# Using the Clinical Interview as a Complementary Assessment for Minority Elementary Students to Determine Their In-depth Understanding of Mathematical Concepts 

Nicola Elinor Crisp<br>Montclair State University

Follow this and additional works at: https://digitalcommons.montclair.edu/etd
Part of the Education Commons

## Recommended Citation

Crisp, Nicola Elinor, "Using the Clinical Interview as a Complementary Assessment for Minority
Elementary Students to Determine Their In-depth Understanding of Mathematical Concepts" (2013).
Theses, Dissertations and Culminating Projects. 24.
https://digitalcommons.montclair.edu/etd/24

[^0]
# USING THE CLINICAL INTERVIEW AS A COMPLEMENTARY ASSESSMENT FOR MINORITY ELEMENTARY STUDENTS TO DETERMINE THEIR IN-DEPTH UNDERSTANDING OF MATHEMATICAL CONCEPTS 

## A DISSERTATION

Submitted to the Faculty of Montclair State University in partial fulfillment of the requirements for the degree of Doctor of Education by NICOLA ELINOR CRISP<br>Montclair State University<br>Upper Montclair, NJ

2012

Dissertation Chair: Dr. Mark Weinstein

Copyright © 2012 by Nicola Elinor Crisp. All rights reserved.

# MONTCLARR STATE UNIVERSITY 

THE GRADUATE SCHOOL

DISSERTATION APPROVAL

We hereby approve the Dissertation
Using the Clinical Interview as a Complementary Assessment for Minority Elementary
Students to Determine their In-Depth Understanding of Mathematical Concepts
of
Nicola Elinor Crisp
Candidate for the Degree:
Doctor of Education in Pedagogy and Philosophy

Department of Educational Foundations
Dissertation Committee:


Dissertation Chair
Certified by:


Dean of The Graduate School


Dr. Ana Maria Villegas


#### Abstract

USING THE CLINICAL INTERVIEW AS A COMPLEMENTARY ASSESSMENT FOR MINORITY ELEMENTARY STUDENTS TO DETERMINE THEIR IN-DEPTH UNDERSTANDING OF MATHEMATICAL CONCEPTS by Nicola Elinor Crisp While some African American students perform as well as or better than their White peers on standardized tests, African Americans as a group attain lower scores on standardized tests than their White peers. This phenomenon has been addressed extensively in educational research. However, not much empirical research has been conducted to investigate whether a complementary assessment, such as the clinical interview, would provide more information about African American students’ mathematical knowledge than a standardized test.

Qualitative clinical interview methodology was used to explore the performance of the student participants on the clinical interviews and on the standardized test as well as teacher feedback about these students’ mathematical knowledge. Data were gathered from three main sources: clinical interviews, standardized test results, and teacher interviews. Data were coded to identify common themes that shed light on the research questions.

Most of the students performed better on specific items from the standardized assessment, the New Jersey Proficiency Assessment of State Standards (NJPASS), than on the clinical interview items. The scores on both assessments revealed much disparity among participants' mathematical competencies. The clinical interviews affirmed or


changed the teachers' opinions of their students. The information from the clinical interviews fostered discussion of pedagogical practices.

## ACKNOWLEDGEMENTS

"I can do all things through Christ who strengthens me." Phillipians 4:13 Thank you, God, for making it possible for me to complete this dissertation.

To my dissertation committee: Dr. Mark Weinstein - Chair, Dr. Ana Maria Villegas, and Dr. Alina Reznitskaya: I deeply appreciate your commitment of time and the care with which you provided excellent feedback. Your guidance and patience were invaluable to me.

To Brenda Sheehan: Thank you for always making the bureaucratic processes easy.

To Joan and Merrie: Thank you for introducing me to the clinical interview as well as for the subsequent collaboration.

To Barb, Deb, Diane, Kathleen, Lynn, Marisa, and Rochelle: Thank you for your encouragement and unstinting assistance. You made all the difference.

To Colleen and Scott: Thanks for the practical help that you provided.
To my doctoral support group, Mary, Winston, and George: Your intellectual and emotional support helped me get through the challenges of completing this dissertation. I look forward to our continued collaboration.

To my children, Craig and Nesa: Thank you for your love and your steadfast belief in my ability.

To my sister, Monica: Thank you for your unwavering support and thoughtful suggestions.

To my husband, Sydney: Your loving and pragmatic contributions to this endeavor have been present from its inception. I thank you for the myriad ways in which you have helped me to complete this project.

## TABLE OF CONTENTS

ABSTRACT iv-V
ACKNOWLEDGEMENTS ..... vi-vii
LIST OF TABLES ..... ix
CHAPTER 1 - STATEMENT OF PROBLEM ..... 1
CHAPTER 2 - REVIEW OF LITERATURE ..... 28
CHAPTER 3 - METHODOLOGY ..... 85
CHAPTER 4 - FINDINGS ..... 122
CHAPTER 5 - DISCUSSION AND SUMMARY ..... 270
REFERENCES ..... 302
APPENDICES ..... 313
A. Student Participants ..... 313
B. Teacher Participants ..... 314-316
C. Mathematical Concepts for Teacher Interview \#1 ..... 317
D. Teacher Survey \#1 ..... 318-320
E. Post Interview Sheet. ..... 321-322
F. Clinical Interview Protocol ..... 323-328
G. Guide for Viewing Clinical Interviews ..... 329-330
H. Comparison Scores for Clinical Interview \& NJPASS Items ..... 331-337
I. NJPASS Individual Performance Profiles ..... 338-341
J-L. Consent Forms ..... 342-356

## LIST OF TABLES

1. Specific Units of Analysis for Each Coding Category ..... 100
2. Common Coding Categories ..... 102-106
3. Problem Solving Rubric ..... 107
4. Themes and Explanations. ..... 110
5. Theme Comment Sheet ..... 111
6. Coding Sheet for Teacher Interview \#1 ..... 114
7. Coding Sheet for Teacher Interview \#2 ..... 115
8. Teacher Coding Categories ..... 116-117
9. Students’ Scores ..... 125
10. Summary of Changes in Teachers’ Attitudes ..... 230
11. Comparison of Challenging Items ..... 242
12. Summary of Correct NJPASS Open-Ended Items ..... 243

## CHAPTER ONE <br> Introduction

Throughout my career as an elementary school teacher in a suburban school district in New Jersey, I noticed a difference in the academic performances of the White students compared to the African American students in my classes. A substantial number of the African American students performed less competently than their White peers. This discrepancy was particularly evident in assessments of all types.

I would teach new concepts to both African American and White students at the same time. The African American students were attentive, participated orally in class, and usually worked hard on assignments. When I worked with them individually, they were usually able to find and correct their errors. Yet year after year, many of them earned lower scores on tests than their White counterparts.

The type of test had no apparent impact on this pattern. The results were the same whether the tests were teacher-created based on the district's curriculum, provided by textbook publishers, or were the state mandated standardized tests that have become so ubiquitous in our educational system. The grades of many African American students on these assessments were consistently lower than those of their White peers. In addition, boys often performed more poorly than girls. It was evident that the majority of the African American students were underperforming in most subject areas including mathematics.

This state of affairs was extremely worrying to me. For most of my career in this suburban district, I was one of two African American teachers in my school. Towards the end of my career in that district, I was the only teacher of color in that school. I found it deeply disturbing, professionally and personally, to witness and listen to the negative comments by a few of my colleagues about the continual underperformance of many, although not all, African American students.

At a personal level, I was the mother of a son who was often the only African American male in his high school honors courses. He received the highest verbal score on the PSATs in the school in his junior year. However, many of his peers, including a next-door neighbor, were mostly enrolled in remedial courses. Motivated by professional and personal reasons, I was inspired to find explanations and potential solutions for the discrepancies in academic performance exemplified by the students in my classes and their White peers, and by my son and his counterparts.

In an attempt to better understand these differences in academic performance, I examined the National Assessment of Educational Progress (NAEP) assessments. Results from NAEP for 2009 showed that African American fourth grade students scored 222 points on the mathematics test compared to a higher score of 248 for their White counterparts. Additionally, the statistics from the National Center for Educational Statistics (NCES) for the Black/White mathematics achievement gap for fourth graders from 2000 to 2009 reveal that while the test score gap has lessened since 2000, it remains quite high. In 2000, the gap was minus 31 (-31). By 2009, the gap had narrowed slightly to minus 26 (-26) in 2009 (NCES, 2009). These scores show that while the test score gap
between Black and White students has decreased, the difference between the scores of the two groups continues to be substantial.

Such empirical evidence confirmed that many African American students tend not to perform as well as their White peers on standardized tests (Flowers, 2007; Gardner \& Miranda, 2001; Haney, 1993; Ladson-Billings, 1994). These differences in test scores have serious implications for teachers.

According to Ferguson (2003), teachers’ perceptions, behaviors, and expectations can help to maintain and even increase the racial/ethnic test score gap. Their views of pupils positively or negatively affect those students’ academic performance (Baker, 1999; Crozier, 2005; Kuklinski \& Weinstein, 2001). As a result of this evidence, standardized test scores and teachers' opinions of their students based on these assessments became the foci of my study. Chosen for the sample were a small group of male African American students and their teachers. Clinical interview methodology allowed me to link the students in the study with their teachers and to see how these connections affected the findings. Outcomes yielded themes and patterns that helped to explain the link between the teachers and their students as well as the resulting level of student performance.

In this study I examined whether a complementary type of assessment, the clinical interview, would allow African American students to show their mathematical abilities better than they are sometimes able to do on mandated standardized tests. The study also explored what the students' teachers could learn about their mathematical abilities from the results of the clinical interviews. I sought answers that looked behind and beyond what the state assessment data provided. This search included examining prior efforts to
remediate the academic performance of African American students. The attempts at remediation appeared to be circuitous. By this I mean that while standards based reform initiatives designed to improve the academic achievement of minority students have been implemented in the United States (Darling-Hammond, 2007), many of these initiatives used standardized testing to assess and remediate African American students’ academic underperformance (Linn, 1995). Since this underperformance was largely determined by standardized test results, using the tests themselves as tools of remediation seemed circular. Nonetheless, widespread mandatory standardized testing has been an important aspect of educational reform in the United States for many years (Wang, Beckett, \& Brown, 2006).

Perhaps, the most notable of these reform initiatives is the No Child Left Behind (NCLB) legislation of 2001. It mandates that all public schools that receive federal funding will be required to show that all students are achieving at a proficient level in reading and mathematics on state standardized tests by 2014 (Linn, 2007; Ravitch, 2010). Unfortunately, many reform initiatives, including NCLB, have not had the desired effect of improving the academic achievement of African American students (Dillon, 2009; William, 2010). They have not eliminated the gaps in scores between African American and White performance on these tests (NCES, 2009). In fact, the disparities in scores between African American students and their White peers were less before NCLB was legislated than after it became law (Ravitch, 2010).

To conclude, my interest in the academic underperformance of many African American students has spanned decades and ultimately led to my conducting this study.

