of learning to teach include successful classroom management, planning curriculum, developing pedagogy for teaching, meeting the needs of diverse learners, and assessment. Where do you see your strengths after two years? Are there areas that still need attention?**

> **Probes**:
> How do you expect to grow as a teacher in the next few years?
> How will you achieve these goals?
> What, if any, of these factors have changed the most in the last few years?
> How and Why?

**8. In early interviews, a number of our participants talked about teaching as a career. There are great rewards in influencing lives, sharing content that you are passionate about…and there are real drawbacks – pay, relative lack of respect for the profession, limited or no opportunities for advancement. How do you feel about teaching as a career at this point? What do you see as your career trajectory at this point?**
> **Probes:**
> Has this changed?
> Do you plan to stay in teaching?
> Are you more or less enthusiastic about teaching as a career choice than when you started?
>
> **Probe:**
> Do you plan to stay at this school next year? If not, where will you go? If yes, will it be the same position?
>
> **Probe:**
> Considering that teacher retention is such a big problem, from your experience, what do you think drives teachers from the profession?

**Part II- TAPL – Teacher Assessment / Pupil Learning**

**9. OK, let's take a look at the assignment you brought. Although we only have one assignment, it would be helpful if you could walk me through the larger unit it draws from. You could work backwards and describe the larger unit or you might want to move chronologically through the unit and describe the pieces that led up to this final assessment.**

**Probe:** How does it fit into a larger unit?
**Probe:** Was this something you devised yourself?
**Probe:** Was any part of this lesson from a preexisting lesson that you adapted?
**Probe:** Why did you decide this lesson/assignment/assessment would be appropriate? How much autonomy did you have in creating the lesson or assignment?

**10. What did you want students to get out of this activity? How do you know whether or not students accomplished what you wanted them to get out of this activity/lesson/unit?**
**Probe:** How did you evaluate these assignments (rubric, scoring, etc.)?

**11. Let's now look at your examples of a high, a medium, and a low-level response. How do these samples compare to the overall class?**
**Probe:** Is this work representative of the class? Is this what you expected?

**12. Did the students who completed these examples meet your expectations? Why or why not?**
**Probe:** What might you do differently in the future for each of these students?

**13. Why did you choose these?**
**Probe:** Tell me about these three students (SPED, ELL, Bilingual).

**Part III. Teacher Development Chart**
**14. Now we are going to move to a different part of the interview that provides you with an opportunity to talk about how your view your development as a teacher.**
**So if you look at this chart and the horizontal axis represents time from prior to being in a teacher education program through the end of the second year of teaching and the vertical axis represents development as a teacher, how would you chart your own development in a general way?**
>    **Probe –** If teacher asks 'What does development mean?' respond by turning it back to the individual 'We want to understand how you would interpret development.'
>    **Probe –** If the first probe is not needed, ask the teacher to explaining their understanding of development after they've completed their line.

**15. Okay now imagine we take your development and think about it in terms of 3 aspects: ---**
>    **Content knowledge** (Red)
>    **Pedagogy & practice (**Blue)
>    **Understanding the role of the teacher (**Green)

**Would you have three different lines? If so, how would you draw them?** (provide 3 different color markers (RED, BLUE, and GREEN) for drawing each line- be sure to reference the key on the blank development chart or the list above for the colors that correspond to the three aspects)

**16. Describe your lines on each chart.**
>    **Probe:** Why does the line drop here?
>    **Probe:** Why is there such a sharp increase in development at this point?

**17. How would you project the continuation of your line in the future?**

**Probe:** 5 years into teaching, 10, 25?

**18. Can you talk about your development toward becoming the best teacher you can be?**
        **Probe:** What, or who, has helped you along the way?  What circumstances might have has held you back?

---

## INTERVIEW 12 (Nov-Jan 3<sup>rd</sup> year of teaching)

*Questions 1, 2, and 3 only if it's a new school context:*

I. School Context Questions
**1. Tell me about your school…how would you describe it?**
        Probes:
- What kind of resources do they have?  Or lack?
- What are the population demographics?
- Are parents involved in the school?
- What kind of goals does the school promote?  Is there a mission statement?  If so, do both faculty and students buy into it?
- Is there anything major that has happened at the school (AYP problems, new principal, new curriculum they have to use, construction)?
- Is this a very different setting from your last teaching experience?

**2. Let's shift to your students for a bit.  I'd like you to describe them to me.  Can you start with some general demographics that describe the pupils in your class(es)?**
        Probes:
- Age, ethnicity, language backgrounds, SES *(How does this compare to last year?)*
- SPEd
- ELL
- Range of abilities across the group(s)
- Did you get some of this information from teachers who had these students        previously? Did you have prior experience with any of these pupils?
- How would you describe classroom dynamics? Do you have difficulty with certain students or a particular class?
- What is the biggest challenge you have faced so far this year?
- 

**3. If the teacher is in the same school start with:**
- Is there anything major that has happened at the school (AYP problems, new principal, new curriculum they have to use, construction) since last June?
- Is there any significant difference in your teaching assignment this year?