In the following sections, I discuss research on the academic achievement of these students. The first section is an exploration of the historical reasons for the general academic underachievement of African Americans. The second section is a discussion of the existing educational problems, with particular emphasis on how they impact minority students. Section three addresses the social and economic ramifications of the disparity in achievement between African Americans and their White peers. Section four presents the statement of the problem investigated in the study. In the subsequent sections, I present the research questions, provide information about the theoretical framework and methodology of the study, and give a brief summary of the findings.

## History

The many issues that affect the academic performance of African American students are extremely complex (Barton \& Coley, 2008; Lee, 2002). The roots of these issues have been present in our nation's history for centuries (Darling-Hammond, 2007). Haskins (1998) writes that the Black-White test score gap can be traced to historical inequities that have existed in the United States for generations. One such inequity, legal segregation, was part of this history for over 200 years. It started when the first Blacks in North America arrived in 1619 and culminated with the Civil Rights Movement of the 1950s and 1960s (Haskins, 1998).

Patterns of segregation that African Americans have historically faced effectively stymied their access to equal economic and educational opportunities (Haskins, 1998). In 1896, the United States Supreme Court legalized racial segregation on the nation’s railroads through the Plessy v. Ferguson ruling (Plessy v. Ferguson, 1896). Haskins
(1998) explains that the law was primarily intended to keep transportation separate but equal for Whites and Blacks. However, southern states approved a deluge of laws with the intent of enforcing the separation of the races in many other areas including the education of Black and White students. Blacks seeking better educational and economic opportunities in northern states often fared as poorly as their counterparts in the South (Haskins, 1998).

Kantor and Brenzel (1992) explain that after World War II, many Blacks and Hispanics flocked to northern cities in search of better economic opportunities. This influx of minority immigrants caused many White middle and lower class citizens to leave the urban areas. They fled to the suburbs not only to avoid living in close proximity to African Americans but also in response to the lure of home ownership. As a result, the cities they left behind were heavily populated by minority residents while the suburbs became predominantly White. The geographical division between these two disparate racial groups was fostered and perpetuated by prejudicial actions against minorities. In the suburbs, real estate covenants prevented white homeowners from selling their properties to potential minority buyers. Further, the Federal Housing Administration did not provide many loans to multi-racial communities. Instead, it preferred to support all White housing in suburban neighborhoods. Finally, the enforcement of discriminatory zoning conventions that barred minorities from living in certain communities successfully limited the ability of minority groups to live in most suburbs. In essence, they were only permitted to live in suburban areas that already had been abandoned by Whites and which were in decline (Kantor \& Brenzel, 1992).

The shift to the suburbs by predominantly White citizens did not only change the racial composition of these neighborhoods. It also had an enormous impact on the economic structure of the city neighborhoods they left behind. The residents of the suburbs earned higher incomes than the denizens of the inner cities. This discrepancy in earning capacity was further exacerbated by the fact that there were many more people living in poverty in urban than in suburban areas (Kantor \& Brenzel, 1992). Since schools are funded by local property taxes, the differences in income between the two groups influenced the quality of education they received.

As Kantor and Brenzel (1992) note, the racial and economic disparities and inequities between White and minority neighborhoods led to vastly differing opportunities and struggles for these two groups. Racial separation and income inequalities restricted the academic progress of low socio-economic and minority students. These inequities were increasingly reflected in inner city public schools that progressively became identified with poor academic achievement (Kantor \& Brenzel, 1992). This was the situation years after the country's judicial system had attempted to redress the widespread educational disparity between the races. In the effort to make academic access more unbiased, the nation's Supreme Court legalized equitable educational opportunities for its African American citizens in 1954. In that year, the landmark decision Brown v. The Board of Education of Topeka, Kansas, declared that separate but unequal educational opportunities for African Americans were illegal (Brown v. Board of Educ., 1954).

This watershed decision should have rendered educational inequality passé in the United States. By the late 1960s, laws providing Blacks with equal access to educational opportunities had been passed. However, due to endemic and institutionalized discrimination, many African Americans were unable to benefit from these new opportunities (Haskins, 1998). Darling-Hammond (2007) explains that this inability to fully utilize new opportunities was due to the fact that legally endorsed injustice had limited Blacks’ ability to receive the same quality education as Whites. Kantor and Brenzel (1992) note that comparisons of suburban and urban schools revealed many disparities, which were paramount to two separate and unequal school systems. After the Brown v. Board of Education ruling of 1954, many urban schools in the Northeast and the Midwest actually became more segregated (Kantor \& Brenzel, 1992).

The reasons for this regression can be partially attributed to subsequent appointees to the Supreme Court. In order to force compliance of the Brown v. Board of Education ruling from recalcitrant school districts, the 1964 Civil Rights Act imposed federally mandated punitive sanctions on any that defied the school desegregation law. As a result, school desegregation was strictly enforced leading to more integrated schools particularly in the South. However, this progress towards school integration was halted under Presidents Nixon and Reagan. Nixon appointees to the Supreme Court effectively dismantled desegregation efforts in addition to stalling equitable school funding practices. President Reagan's nomination of William Rehnquist as Chief Justice furthered the reversal of desegregation efforts. Rehnquist persistently voted against the
law that he held in acute disfavor. So despite previous progress, segregation regained its status in cities such as Birmingham and Atlanta (Frankenberg, Lee, \& Orfield, 2003).

Kantor and Brenzel (1992) explain that in many urban areas in the Northeast, Midwest, and West, middle class Whites, uncertain about the effects of court mandated school integration on their children, sent these students to private schools or moved to the suburbs. This led to the increased segregation of the inner cities that had already been negatively affected by enduring racial separations and income inequalities. In the 1974 Milliken v. Bradley ruling, the Supreme Court ignored civil rights challenges to the resegregation of inner city schools in the Northeast and Midwest by saying that suburbs could decide how their schools operated. The court further restricted desegregation efforts to urban areas, a decision that was pointless as these predominantly minority communities were already suffering from a dearth of White students. Since most of these students were also low socio-economic (SES) students, this meant that in addition to being increasingly Black, these schools would also become sites that housed more students who were from low-income families. In turn, this re-segregation proscribed the academic progress of low SES and minority students (Kantor \& Brenzel, 1992).

In contrast, as Frankenberg, Lee, and Orfield (2003) note, when schools were held accountable by very strict desegregation mandates they experienced more integration and less White flight. The period of desegregation contributed to more favorable racial perceptions especially in the South and was also the time when the achievement gap between the races dropped considerably. Schools that are racially separated tend to be poorer. Their students take less challenging courses and have teachers who are less
qualified. In schools that are more racially integrated, minority students are able to raise their levels of academic achievement despite the fact that their attainments are still below their White peers. Their chances of attending college and being successful in a multiracial work environment are enhanced (Frankenberg, Lee, \& Orfield, 2003). In these ways, a person's socio-economic class and background, his or her command of the English language, as well as where he or she resides, have tremendous impacts on the quality of their educational opportunities (Darling-Hammond, 2007).

Helped by the re-segregation of schools that occurred during the Nixon and Reagan administrations (Frankenberg, Lee, \& Orfield, 2003), the inequities in access to adequate schooling became ubiquitous features of the educational landscape of the United States (Darling-Hammond, 2007). As Flowers (2007) confirms, these inequalities are reflected in the Black-White test score gap. Race, poverty, and inner city education formed an unwelcomed trifecta in our nation's educational system (Kantor \& Brenzel, 1992). This trifecta caused urban school districts to be increasingly labeled as areas of low academic achievement as the educational attainment of the students in these schools was consistently poorer than that of their peers in the suburbs. As evidenced by school attainment and completion data, this discrepancy increased exponentially the more segregated and economically impoverished the students were. These students perform more poorly on achievement tests, and dropout rates in urban areas are higher than in suburban or rural areas particularly in the Northeast and Midwest (Kantor \& Brenzel, 1992).

In addition to legalized discrimination, segregation, desegregation, and resegregation, Jencks and Phillips (1998) posit other explanations for this persistent academic performance gap between African Americans and their White peers. They and other scholars believe that the achievement gap can be attributed to a combination of socio-cultural, and school related factors, including sub-standard school facilities, inequitable access to funding and educational opportunities, social class, and cultural differences between African-American students and their White teachers (Boykin, Tyler, Watkins-Lewis, \& Kizzie, 2006); Jencks \& Phillips, 1998; Roscigno, 1998).

In sum, historically legalized discrimination prevented African Americans from receiving equal educational opportunities, and these inequities left many students unprepared to benefit from access to more equitable education. The discriminatory practices that were endemic during segregation and re-segregation have resulted in a combination of factors that contribute to current educational inequities. As the previously cited research indicates, these issues have plagued the educational system of the United States for decades and continue to contribute to its present problems.

## Current Educational Problems

Before continuing this discussion, it should be emphasized that the words "poor" and "minority" are not synonymous. Thus, not all poor students are minorities and not all minorities are poor. Nonetheless, in many cases, students from low socio-economic backgrounds are also minority students. However, it is social class, perhaps more than race or culture, that influences children's success in school (Rothstein, 2004). "Much of the differences between the average performance of black and white children can
probably be traced to differences in their social class characteristics (Rothstein, 2004, p.4).

Educational research has shown that students from low socio-economic backgrounds suffer from a paucity of educational resources and opportunities (Hill, Guin, \& Celio, 2003; Kozol, 2005; Lareau, 2003; Rothstein, 2004). There are many examples that illustrate this scarcity of academic assets. Schools that enroll a predominantly African-American student population are hampered by numerous inadequacies. For example, the school buildings are often dilapidated and poorly maintained (Kozol, 2005). This situation is problematic as the condition of the structures in which students attend school affects their academic performance.

Poor building conditions are linked to a drop in student achievement. Differences in scores for students working in unsatisfactory buildings can range between 5 and 17 percentile points lower in comparison with students who are educated in satisfactory structures even after taking socio-economic status into consideration (Earthman, 2004). Substandard physical facilities have their causes in lack of adequate funding. Hill, Guin and Celio (2003) state, "...elementary schools in low-income areas nationwide receive from 10 to 30 percent less per pupil than higher-income schools in the same districts" (p. 55). Urban school systems experienced an erosion of their tax base as a result of white flight to suburbia. Additionally, due to increasing demands on their funds for other services, cities could not devote larger sums to education. This led to their inability to fiscally underwrite quality education for poor and minority students. Attempts to redress
fiscal inequalities in educational funding have largely been unsuccessful (Kantor \& Brenzel, 1992).

As Darling-Hammond (2007) observes, standards-based reforms have been lauded as the means of providing more equitable opportunities for minority students. However, most states have not provided sufficient resources such as adequate physical facilities and equitable funding to enhance learning opportunities. The result is diminished access to education for many students (Darling-Hammond, 2007). Lack of access has therefore been as troubling as need or scarcity in many schools located in poor communities especially those that house predominantly African American or minority students.

This difference in access to educational opportunities has a negative impact on the academic progress of students who are poor and of racial/ethnic minority backgrounds. Kozol (2005) explains that in urban neighborhoods, a large number of children have received no preschool education and enter kindergarten with limited social skills as well as minimal early learning skills. They may not know the correct way to hold a pencil, be able to identify colors and shapes, or realize that words on a printed page in English are read from left to right (Kozol, 2005). In agreement, Farkas (2003) states, "On average, African American children begin kindergarten or first grade with less oral language, prereading and pre-mathematics skills, as well as less general knowledge than that possessed by White and Asian children" (p. 2).

From birth to school graduation (that is, if they do not drop out), minority and poor children have been overwhelmingly exposed to situations in their educational
experiences that prevent them from achieving to the same degree as their White peers (Barton, 2003). It is therefore not surprising that children from lower socio-economic classes, even if they attend good schools, will not be as academically successful overall as their middle-class peers (Rothstein, 2004).

Academic underachievement is not only problematic for low-income African American students, but also for middle class African American students. They too often do not perform at the same level as their White peers (Ferguson, Clark, \& Stewart, 2002; Viadero, 2000). Evans (2005) notes that even belonging to a higher socio-economic class and thus having access to better educational opportunities sometimes fails to prevent African American students from experiencing lower academic performance. Some researchers like Ogbu and Simons (1998) posit psychological reasons to explain the lower level of academic performance of some African American students. In his thesis, Ogbu (1993) describes different types of minorities and classifies them into voluntary and involuntary groups. African Americans who were brought here unwillingly as slaves are designated as involuntary minorities. As such, unlike voluntary minorities who came here willingly, many African Americans do not have much faith in public schools that are controlled by Whites because of historic and current discrimination. They view behaving in ways that facilitate academic success in schools as "acting white" and as harmful to their cultural and minority identities (Ogbu, 1993). While involuntary minority parents do want their children to do well in school they are ambivalent about the means of achieving this success. "The double message that involuntary minority parents and communities send to their children is to do well in school, but be wary of your teachers,
school officials, and the curriculum because they are a part of white institutions that cannot be trusted" (Ogbu \& Simons, 1998, p. 13). Additionally, while they believe that education is necessary for success, due to their experiences with systematic discrimination in the labor market, they are not convinced that education is the path for getting ahead in the United States (Ogbu \& Simons, 1998). In response to this idea, the role models of many African American youth are not intellectuals or corporate professionals. Rather, the role models come from the ranks of entertainers or athletes as these pursuits are not dependent on education but on talent and physical strength (Ogbu \& Simons, 1998).

Psychological tensions where children from involuntary minority cultures view the adoption of White educational habits as a rejection of their own cultures and identities have been proffered as partial explanations for the lower academic achievement of African American students in middle class environments. In this way, psychologically based cultural conflicts regarding education are believed to contribute to the discrepancies in achievement that exist between the races in middle class neighborhoods. These variations in educational attainment lead some educators and social workers in school districts to believe that many intelligent African American students are not smart enough, lack the attitudes deemed appropriate for school, or belong to families who are unable to provide them with the requisite emotional and academic support (Landsman, 2004). Data from national testing sites (NCES, 2009) show that the professionals who make these adverse judgments are often members of the middle class. Further, they are usually white. In 2007-2008, 82 percent of elementary public school teachers in the

United States were White, compared with 7.4 percent who were Black. Additionally, 84.4 percent of teachers in these schools were female (NCES, 2009). This combination of Black students and White, female, middle-class teachers, sometimes creates cultural dissonance between the two groups (Boykin et al, 2006; Ladson-Billings, 1994; Villegas \& Lucas, 2002). However, these data do not mean that all White teachers are biased against their minority students. There are many White teachers who educate their Black students caringly and competently. Conversely, some Black teachers treat their minority students disrespectfully and are not invested in securing these pupils' academic success (Irvine, 1990). She further elaborates:

It would be unfair to imply that teachers are solely a product of their own cultural/racial reference group. Cultural identity as well as individual characteristics determine attitudes and behaviors. Although they do not have cultural and racial identities with Black students, some White teachers are effective with Black students because of their personal traits and teaching expertise. (p.49)

Nonetheless, other White teachers are less effective educators of minority students due to cultural differences. Crozier (2005) writes that as a result, the focus of some of these educators is to improve the behavior of African American students rather than to educate them. Consequently, the punishments they receive are often not commensurate with the degree of their infractions (Crozier, 2005). Ferguson (2001) concurs. She states that in many schools to discipline African American students means to mete out punishment. Harsh discipline by school authorities for even minor misdemeanors committed by these students is commonplace. This style of discipline robs these students of valuable time in class. Losing class time to suspension or expulsion worsens the academic futures of these students (Ferguson, 2001).

As a result, Rousseau and Tate (2006) find that "...these African American students are often overrepresented in the lower rungs of achievement distributions and underrepresented in the upper rungs" (p. 210). There is an insignificant amount of African American male students in academically challenging programs as opposed to those who are examples of "suspensions, expulsions, non-promotions, dropouts, and special education placements" (Garibaldi, 2007, p. 324).

In 2009, the drop out rate for Black high school students was 9.3 percent compared to 5.2 percent for White students (NCES, 2009). Lan and Lanthier (2003) state, "The high school drop out rate is a major educational problem for dropout students and for society" (p. 309). Darling-Hammond (2007) points out that with our economy increasingly requiring workers who possess high academic qualifications, the effects of dropping out are more detrimental than they have been in the past. While these effects are applicable to all students, they are particularly relevant for minority students (Darling-Hammond, 2007).

However, African American students do not only drop out of school because they, especially the males, are exposed to unfair disciplinary practices. They also drop out of school because they receive narrowed academic instruction that focuses primarily on test preparation for state mandated tests (Orfield \& Wald, 2000). As Mathis (2005) notes, "The poor schools, however, find themselves increasingly trapped into a dull and spiritless routine of drill and practice with the narrow objective of passing the examinations" (p. 592). In corroboration, Darling-Hammond (2007) asserts that some young citizens of the United States, particularly minority students in the lower socio-
economic classes, have not received the kind of education they need in order to become viable members of the labor market.

So, in conclusion, despite the Brown v. Board of Education ruling in 1954, many African Americans experience de facto if not de jure educational discrimination. They experience a lack of resources and opportunities that negatively impact their academic achievement. The adverse effects are apparent from the very beginning of these students' educational experiences and are evident across socio-economic classes. My study's inquiry into a complementary assessment is designed to explore one aspect of this problem.

## Social and Economic Ramifications

The deleterious effects of these socio-cultural and school related factors on the academic performance of African American students greatly concern many educators, business leaders, and some politicians as evidenced by media reports. This is because these factors negatively impact not only the democratic underpinnings of our society (Ravitch, 2010), but also our global economic competitiveness (Barton \& Coley, 2008; Friedman, 2005). The consistent differences in academic achievement among the nation's students place the United States at risk of losing its ability to compete satisfactorily on a global scale (Barton \& Coley, 2008). Others, Johnston and Viadero (2000) and Hale (2004), are also concerned about the unequal education and performance of some of our students.

Johnston and Viadero (2000) believe that there will be serious consequences for the United States if minority students are not adequately educated so that they can fully
participate in an economic environment that is rooted in technology. The proper education of both low income and middle class African American students is necessary for the continued wellbeing of the entire nation (Hale, 2004). This is because: "The state of the United States depends on the ingenuity of all its citizens, irrespective of race, to compete in a largely international economy" (Hunter \& Bartee, 2003, p. 157).

Extrapolating from the studies discussed above, it can be inferred that African American students need to be educated well enough so that they can play a positive role in the country's future growth. Unfortunately, in the decades since separate education for the races was declared illegal, and after other civil rights were written into law, sociocultural factors still prevent many African Americans from experiencing equal educational opportunities. For example, limited curricular choices have a negative impact on African American's ability to receive adequate teaching in the STEM (Science, Technology, Engineering, and Mathematics) subjects that are deemed to be critical for the future economic success of the United States. Therefore, for African Americans to become competitive in today's global environment, improving their mathematical competence or literacy will be an important factor of their overall academic achievement.

The critical role that mathematics literacy now plays in people's lives is acknowledged in the work of Robert Moses, a highly respected civil rights leader. According to Moses (2001), access to economic success within the global world we have become depends largely on a person's mathematical skills. Moses (2001) makes a compelling argument that in today's world, mathematical literacy actually trumps
language arts literacy. Therefore, he believes that mathematics education is a new, fundamental civil right.

In a powerful analogy, Moses (2001) compares the lack of mathematical literacy among people of color today to Black people's lack of access to voting rights in Mississippi in 1961. As he sees it, immediate attention must be paid to the development of mathematical literacy for people of color in order to improve their economic opportunities (Moses, 2001). Schoenfeld (2002) concurs. He states, "More than ever before, today's students need to learn to reason and communicate using mathematical ideas" (p. 13).

Currently, access to society's resources depends to a great extent on one's educational qualifications (Stringfield, 2007). These credentials are gained in school through the accountability tests that are used to measure the performance of the nation's students (Bainbridge \& Lasley, 2002; Linn, 2007). In this role, they have tremendous influence on students' future success (Stringfield, 2007). For students whose opportunities to learn have been compromised because of a variety of factors, being held to the same achievement standards as their better educated peers on accountability tests is therefore problematic (Nichols \& Berliner, 2008). It can be inferred that this practice would contribute to the academic underperformance of the students who fall in this category, many of whom are African American. Thus, the potential role of other assessment options to counter the negative consequences of accountability testing in our educational system is a key element of this study's inquiry.

## Statement of the Problem

Accountability testing is very evident in the nation's current educational landscape (Miller, Linn, \& Gronlund, 2009). Evidence of this lies in the fact that most states have mandatory tests in many grade levels due to the requirements of NCLB. The specifications have placed significant negative pressure on the denizens of the nation's schools. While school officials agree that students’ academic knowledge needs to be assessed, the current punitive measures that are attached to state mandated assessments are unnecessary burdens on public school administrators, teachers, and students (Ravitch, 2010). With regards to students, it was the academic underachievement of many African Americans that prompted the dramatic increase in accountability testing (Miller, Linn, \& Gronlund, 2009). However, ironically the data show that many African American students still underperform on these assessments (Hunter \& Bartee, 2003). The difference in achievement is not only evident on state standardized tests but also on national tests such as the NAEP assessments (NAEP, 2009).

Such disparities in achievement among the nation's ethnic groups begin early in students' educational experiences and span all grade levels (Grodsky, Warren, \& Telfts, 2008). Inequalities in performance are also seen among socio-economic groups with students from the lower socio-economic groups generally performing more poorly than more affluent peers (Gardner \& Miranda, 2001; Rothstein, 2004). However, among African Americans, some middle-class students earn lower scores on standardized tests than their White counterparts (Viadero, 2000). Additionally, gender differences contribute to the inequities in performance of African Americans as on most local and
national measures of academic achievement African American boys as a group are underachieving significantly (Garibaldi, 2007).

The chronic academic underachievement of African American students has its roots in systemic discriminatory practices that resulted in Whites having access to better educational opportunities than Blacks (Frankenberg, Lee, \& Orfield, 2003; Kantor \& Brenzel, 1992). More minority students live in poverty, and its negative effects on academic achievement contribute to the discrepancy in academic outcome between the races (Kozol, 2005; Rothstein, 2004). Out of school factors (Berliner, 2009) such as inadequate health care before and after birth, exposure to pollutants such as lead and mercury, familial violence, and unsafe neighborhoods all contribute to the unsatisfactory academic performance of African Americans on standardized tests. In our discourse about the achievement gap between the races more attention needs to be paid to the continuing presence of these impediments in the lives of impoverished and minority students. These elements create tremendous roadblocks to the academic achievement of low SES and minority students, but these factors are often overlooked due a preoccupation with test results (Berliner, 2009).