II. TAPL Section (only related to teacher task)
*Teachers are to bring 2 samples of culminating assignments, if possible from different subject areas or classes, i.e. in Elementary, a teacher might bring a math and literacy assignment and in Secondary, a teacher might bring an assignment from A.P. World History and Sophomore American History.*
**4. Describe the assignments. How do these assignments fit into a larger unit?**
        Probe:
- Was this something you devised yourself? How much autonomy did you have in creating the lesson or assignment?

- Was any part of this lesson from a pre-existing lesson that you adapted? Is this something you've used before?
- Why did you decide this lesson/assignment/assessment would be appropriate?

**5. What did you want students to get out of this activity? How do you know whether or not students accomplished what you wanted them to get out of this activity/lesson/unit?**
    Probe:
- How did you evaluate these assignments (rubric, scoring, etc.)?

**6. Is there anything you would change about this lesson or assignment or unit? What? Why?**

**7. What role do formal and informal assessments play on a day-to-day basis in your teaching? How has this changed over time?**

<u>III. Questions related to satisfaction and retention</u>

**8. Has your workload changed?**

**9. What is the school's investment in teachers?**
    Probe:
- What professional development options are available to you?
- What professional development have you participated in? How was it?
- What kinds of leadership roles have you played, could you play?
- How would you describe the collegiality of the faculty/staff?
- What are your professional and social relationships like?
- Does your school have teacher learning communities?

**10. How has teaching affected your personal life? How does your personal life affect teaching?**
    Probe:
- Has that changed over the three years?
- How have recent life decisions affected your career plans?

**11. What is the impact of the school culture on your commitment to teaching?**
    Probe:
- To what extent do you feel that people in your school are working collaboratively toward common goals?
- Some researchers have described three types of school cultures: the veteran culture in which nearly all the teachers have been at the school for many years, the novice culture in which many or most of the teachers are new to the school and to teaching, and what is called the integrated culture, in which there is a mix of veterans and novices. How would you describe the culture of your school? What influence has this had on your experience at the school and your desire to stay?

**12. Some people say that teaching gives them energy whereas others say it takes it away—so, whereas some feel energized by the practice of teaching, others feel drained by it. How would you describe your energy at the end of the day?**

    Probe:
- On your way into school, how are you generally feeling about getting to your classroom?

**13. Have there been any issues related to your gender that have affected your role as a teacher or your ideas about teaching?**

**14. Are you a better teacher now than you think you were at the beginning? In what ways? What do you think might be your next area/s for growth/learning?**

**15. Do you think your understanding of teaching for social justice has changed?**

      Probe:
- Are you attending to teaching for social justice more now? Moving away from it?
- Do you feel it's possible to teach for social justice in the context in which you work?

**16. How long do you see yourself teaching?**

      Probe:
- What else could you see yourself doing?
- What would keep you in the classroom?
- People who have studied teacher retention would describe you as a "stayer" which means you've stayed at the same school (or "mover" because you've stayed in teaching but moved to a different school). According to statistics which demonstrate that so many people leave teaching, how would you describe your experience as a "stayer"/"mover"?

# Appendix B

# TAPL Protocol

**TAPL – Teacher Assessment / Pupil Learning**

**1. OK, let's take a look at the assignment you brought. Although we only have one assignment, it would be helpful if you could walk me through the larger unit it draws from. You could work backwards and describe the larger unit or you might want to move chronologically through the unit and describe the pieces that led up to this final assessment.**

**Probe:** How does it fit into a larger unit?
**Probe:** Was this something you devised yourself?
**Probe:** Was any part of this lesson from a preexisting lesson that you adapted?
**Probe:** Why did you decide this lesson/assignment/assessment would be appropriate? How much autonomy did you have in creating the lesson or assignment?

**2. What did you want students to get out of this activity? How do you know whether or not students accomplished what you wanted them to get out of this activity/lesson/unit?**
**Probe:** How did you evaluate these assignments (rubric, scoring, etc.)?

**3. Let's now look at your examples of a high, a medium, and a low-level response. How do these samples compare to the overall class?**
**Probe:** Is this work representative of the class? Is this what you expected?

**4. Did the students who completed these examples meet your expectations? Why or why not?**
**Probe:** What might you do differently in the future for each of these students?

**5. Why did you choose these?**
**Probe:** Tell me about these three students (SPED, ELL, Bilingual).