The scores of standardized assessments chronicle the differences in academic performances among the nation's students, and the divergence in scores are frequently reported in the news as well as the academic media (Bainbridge \& Lasley, 2002). Further, due to the increasing rise of accountability testing in our schools (Solley, 2007) students' test scores have become even more important to teachers as they have potentially negative consequences for their professional careers (Ravitch, 2010).

In their quest to ensure that their students achieve proficiency levels on standardized tests, teachers make a variety of pedagogical decisions including limiting instruction (Orfield \& Wald, 2000) and emphasizing basic skills in an attempt to address perceived gaps in instruction from previous grades (Harris, 2012). Decisions about how to conduct their students' instruction are influenced by their opinions about these students’ academic abilities (Harris, 2012; Montague \& Rinaldi, 2001). Teachers use class assignments and standardized test scores to make judgments about their students’ educational capabilities (English, 2002). As such, test scores are one of the means by which teachers' expectations about their students’ academic competencies are formulated. Using test scores in this manner is important as these expectations affect teacher practice and student learning (Baker, 1999; Kuklinski \& Weinstein, 2001; Rosenthal \& Jacobson, 1968).

Given the influence that standardized tests can exert over teachers' opinions of their students’ academic capabilities, providing teachers with multiple and varied data sources from which to make valid evaluations about their students’ academic competencies is desirable. Gordon (1995) addresses this issue. He suggests that one solution is to use tests and testing procedures that encourage the analyses of cognitive functions. Such new tests or procedures could facilitate the observation of problem solving requiring different cognitive skills and styles (Gordon, 1995). Farkas (2003) also suggests experimental means should be explored to find interventions that address the causes for the academic disparities that exist in the educational experiences of many African Americans.

The clinical interview as a complementary assessment could be one type of data source. It provides another "window" into these students’ academic performance, broadening the picture of how much they know and are able to do. This in turn may help practitioners to confirm or reevaluate the expectations they held about the academic performance of the African American students in their classes.

It should be noted that the clinical interview would not replace the standardized test. It would be used in conjunction with standardized tests. By revealing what students know about mathematical concepts, the primary function of the interview would be to provide teachers with an alternative source of information about their African American students. A crucial difference between the clinical interview and the standardized test is that the clinical interview permits deeper insight into student's mental processing. This data can provide teachers with a better view of students' strengths and weaknesses that some can use to implement changes in classroom pedagogical practices. In this way, teachers do not only receive a score indicating a certain level of proficiency, but also receive other pertinent information about the students' performances. In addition to the numerical data that is accessible through test scores, teachers would also receive detailed information about their students.

As Barton (1999) notes, while standardized testing has provided much numerical data by documenting test scores, some of this information has not always been helpful to teachers in their efforts to improve student performance (Barton, 1999). If teachers are given pertinent and helpful information about what their students are learning, they can help children achieve important mathematical objectives (NCTM, 2000). By offering a
"window" to the problem solving processes of students, the clinical interview could help teachers learn more about their African American students' academic ability. This could be particularly relevant for less careful or inconsistent achievers. This information could in turn lead to more accurate perceptions about these students that could in turn positively impact their academic futures. These possibilities were examined through the study's research questions.

## Research Questions

This study explored the following questions: 1) Will the clinical interview show that African American students actually know more about mathematics than they are able to demonstrate on a standardized test? 2) What might teachers learn about their students’ mathematical ability, and how might this knowledge affect their perceptions of these students?

## Overview of the Methodology

This was a qualitative study that examined the research questions by utilizing clinical interview methodology. The participants were seven African American male elementary school students and their White (non-Hispanic) teachers in a suburban school district in New Jersey. Data were collected from the School Report Card, teacher interview and survey sheets, teacher and investigator notes, teacher interview transcripts, clinical interview protocols and transcripts, and test scores from the New Jersey Proficiency Assessment of State Standards (NJPASS). Data were coded to identify common themes as they related to the exploration of the relationship among the teacher's
evaluations of their students' mathematical abilities, the students' performances during the clinical interviews, and their standardized test scores.

## Overview of Findings

Most of the students performed better on the standardized test than on the clinical interview. Only one of the teachers indicated that she learned anything new about the mathematical abilities of her student as a result of the clinical interview. Another teacher planned to make pedagogical changes in her classroom as a result of the information gathered from the clinical interview.

## Summary

This first chapter introduced the dissertation and some of the topics discussed here will be expanded upon in greater detail in the subsequent chapters. It began by describing the development of my personal interest in the underperformance of my African American students in comparison with their White peers. This interest led to a desire to explore what could be done to address the issue of their differences in academic performances.

A discussion of reform initiatives and the lack of definitive resolution of the performance gap followed. A long history of educational and other inequities for African Americans in the United States has contributed to the lack of a clear solution for this problem. The chapter progressed by exploring the effects of these historical inequities on African Americans. It then moved from a historical perspective to exploring current problems surrounding the achievement gap and its national and global implications. Part of these implications included the importance of mathematics literacy for African

Americans in today's global environment. This discussion led to a statement of the problem investigated in the study. The chapter concluded with a description of the research questions, an overview of the methodology, and a summary of the findings.

Chapter two will elaborate on all the themes that have been introduced and discussed in chapter one. The literature review provides the theoretical framework for the study as it addresses the problem and the suggested solution.

Descriptions of the participants, the setting, data sources, method and rationale for the study are found in chapter three.

The findings of the study are contained in chapter four. Examples from both the clinical and teacher interviews are provided for clarification.

Chapter five situates the findings in terms of the larger educational and social issues that prompted the research in the study. This chapter also includes the responses of the teacher participants to the findings of the study. Suggestions for future research about the education of African Americans in the context of current American education are included. The limitations of the study are also discussed.

## CHAPTER TWO

## Literature Review <br> Introduction

Chapter two reviews the pertinent bodies of literature that provide the framework for this study. The review begins with an examination of minority achievement on standardized tests. It continues by analyzing the advantages and disadvantages of accountability testing in the nation's current educational environment. The multifaceted consequences of teacher expectations on student performance are addressed next. That discussion is followed by an exploration of the use of clinical interviews in educational research. The study proposes using the clinical interview as a complementary assessment to the standardized test for African American students. Consequently, the focus of the literature review moves to explore the use of clinical interviews in educational research. The ideas of Freud, Piaget, and Vygotsky represent the theoretical validation for the clinical interview assessment, and their singular and collective contributions are discussed.

Using the clinical interview in educational research is not a prevalent practice, though its use has become increasingly popular. Therefore, empirical support for choosing this type of assessment in this study is provided. Then, the actual process of the clinical interview as used in an educational setting is described. This description is one way of clarifying the differences between the clinical interview and dynamic assessment. In the last section of the literature review, I first discuss the similarities between these two assessments. Then, I present their salient differences. Finally I clearly identify the reasons for choosing the clinical interview as the research tool in this study.

The study explores the use of a complementary type of assessment, the clinical interview, as a means of allowing African American students to show that they know more about mathematical concepts than they are usually able to reveal on standardized tests. The literature review begins by presenting the problem of African American student underperformance on standardized assessments and ends by examining one potential solution to this pervasive and persistent problem of our educational system.

African American Achievement and Standardized Tests

## Test Scores

The achievement of African American students on standardized tests has been documented over the years by test scores. The intransigently lower scores for urban and suburban African American students are particularly worrisome as test results are used in ways that significantly affect students' futures as well as their everyday lives in schools. For example, these scores are used to determine crucial educational issues such as promotions between school grades and graduation from high school (Orfield \& Wald, 2000). Stringfield (2007) adds that test scores influence the resources that students receive in school. They are instrumental in determining which students are admitted to academically prestigious and challenging programs, schools, colleges, and graduate schools. Test scores affect students' lives beyond school as they influence career opportunities (Stringfield, 2007). The scores consistently show that most African American students underperform academically in comparison with their White peers regardless of their socio-economic status (Ferguson et al, 2002; Gardner \& TalbertJohnson, 2001; NCES, 2009; Viadero, 2000).

As these examples indicate, test scores are used to make decisions that have significant consequences for the current and future success of students in our educational system. The persistent discrepancy in academic performance between many African Americans and their White peers has therefore been of considerable concern for many years. This has led to a variety of federally mandated reform initiatives such as the Title I program. In response to the inequities in educational opportunities and unsatisfactory student performance, the federal government instituted the Title I program providing compensatory education as part of the Elementary and Secondary Education Act (ESEA) of 1965 (Miller, Linn, \& Gronlund, 2009). The ESEA was not successful in improving academic achievement as was revealed in the 1983 report entitled, A Nation at Risk: The Imperative for Educational Reform. The report suggested that higher standards were needed in many areas in order to improve the academic performance of the nation's students (A Nation at Risk, 1983). It initiated the use of standards-based assessments culminating with the NCLB Act of 2001(Linn, 2006).

Issues with Using Testing to Assess Achievement
While the purpose of reform measures was and is to improve African American student performance on standardized tests, Linn and Hambleton (1991) note that these assessments are often used not only to determine the effectiveness of the reform initiatives, but also as the actual instruments of the required improvements. Increased testing was viewed as the means by which reform measures would be instituted (Linn, 1995). Thus, using standardized tests as tools of reform has become prevalent in our educational environment.

As Linn (2006) explains, these assessments are expected to fulfill important societal roles such as engineering broad educational changes. Linn writes the following:

These tests have been expected to: help clarify expectations for teaching and learning; monitor educational progress of schools and students; monitor the progress of demographic subgroups of students and the gaps in achievement of those subgroups; encourage the closing of the gaps in the performance among racial/ethnic subgroups and between economically disadvantaged students and their more affluent peers; motivate greater effort on the part of students, teachers, and school administrators; contribute to the evaluation of educational programs and schools; identify schools and programs that need to be improved; and provide a basis for the distribution of rewards and sanctions to schools and students. (pp. 2-3)

Hunter and Bartee (2003) also discussed the expanded use of standardized testing. They note that test scores are being increasingly used to give the stamp of approval to the nation's public education system. They contend, "NCLB is based on the market-based contextual framework of choice and competition in public education, which is substantiated by standardized testing. This approach is also based on the presumption of equitable access and equal opportunity in American schools and society" (p. 153). Hunter and Bartee (2003) also believe that disparities between the academic performance of African American and White students are viewed negatively. The critics who hold these unflattering opinions ignore the fact that mandatory standardized testing does not mean that all students have equal access to a good education (Hunter \& Bartee, (2003). In addition, Hunter and Bartee (2003) declare that NCLB is flawed because "...it assumes a blame the victim approach in its attempt to remedy educational inequities" (p. 158). English (2002) agrees that there is widespread belief in the idea that standardized testing is inherently non-discriminatory. However, in debunking this theory, English (2002) refers to the belief in the inefficacy of public education. He states that the idea that public education is ineffective, especially in many inner cities, can be traced
back to the inadequate test scores that are associated with students in these areas (English, 2002). Students who attend schools in urban areas are predominantly from low income and minority backgrounds (Kantor \& Brenzel, 1992). The social and economic circumstances in which these students reside often do not provide them with the kind of knowledge that is preferred by schools (Rothstein, 2004). In essence, they lack the requisite "cultural capital" that is necessary for academic success.

The term cultural capital was coined by Pierre Bourdieu, and in his explanation of how the phrase originated, Bourdieu (1986) states that it resulted from his desire to understand the reasons for the disparate levels of academic achievement within the different social classes. He wanted to distinguish between natural intellectual acumen and the cultural capital that was bequeathed to students by their families. Bourdieu (1986) posited that, "...the scholastic yield from educational action depends on the cultural capital previously invested by the family. Moreover, the economic and social yield of the educational qualification depends on the social capital, again inherited, which can be used to back it up" (p. 48). His ideas refuting the notion that academic success was entirely due to innate intellectual ability that was not affected by extraneous influences continue to be pertinent in today's educational discourse.

For example, the lack of certain kinds of cultural knowledge is sometimes considered equivalent to being academically incompetent (English, 2002). While a lack of cultural capital is not synonymous with a paucity of intelligence, it does negatively impact students’ academic performance (Hale, 2001). English (2002) asserts that if the impact of cultural capital on students' performances on standardized tests continues to be discounted then differences in academic performance among students will persist.

This idea that the impact of cultural capital on the academic achievement of low socio-economic and minority students has been insufficiently addressed in our educational discourse has caused critics of standardized testing to complain that many tests are culturally unfair (English, 2002). Rothstein (2004) states that increased financial security and educational achievement are not always immediately reflected in childrearing practices such as taking children to museums and involving them in other intellectual extra-curricular activities. Thus, people who have achieved middle-class status for the first time are not as imbued with the benefits of education and are not as facile in providing their children with additional exposure to activities that enhance their children's literacy skills (Rothstein, 2004). Berliner (2009) is also aware of the importance of what he calls "extended learning opportunities" (p. 37) and their effects on achievement. In the section of the article in which he discusses the Out-of-School Factors (OSFs) that positively impacted children, especially those who are from lower socio-economic classes, Berliner (2009) writes:

In this final, short section, some programs that can improve the education of youth, particularly poor youth are discussed. These programs offer extended learning opportunities, and so they are all educational programs; however, they typically operate separately from traditional school programs, and thus they are classified as an OSF influencing achievement. This section is also a reminder that education does not take place only within schools. The many opportunities for learning outside of school, where some students learn more (and more easily), are not equally available across income groups. In addition, poorer students or their families who are not or cannot be motivated to take part in available out-of-school learning opportunities will not learn as well or as much as those who do. (p. 37)

As these examples illustrate, standardized tests do not always adequately take into account the cultural and socio-economic differences among students. Further, as Ravitch
(2010) explains, there is another troubling factor with the increased use of standardized testing in our schools. Due to the importance of test scores, there has been an increase in the amount of test preparation occurring in many school districts. This is particularly prevalent in many inner city schools where students' daily instruction is constituted primarily of practice that is designed to improve performance on state mandated assessments (Ravitch, 2010). Consequently, test preparation replaces teaching in poor districts more frequently than in more prosperous ones (Orfield \& Wald, 2000).

Ironically, this emphasis on test preparation has the opposite effect of the desired outcome. The data show that the students in the more impoverished districts still underperform on standardized tests (Gardner \& Miranda, 2001). As a result, the gaps in achievement between African American and White students continue to be widened (Dillon, 2009; William, 2010). Thus, narrowing the curriculum may help students pass one specific standardized test, but it may not help them to pass a different assessment or even apply that limited knowledge to their everyday lives (Koretz, 2008).

A narrowed curriculum is particularly devastating for students who begin school with limited vocabulary exposure. Barton and Coley (2008) describe a study in which it was found that, "By age 4, the average child in a professional family in the study heard about 20 million more words than the average child in a working-class family, and about 35 million more words than the average child in a welfare family" (p. 3). It was estimated that it would take several hours per week (approximately equivalent to a 40 hour work week) of extra-curricular vocabulary rich experiences in order to ameliorate this huge language deficit (Barton \& Coley, 2008). However, since this type of remediation is not usually possible, language
deficiencies are another factor that can have adverse affects on minority students’ academic achievement. This finding is supported by the research that suggests that one reason for the disparity in standardized test scores is the type of language used in the phrasing of test questions. As Boaler (2003) notes, the language used in questions can negatively impact students’ ability to succeed on standardized tests.

More specifically, Matteson (2006) states that mathematics has its own unique language. Since it has so many specialized terms, it is akin to learning a foreign language. So, as teachers instruct students about the mathematical terms that are used in textbooks and on standardized tests, they need to ensure that the children are also able to understand the contexts in which these terms may appear. Consequently, mathematics teachers need to realize the importance of reading comprehension on standardized tests. This is particularly important with multiple-choice questions. Students need to understand not only the mathematical terms that are used, but also the rest of the language in the questions (Matteson, 2006). Mastery of terminology and context will facilitate reading comprehension that Matteson (2006) believes is a requisite skill for improving students’ ability to engage in mathematical communication. As Draper and Siebert (2004) state, "In mathematics education, teachers and educators have long recognized the importance of reading, writing, symbolizing, and communicating" (p. 928). This viewpoint is echoed by the recommendations of the National Council of Teachers of Mathematics (NCTM).

The communication section of the NCTM guidelines of 2000 states that in order for students to do mathematics successfully, instructional programs from prekindergarten to grade 12 should enable all students to, " organize and consolidate their mathematical thinking
through communication; communicate their mathematical thinking coherently and clearly to peers, teachers, and others; analyze and evaluate the mathematical thinking and strategies of others; and use the language of mathematics to express mathematical ideas precisely" (NCTM, 2000). Based on this data, it can therefore be concluded that language is not only relevant on reading and language arts tests, but it is also important in other subjects such as mathematics that are assessed on standardized assessments. It is one more element that inhibits academic success on standardized tests.

Teaching to the test, cultural test bias, and differences between the language used at home and the language used in school are not the only factors that can negatively affect African American students' achievement on standardized tests. Other studies such as the one conducted by Hill, Guin, and Celio (2003) indicate that this achievement is further inhibited by the inequitable distribution of the financial resources among the nation's students and their parents. Many African American low-income students in poorer districts nationwide do not have equitable access to the educational resources they need in order to succeed academically (Darling-Hammond, 2007; Hill, Guin, \& Celio, 2003: Kozol, 2005). For example, inequitable funding for predominantly minority students in high poverty areas denies them access to preschool programs such as Head Start, exposes them to narrowed curricula, subjects them to more inexperienced teachers, provides them with little or no extra-curricular activities, and houses them in physical plants that are often in drastic need of repair (Kozol, 2005).

Berliner (2009) describes situations that, although they do not actually occur inside school buildings, seriously affect the academic achievement of low SES and minority children. Scant emphasis is given to the many non-school impediments that
affect the learning potential of economically disadvantaged students as there is an almost overwhelming belief that schools are the means of bettering the academic achievement of students who live in poverty (Berliner, 2009).

It is within this context that this brief offers a reminder that inputs, including many of the equity issues that have dropped largely out of sight, have never stopped affecting the achievement of our most impoverished youth. In fact, it is the position taken here that we can never reduce the achievement gap between poor and non-poor children, between African American and white children, or Hispanic and Anglo children, unless OSFs that positively or negatively affect achievement are more equitably distributed. In the U. S. today, too many OSFs are strongly correlated with class, race, and ethnicity, and too many children are in schools segregated by those very same characteristics (Berliner, 2009, p. 7).

Despite the problems that many poor and minority students often bring to school, the institutions that teach these pupils are expected to raise their test scores. The mandate is to improve the academic performances of these students, but this quest ignores the many outside factors that impede their academic progress. In recognition of this lapse, Berliner (2009) identifies six out-of-school factors that are especially relevant for impoverished students. They are: poor pre-natal conditions, medical care, food insecurity, pollutants, family relations and stress, and neighborhood norms. Despite the emphasis on results in the current discussion on academic achievement, Berliner (2009) believes that the obstacles such as those listed above should be included in the conversation given their powerful impact on educational success.

The salient points that Berliner (2009) makes about these out-of-school factors and their impact on academic achievement can be summarized as follows. Inadequate access to health care among low SES students affects their learning and behavior in school, and they are more likely to have high rates of absenteeism due to illnesses such as
asthma. Substandard nutrition alters their cognitive functioning and their ability to participate in higher levels of learning while pollutants impair their behavior and intellectual development. Children who are exposed to familial violence experience "...higher rates of aggressive behavior, depression, anxiety, decreased social competence, and diminished academic performance" (p. 25). Further, their peers in school suffer not only from increased incidences of inappropriate behavior but also from a decline in their test scores. Impoverished neighborhoods have too many troubled youth who lack the necessary role models that encourage striving for educational success. These youth engage in antisocial activities that undermine the family values of parents in these communities (Berliner, 2009). These factors as well as those described earlier highlight the significant barriers to academic success that are faced by many African American students who live in poor communities.

However, even when African American students are educated in adequately funded districts in suburban neighborhoods, Hale (2001) explains that they sometimes do not have the same level of economic income as their White neighbors. Rothstein (2004) agrees and notes that many black middle-class families do not have the same level of financial assets as their white counterparts. So, in order to live in middle-class districts, African American families often have to work much harder to maintain their homes (Hale, 2001). Their middleclass status is more tenuous as they often do not have family members who have attained financial security (Rothstein, 2004). Hale (2001) adds that they often do not have the cultural capital that enables them to deal effectively with their children's teachers. As a result, some of the African American parents in these neighborhoods sometimes do not have the time,
energy, or knowledge to negotiate the culture of school on behalf of their children (Hale, 2001).

The concerns described above are evidence that standardized tests are increasingly used in ways that do not sufficiently adjust for the differences in educational opportunities (Kozol, 2005) and cultural capital (English, 2002) that are experienced by some African American students. This is exemplified by NCLB (2001) that mandates that all students attain scores of proficiency by 2014 and which specifies punitive measures for those failing to meet those standards (Darling-Hammond, 2007; Linn, 2007; Ravitch, 2010). The nation’s students reflect a wide spectrum of academic abilities and performance ranges as evidenced by their test scores (NAEP, 2009; NCES, 2009). They experience disparate access to household resources, educational opportunities and qualified teachers (Bainbridge \& Lasley, II, 2002; Barton \& Coley, 2008; Farkas, 2003; Hale, 2004; Hill, Guin \& Celio, 2003). Therefore, expecting all the nation's students to be in compliance with uniform standards on state or national tests is not feasible (Ravitch, 2010).

These examples highlight some of the reasons why on average, as Flowers (2007) notes, African American students earn lower scores on standardized tests not only in comparison with White students but also with other racial and ethnic groups. There are, however, some exceptions to these findings although they are not as prevalent in our educational discourse on this topic. Stinson (2006) states that with reference to African American students, "The discourse of achievement has been the least researched and theorized discourse." (p. 495). However, there are studies that highlight African American students who succeed academically despite marginalizing experiences (Stinson, 2006). In his
study Stinson (2006) examined the sociocultural factors that affected the schooling experiences of four successful African American male students. Later in his article, Stinson (2006) referred to a few similar studies. Hebert and Reis (1999) examined an ethnically mixed group of students who attended an urban school under adverse circumstances but who had managed to do well academically. O'Connor (1997) conducted a study that featured lowincome African Americans who had also succeeded academically against tremendous odds. The impact of their peers on mathematically successful minority students was the focus of yet another study that Walker (2006) administered. The unifying factor was that the studies all highlighted African American and Hispanic students who succeeded academically despite enormous challenges.

These studies showed that some African American students achieve success in schools and on standardized assessments. Noguera (2003) in his discussion about the educational issues of African American male students corroborated this evidence as he indicated that there were schools where academic success for African American students was the norm and not the exception, and where African American males did well academically. Finally, in another example of African American students being successful on standardized tests, Slavin and Madden (2006) focused on the impact of the "Success for All" program on the achievement of high poverty, predominantly African American children. The evidence from their data suggested that the program could help in reducing the achievement gap between these students and their peers.

These studies indicate that despite the preponderance of evidence to the contrary, there are African American students who are able to succeed academically on standardized tests.

As Stinson (2006) suggests, more research needs to be done in order to ascertain the factors that contribute to these students' academic success. He believes that an emphasis on the discourse of achievement could facilitate new ideas and pedagogical practices that could be useful for African American students particularly in the area of mathematics education (Stinson, 2006).

## Conclusion

This discourse of achievement could be a new lens through which to view the academic accomplishments of African American students. The information obtained from complementary assessments such as the clinical interview could help refocus the "camera" that captures images of underperformance. Instead of capturing and recording poor academic achievement, the camera can highlight images of successful students that can then be critically examined for nuggets of data that can be implemented with less able pupils. Such a change in perspective would be paradigmatic. Historically, the discussion of African American achievement has usually been framed around the viewpoint of reform. However, reform initiatives that were implemented to counter the adverse effects of under-achievement on standardized testing usually took the form of increased assessment and this has not ameliorated the problem. This is because these initiatives ignored educational inequities including, but not limited to, highly damaging out of school barriers that adversely affect the academic achievement of African American students.

Other inequalities include cultural bias in testing, differences in language usage, narrowed curricula and teaching to the test, language deficits, inequitable funding, and differential financial assets. Despite these inequalities, and the wide array of educational
access and abilities that exist in the nation, under NCLB, all students are nonetheless expected to attain similar levels of proficiency on standardized tests. Despite these unfavorable odds, some African American students across income brackets do achieve academic success.

This is promising and provides hope for the future academic achievement of these students. However, these students are in the minority and they along with their less successful peers face many challenges that are the result of using standardized tests to improve their academic accomplishments. For the student participants in my study, the NJPASS represented the first exposure to accountability testing. Due to the increasing popularity of high stakes testing in our educational system, this encounter was merely a prelude of the future that will be repeated annually as they progress through the grades. Thus, along with their peers from the other ethnic groups, these participants as well as other African American students, regardless of academic ability, face the increased pressure of high stakes accountability testing.

## High Stakes Accountability Testing

## The Shift from Standardized Testing to High Stakes Accountability Testing

A look at the nation's educational history reveals that standardized tests were not always used as measurements of accountability. As Baker (2008) notes, achievement testing was originally designed to appraise students’ academic progress. It was meant to evaluate whether the academic programs of the nation's public schools adequately fulfilled students' learning needs. The information provided by achievement testing was to be used to create strategies that met the educational needs of the nation's students including those in need of special assistance (Baker, 2008). Testing was also seen as a
means of controlling and managing the nation's burgeoning and increasingly ethnically diverse school population (Linn, 2001). However, with the increasing use of testing to redress poor academic performance issues among the nation's students (Linn, 1995), the emphasis on testing shifted from evaluation for the purpose of providing students with assistance to evaluation for accountability purposes.

As Miller, Linn, and Gronlund (2009) explain, federal officials requested accountability through the use of standardized testing for programs funded under Title I of the Elementary and Secondary Education Act (ESEA) of 1965. Title I students were required to be evaluated twice a year. The Title I scores from these evaluations showed that student achievement had not improved significantly for students in the program (Miller, Linn, \& Gronlund, 2009). Schools were seen as being responsible for this failure, and testing was seen as a way to make educators more responsible for student learning and also as a means of improving classroom instruction (Linn, 2006).

From the mid 1960s through the 1980s, the academic achievement of the nations’ students did not improve sufficiently despite increased testing. This decline in academic progress was documented in reports such as A Nation at Risk: the Imperative for Educational Reform (1983). Widely publicized and viewed as a clarion call for educational reform, this report gained the attention of the general public (Ravitch, 2010). Equally worried about the state of public education, policy makers and business leaders became increasingly involved in educational reform. Wanting to make educators and students more accountable, they proposed using quantifiable measures to evaluate educational progress (Ravitch, 2010).

Standardized tests were seen as the vehicle through which the academic progress of the nations’ students could be assessed (Barton, 1999). These tests necessitated the implementation of accountability mechanisms to track the performance of schools, administrators, teachers, and students (Miller, Linn, \& Gronlund, 2009). In an attempt to make these accountability mechanisms more cohesive and forceful, the No Child Left Behind Act (NCLB) was signed into law in 2001. In response to its requirements, all states instituted mandated proficiency tests designed to measure the academic progress of their students. In this way, NCLB created an added dimension to the testing and accountability practices already in place in the nation’s public schools (Ravitch, 2010). In addition to federally mandated state tests, American students also take assessments such as the National Assessment of Educational Progress (NAEP) in order to get national statistical data (Miller, Linn, \& Gronlund, 2009). This plethora of tests and testing has disadvantages and advantages.

## Effects of High Stakes Accountability Testing

Standardized tests, particularly high stakes accountability tests, are an area of substantive controversy in our nation's educational dialogue. There are proponents who cite the advantages of these assessments as well as opponents who list their disadvantages. Both groups provide data that is intended to substantiate the antithetical viewpoints. In this section, I present the differing perspectives of some of the more common arguments surrounding standardized tests.

In one difference of opinion, supporters of standardized tests claim that more than ninety percent of studies show that student assessment, including the prevalent high-
stakes testing, is beneficial to students' academic accomplishment (Phelps, 2001). Phelps (2001) based his finding on an analysis of a hundred years of research that evaluated the effects of testing on students. Countering that claim, opponents of standardized testing assert that student achievement has declined despite the rise in accountability testing. To illustrate their point, they cite the decreasing scores American students earn on international tests such as the Programme for International Student Assessment (PISA). Further they state that educators and policy makers have not yet discerned how to use assessments to ameliorate student achievement in particular and the educational system in general (Incentives and Test-Based Accountability in Education, 2011).

The positive or negative effects of standardized tests on student achievement are not the only area in which there is disagreement about the desirability of the assessments. The issue of objectivity is another source of dissent. Advocates of accountability testing emphasize that the tests are objective and reliable especially in contrast to the subjective evaluations done by educators (Phelps, 2002). Adversaries argue that standardized tests are not very reliable measures of student attainment. They believe that increases in students' test scores are ephemeral, and that they do not represent enduring improvements in learning (Olson, 2001). Further, scores are affected by how the tests are designed and administered. For example, Koretz (2002) points out that changing the relative weight of two of the topics on the NAEP amended the performance gap between African American students and their White peers.

In defending their assertions of neutrality, those who are in favor of standardized tests believe that multiple-choice tests are particularly impartial as they are scored by
machines and not by humans (Phelps, 2002). Multiple-choice tests are not only considered to be objective. They are also believed to provide the type of precise information that is necessary for improving the nation's schools (Center for Teaching Excellence, 2011). Proponents of standardized tests state that the current examples of multiple-choice questions require more critical thinking and use more complicated calculations than previously (Mitchell, 2006). Critics decry these claims and assert that the multiple-choice format in assessment is not desirable due to its emphasis on correct or incorrect answers. The thinking is that this basic duality does not encourage the complexity of thought that is required in daily interactions. The challengers also note that women do not adjust as well as males to this type of testing. Men fare better because getting multiple-choice questions correct is similar to scoring points in games (Sacks, 2001). Thus, some critics charge that tests are biased with reference to gender. They further assert that the lack of neutrality occurs in other aspects of standardized testing.

Challengers declare that standardized tests are designed to benefit White students who have no learning disabilities. As evidence, they note that special education students can sometimes only use a few of the accommodations that are listed on their Individual Education Plans (IEPs) (McKnight, 2011). Additionally, second language learners are required to take these tests although they have inadequately mastered English. Defenders of the benefits of standardized testing further bolster their claims of impartiality by noting that since the tests require the same content and testing parameters for all students, this makes the tests less biased towards minorities and students with disabilities (Rhee, 2011). Regardless of the opposing viewpoints, since state accountability test scores are
disaggregated, the performance of students with IEPs will have an enormous impact on school scores (Miller, Linn, \& Gronlund, 2009). The scores of all students will now be reported.

As Ravitch (2010) explains, NCLB mandated that the academic achievement of poorly performing demographic subgroups would now be highlighted and would not be subsumed under other categories. Previously, their underperformance could be hidden within school district and state averages (Ravitch, 2010). Thus, school districts were intended to do all they could to bring their struggling students up to par educationally. NCLB focused attention on the disparities in academic performance among the races and enabled students with academic needs to receive remediation (ETS Policy Notes, 2009). With this focus on helping students receive the academic assistance they need, test scores not only provide information about student needs but also indicate how quickly their problems are mediated (Stringfield, 2007). In this way high stakes state accountability tests have been instrumental in improving student learning (Wang et al, 2006).

Just as published test scores can be used to exhort educators into providing academic help for needy students, supporters of standardized testing believe that it is beneficial for teachers to use the content of the assessments to prepare their students. These proponents think that it is constructive for schoolteachers to focus on the material that is presented on the tests because that instruction enhances student performance on these examinations. Further, advocates contend that if this content is taught well, the students will master not only those topics but also other content areas as well (U. S. Department of Education, 2004). The opposing perspective is that when children are
taught to the test, teachers abandon proven pedagogical habits and employ strategies that emphasize rote learning to the detriment of more cognitively complex activities (Jacobs, 2007; Valli \& Croninger, 2011).

In agreement with the last assertion, Causey-Bush (2005) notes that strict adherence to state standards by districts can cause some teachers to stop using culturally salient practices that actually benefited students in their classrooms. She further states that some educators erroneously believe that teaching students strategies aimed at improving their performance on standardized tests is more necessary than fostering student creativity and interests (Causey-Bush, 2005). Many fear that the country is losing its creative edge because there has been insufficient attention paid to innovative characteristics in students due to the rise in standardized testing (Zagursky, 2011). They point out that many of the qualities that are important in developing well-rounded students are not measured on standardized tests, and are consequently being ignored in our nation’s schools (Strauss, 2011). Acknowledging the importance of personal qualities that are not measurable through standardized assessments, Rothstein (2005) indicates that standardized tests are unable to evaluate "creativity, insight, reasoning, and the application of knowledge to unrehearsed situations" (p. 34).

The controversy surrounding teaching to the test and the purported lessening of creativity in the nation's students is related to another area of contention - the narrowing of the curriculum due to the excessive pressures of accountability testing. Proponents of standardized testing state that rather than narrowing the curriculum, the tests provide the opportunity for students to master the necessary basic skills. In addition, these advocates
state that the tests help teachers make positive changes to the curriculum while increasing students’ academic accomplishment (Yeh, 2005). The opposing viewpoint is that teachers might neglect subject areas that are not currently tested and focus mainly on those that are assessed on standardized tests (Orfield \& Wald, 2000; Ravitch, 2010; Shepard, 2000). Additionally, if students are only taught material that applies to a specific standardized assessment, they will not be able to use that knowledge in other academic or social contexts (Koretz, 2008). Further, opponents assert that the evidence shows that over forty percent of school districts have decreased the periods of time (by about two hours weekly) that teachers taught science, social studies, and the arts. This was done to increase the amount of time spent on instruction in reading and mathematics (Ravitch, 2010).

Students are not only affected by the amount of instructional time they receive in certain subjects that are taught in schools. Some critics of standardized tests claim that the assessments are a source of tremendous stress in younger pupils. This pressure can cause these students to vomit or cry (Ohanian, 2002). In contradiction to these claims, the U. S. Department of Education asserted that standardized tests are a normal and accepted part of students' education to which children have grown quite accustomed. Thus, they seldom emit due to undue stress caused by these tests (Mitchell, 2006; Mulvenon, Connors, \& Lenares, 2011). In addition to being concerned about increased stress levels in students, detractors also state that accountability testing has lessened the opportunities children have to engage in quintessentially childlike activities during the school day. To support this argument, they note that state mandated testing begins in
third grade and that there has been less time allotted for play in kindergartens (Kozol, 2005; Miller \& Almon, 2009).

Orfield and Wald (2000) posit that another challenge for students as a result of high stakes accountability testing is that the scores from these assessments are used to determine which students are promoted to the next grades, a practice they call unsound. Miller, Linn and Gronlund (2009) agree that using assessments to determine grade retention is not a viable option, and the possible adverse effects of this practice on students should be weighed against its usefulness for educators. Deciding on whether or not to retain a student in a grade should not be done based only on the scores of an accountability test (Miller, Linn, \& Gronlund, 2009). In opposition, supporters of standardized tests say the results are a deterrent to the practice of advancing children through the grades regardless of whether or not they have achieved the necessary academic competency for the grade level (U. S. Department of Education, 1999). Retaining students was actually shown to have improved students' test scores in reading and math (Greene \& Winters, 2004).

In addition to being used to decide whether children will be promoted in schools, these test scores also determine who will graduate from high school (Orfield \& Wald, 2000). This has important ramifications for African American students given their generally poorer performance on accountability tests because test scores and course grades are still the most effective indicators of students who will drop out of school (Lan \& Lanthier, 2003). Thus, Orfield and Wald (2000) note that test scores can contribute to higher dropout rates especially among minority students. However, even if test
preparation were to improve the test results of these students that would not necessarily guarantee their ability to think critically as adults (Xueqin, 2010; Zhao, 2010).

While excessive accountability testing prepares students to become good test takers, critics state that it does not really prepare them to be productive adults. Detractors note that China, whose students receive extensive standardized test preparation, is trying to limit the extent of this practice. The Chinese have recognized that rote test preparation has not produced high caliber critical thinkers among their students (Xueqin, 2010; Zhao, 2010). In contrast, proponents of standardized accountability testing state that the increase in standardized testing along with the accompanying higher standards are doing a better job at preparing students since NCLB. This opinion was based on a survey of college professors in 1998. The results showed that $66 \%$ of them believed that schools were not adequately preparing students. In 2002, this percentage dropped to $47 \%$ (Gershon Jr., 2002; Public Agenda, 2002).

Much of the debate about standardized tests focuses on their effects on students, but these assessments also affect states, school districts, and teachers. Supporters of standardized testing believe that they are an inexpensive way to gather helpful data in a relatively small amount of time (Phelps, 2002; Walberg, 2011). The countervailing perspective states that standardized tests are increasingly costly especially since the passage of NCLB and are a strain on the fiscal solvency of many states (Martinez, 2011). Additionally, some opponents think that the high stakes that are attached to standardized tests incite teachers and schools to engage in cheating (Strauss, 2010; Toppo, 2011). Proponents counter that teachers and administrators rarely participate in deceitful
practices. They feel that the minimal amounts of cheating that occur are insufficient reason to abandon testing (Cizek, 2005; Rosenthal, 2011). A final area of dissension deals with the fact that each state develops its own accountability tests. Supporters argue that student progress can still be compared through NAEP scores. The comparisons provide proof that the various state assessments are equivalent in difficulty and that any improvements in test scores are valid (Bourque, 2005). Opponents counter that because each state creates its own tests based on NCLB requirements, there is no real ability for true discrimination among them because the tests are different. They cite, for example, that some tests are multiple-choice while others are a combination of open ended and multiple-choice questions (Rizga, 2011). Nonetheless, despite the differences among the tests, if their validity and reliability are verified, they can provide important information for policy makers and educators.

Ravitch (2010) delineates the types of information that can be provided by valid and reliable standardized tests. She states:

The results can show students what they have learned, what they have not learned, and where they need to improve. They can tell parents how their children are doing compared to others of their age and grade. They can inform teachers whether their students understoodwhat they were taught. They can enable teachers and school administrators to determine which students need additional help or different methods of instruction. They can identify students who need help in learning English or special education services. They can inform educational leaders and policy-makers about the progress of the education system as a whole. They can show which programs are making a difference and which are not, which should be expanded and which should be terminated. They can help to direct additional support, training, and resources to teachers and schools that they need them. (pp. 150-151)

## Conclusion

Opinions about standardized testing are evenly divided between its advocates and its opponents. Experts suggest, however, that they be used along with other relevant measures when assessing academic progress (Ravitch, 2010). Basing important educational decisions on standardized tests alone is not professionally recommended as no test is capable of achieving total perfection (Baker, 2008). Hunter and Bartee (2003) believe that standardized test scores should be only one factor in determining academic learning, and so "standardized testing should not be considered as the ultimate measuring stick of educational attainment" (p. 153). These caveats are sometimes overlooked by accountability testing advocates. However, adherence to these suggestions would ultimately enhance the validity of the standardized tests and could reduce their current overwhelming importance in the lives of our nation's students.

Throughout our country's history, various reform efforts have had accountability testing as a key element in improving the academic performance of African American students (Wang et al, 2006). Yet, accountability testing has not succeeded in eradicating this intransigent problem in our educational system (William, 2010). The academic performance gap persists, and the scores of these tests can influence the expectations teachers have about the academic abilities of their African American students. With the emphasis that accountability assessment places on improving students' performances, test scores impact not only teachers’ pedagogical decisions but also their opinions of students’ academic aptitudes. In a cyclical manner, teacher expectations of their students' intellectual traits also affect the latter's academic success.

## Teacher Expectations and Academic Achievement

## Teacher Expectation Research Findings

The pioneering work of Rosenthal and Jacobson in their 1968 study was seminal to teacher expectation research (Rubie-Davis, 2010). The main thrust of the study was that one person's belief about another person's intellectual capacity could actually influence the latter’s academic performance (Rosenthal \& Jacobson, 1968). As RubieDavis (2010) explains, the Rosenthal and Jacobson study seemed to demonstrate that when teachers expected good academic performance from their students they were not disappointed. When they expected less of their students, the latter's academic performance was not as good, exemplifying what has been described as the self-fulfilling prophecy effect. Additionally the study found that teachers who believed that their students were intellectually competent taught these pupils in ways that honed their academic skills (Rubie-Davis, 2010).

Other related findings of the Rosenthal and Jacobson (1968) study were that teachers often based their decisions about student ability on how well these children met middle class standards. Teachers evaluated students based on their physical appearance skin color, clothing, and cleanliness - as well as their speech and their reading ability. Also, programs designed to improve the academic achievement of low-income students tended to create negative expectations in their teachers. The programs fostered the assumption by teachers that low-income students would be academically less competent than their middle-class peers. As a result, the educators treated the low socio-economic
students in ways that did not enhance their academic success (Rosenthal \& Jacobson, 1968).

These findings fueled more empirical research into teacher expectations and their effects on students' academic performance. Other studies confirmed and elaborated on the landmark study. For example, it was found that when teachers treated students very differently because of disparate teacher expectations, and favored more academically competent students over their less competent peers this compounded the negative consequences for the lower achievers (Kuklinski \& Weinstein, 2001). When the classroom climate was biased in favor of higher achievers, and when there were clear differences in the interactions between the teacher and the two levels of achievers this created overt and covert effects on children's ability to adjust to the classroom - Babad's study (as cited in Kuklinski \& Weinstein, 2001). Further, when teachers limited access to certain curriculum and also employed preferential grouping practices this caused students to receive disparate sets of skills and knowledge - Harris \& Rosenthal’s study (as cited in Kuklinski \& Weinstein, 2001). This was particularly relevant in the elementary school grades (Kuklinski \& Weinstein, 2001). Teacher expectations tend to vary according to grade levels as well as on the type (homogeneous versus heterogeneous) and size (small versus large group) of the classes. Since elementary school teachers have more sustained contact with their students in comparison with middle and high school educators their expectations are also more concentrated (Irvine, 1990).

Educational achievement also varied because teachers accepted sub par work from students for whom they had low expectations -Brophy's study (as cited in Rubie-

Davis, 2010), gave them less wait time -Jussim \& Harber’s study (as cited in Tyler \& Boelter, 2008), provided them with answers rather than giving them the opportunity to deduce the solutions with the help of reworded questions and hints -Tenenbaum \& Puck's study (as cited in Tyler \& Boalter, 2008), complimented non intellectual accomplishments while disparaging unsuccessful efforts -Warren’s study (as cited in Tyler \& Boalter, 2008), and paid more attention to the erroneous answers and unacceptable behaviors of their less academically capable students than to those of their more academically competent pupils - Cooper \& Good's study (as cited in Rubie-Davis, 2010). This kind of differential treatment contributed to the self-fulfilling prophecy effect of teachers on their individual students (Rist, 1970). As Rist (1970) notes, students who received favorable treatment from their teachers achieved well, and those who were treated less favorably had negative academic and emotional responses to this biased treatment.

Students who feel disliked by their teachers often react by disliking themselves and, by extension, become averse to school. "These students feel isolated, discouraged, and eventually fail academically" (Irvine, 1990, p. 48). These consequences are especially relevant for students from minority backgrounds as they rely heavily on teachers’ perceptions of them as they create their personal perspectives of their academic competency (Irvine, 1990).

As these examples show, the variations in teachers' treatment of students based on their achievement levels are quite distinct. The consequences of this inequitable treatment are particularly evident in the case of African American males. Although there
are exceptions, Garibaldi (2007) notes that in general, African American males do not perform well on state and federally mandated standardized tests. Additionally, they often are predominantly represented in unsatisfactory school statistics such as suspensions and expulsions, while having only minimal representation in academically challenging programs (Garibaldi, 2007). Crozier (2005) states that many teachers are primarily interested in improving the behavior of their minority male students. They are less concerned with helping them to be academically successful and are more interested in controlling these students' behavior. They constantly criticize actual and perceived infractions. These misplaced emphases often distort their expectations of these students’ academic abilities (Crozier, 2005).

There are serious ramifications of teachers having inaccurate expectations of their students since Irvine (1990) states that by the second month of school, many teachers develop expectations of their students that remain unchanged for the rest of the year. As Irvine (1990) notes,

The significance of the findings that teachers form stable, unchanging, and often inaccurate achievement expectations, in particular about black males, is of concern to educators. It may be that teachers’ expectations for black males’ achievement are more influenced by their stereotypes of black males as potential disruptions than by their academic ability. (p. 77)

The behavior of these teachers exemplifies the cultural and class dissonance that can occur between some African American male students and their female, White teachers. Many of these students often do not behave in ways that conform to White, middle-class mores. Tension and conflict are thus created between them and their teachers who adhere to these values and who reproduce them in their classrooms
(Villegas \& Lucas, 2002). Additionally, sometimes the culturally grounded mores and behaviors that minority students bring to school are not adequately represented in the pedagogical practices to which they are exposed (Boykin et al., 2006).

For example, Ladson-Billings (1997) notes that in most of our nation's schools mathematics is taught in a way that emphasizes predictable methods of achieving desired answers with little incentive to stray from the routine. Mathematics is taught in a way that sometimes makes it difficult for students to connect what they are learning with their everyday activities. Schorr, Firestone, and Monfils (2003) note that in many classrooms mathematics instruction involves much repetition with little or no real understanding of rules or strategies on the part of the students. This type of pedagogy is devoid of academic challenge and is irrelevant to children's everyday experiences (Schorr et al., 2003). It exemplifies the instrumentalist view of mathematics that pays scant attention to students' comprehension of underlying mathematical concepts or procedures (Draper \& Siebert, 2004). As Draper and Siebert (2004) note, NCTM found the lack of understanding that is promoted by the instrumentalist approach to mathematics instruction troubling and antithetical to their standards for mathematics education. Currently, many mathematics professionals are focusing on teaching mathematics in a way that is meaningful to students (Draper \& Siebert, 2004). This new instructional emphasis would facilitate the implementation of culturally relevant pedagogical practices such as those that Villegas and Lucas (2002) recommend.

Cultural modes of expression in mathematics instruction are lacking for all students, but this deficit may be particularly detrimental for African American students
because of their cultural preferences (Ladson-Billings, 1997). Weisglass (2002)
acknowledges that culture plays an important role in mathematics education. The sociocultural mores and historical contexts that influence mathematics instruction can greatly influence the mathematics learning and achievement of students, particularly those who have been exposed to consistent educational inequities (Weisglass, 2002).

Ladson-Billings (1997) summarizes some of these impacts:
Low-income African American students are more likely to be clustered in low ability mathematics classes. As a school's African American enrollment increases, the proportion of classes identified as high ability diminishes. Schools where African American students constitute the majority have less extensive and less demanding mathematics programs and offer fewer opportunities for students to take such gatekeeper courses as algebra and calculus that lead to increased opportunities at the college level and beyond. Schools with high concentrations of African American students tend to have fewer teachers judged to be highly qualified in mathematics. (p. 701)

Faced with these statistics that show minority students consistently falling behind in mathematics some researchers believe that teachers need to be encouraged to be more reflective about how their pedagogical practices address issues of equity (Lubienski \& Bowen, 2000; Struchens, Lubienski, McGraw, \& Westbrook, 2004; Rousseau \& Tate, 2003). They need to be encouraged to become proactive teachers with correct yet flexible expectations who are able to design pertinent learning opportunities for both their high and low achieving students (Irvine, 1990). Teachers' reflection about their instructional practices and the potential impact of their praxis on students is important. Failure to engage in this type of deliberation can lead to inappropriate responses on the part of teachers. All too often, such responses result in negative academic outcomes for students, especially African American males (Rousseau \& Tate, 2003). Irvine (1990)
posits that, "Certainly good teaching benefits all students’ learning, but under achieving Black students need to be taught by teachers recognized as effective and experienced" (p. 93).

In order to minimize or eliminate adverse effects on their students, teachers need to be encouraged to think about the type of instruction that would improve pupils’ academic performance. Ladson-Billings (1997) believes that a paradigmatic shift needs to be made by teachers in their notions of pedagogy so that more and more students are attracted to pursue higher levels of mathematics. To successfully teach mathematics, concepts may need to be rooted in their daily experiences. Their preferred culturally influenced learning styles should be taken into account when pedagogical strategies are considered. For example, Ladson-Billings (1997) states that many African American students view oral interactions very positively. In supporting this opinion, Kunjufu (2002) writes that male remedial minority students earned higher scores on tests when the questions were read aloud to them. Additionally, African American and Latino students understood history concepts much better if they listened to a recording of the text while reading it. The students’ academic performance was enhanced by the use of verbal rather than written interaction. Their affinity to oral communication can be partially attributed to the "call and response" feature of Black cultural interaction (Kunjufu, 2002, p. 98). During "call and response" the audience and the speaker are engaged in active communication with each other. This type of verbal interaction is part of the African American tradition of oral history and is integral to Black culture (Kunjufu, 2002). Given African American students' cultural preference for oral forms of communication,
interviewing them and discussing their interests could be very beneficial in keeping these students academically engaged (Ladson-Billings, 1997). This cultural dynamic supports the use of the clinical interview to ascertain African American students' understanding of certain mathematical concepts. The communication during the interview is primarily oral and as such aligns well with these students' cultural preferences.

However, as Rousseau and Tate (2003) point out, some teachers resist using this cultural dynamic in their pedagogical practices. They believe that education should be neutral and that treating all students equally is sufficient. However, equity for minority underachievers should not be based on "equality of process" (Rousseau and Tate, 2003, p. 213) or on treating all students equally. Rather, the focus should be on equality of outcome, on helping all students to perform to the best of their ability by providing them with the structures and strategies they need in order to be successful. Teachers need to be made cognizant of how cognitively and emotionally damaging, purportedly, neutral expectations can be to their students (Rousseau \& Tate, 2003).

For example, this neutrality when aligned with cultural and racial prejudices sometimes causes teachers to treat African American students differently than their White peers (Rousseau \& Tate, 2003). These researchers insist that teachers who deny that race matters when it comes to teacher expectations and academic outcomes are in essence refusing to acknowledge "...the race-related patterns in achievement and the potential role of racism in the underachievement of students of color" (p. 213). By insisting that students' skin color is irrelevant to their instruction, these teachers ironically perpetrate
and reinforce behaviors that culminate in poor academic performance by African American students (Rousseau \& Tate, 2003). Irvine (1990) concurs. She writes, By ignoring students' most obvious physical characteristic, race, these teachers are also disregarding students’ unique cultural behaviors, beliefs, and perceptions - important factors that teachers should incorporate, not eliminate, in their instructional strategies and individualized approaches to learning. Such treatment contributes to perceptions of inferiority about Black culture and life and to denial and self-hatred by Black children. (p. 54)

In contrast to these color-blind perspectives, Ladson-Billings (1997) believes that the battle to improve the mathematics performance of African American students needs to be fought from three perspectives:

The challenge of improving the mathematical performance of African American students must be fought on three fronts: programmatic, personal, and political. Programmatically, we must participate in the development of meaningful and challenging curricula. Personally, we must come to develop caring and compassionate relationships with students - relationships born of informed empathy, not sympathy. Politically, we must understand that our future as a people is directly tied to our children's ability to make the most of their education - to use it not merely for their own economic gain and personal aggrandizement, but rather for a restructuring of an inequitable, unjust society. (pp. 706-707) Ladson-Billings’ (1997) second and third goals can only be truly attained if teachers have positive expectations of their students regardless of their race or socio-economic status. Additionally, teachers need to recognize the importance of their pedagogical practices and expectations, as these factors not only affect African American students, but also have great significance for the continued health of our democracy.

## Conclusion

Teachers are extremely important in the lives of their students as they affect not only the academic achievement but also the self-esteem of their students, and this is especially relevant in reference to their Black students (Irvine, 1990). Therefore, a
positive change in teacher expectations can and does lead to improved intellectual performance. For example, classrooms of teachers with positive teacher expectations are associated with better academic performance and with students who have happier emotional dispositions (Baker, 1999). Also, the academic achievement of weaker students showed improvement with positive expectations, above the more academically talented students (Madon, Jussim, \& Eccles, 1997). Further, high teacher expectations are linked to improved educational achievement (Benner \& Mistry, 2007). This is in contrast with low teacher expectations that are linked to poorer performance (Lubienski \& Bowen, 2000) and which can be due to curricular variations to which students are exposed (Kuklinski \& Weinstein, 2001).

In today's inter-connected world, the continued success of the nation depends on its people's ability to compete in an increasingly global economy (Hunter \& Bartee, 2003). I support this perspective, and as our country strives to maintain its economic viability in our shrinking world, recognizing and addressing the deleterious effects of low teacher expectations on a significant number of its citizens become even more critical.

My emphasis on the importance of making African American students an integral part of the economic viability of our nation is deliberate. It is based on the realization that previous attempts at legislating academic equity such as Brown v. Board of Education (1954) and NCLB (2001) have not succeeded in eradicating the inequities in our educational system that still disproportionately plague African American students. While I am still deeply committed to reducing the achievement gap and to working towards academic equity for these students, I believe that pointing out the adverse effects
on the nation as a whole rather than on the group or individual members of the group such as the students in the study, will be more influential in effecting this change. In our world, money has often proven to be a more powerful motivator of change than appeals to morality or fairness. Hopefully a real threat to the earning power of a substantial number of Americans will create a full-fledged national drive aimed at eliminating or significantly remediating the underlying socio-economic causes of the educational disparities that negatively impact many minority students.

Further, I believe that the effects of teacher expectations on students' academic success should continue to be among the important factors that are considered when educators and policy makers ponder solutions to the achievement gap between African American students and their White peers. Extraordinary emphasis is placed on accountability testing in today's national educational framework (Linn, 2007; Ravitch, 2010). Tests continue to outline the discrepancies among the scores of the nation's students. Thus, given the direct impact of teacher expectations on student performance, the effects of these expectations on African American student performance should be an integral part of the ongoing, multi-faceted task of educational reform.

It is important that since all students are required to meet a mandated level of proficiency by 2014, (Linn, 2007; Ravitch, 2010) that we provide these students with teachers who are concerned about the effects of lack of equity on their students' academic performance (Rousseau \& Tate, 2003). Additionally, providing teachers with multiple avenues of gaining deeper insight into what their students really know could foster positive teacher expectations. A complementary assessment such as the clinical
interview could serve as a "window" into the abilities of historically poor performers. This detailed perspective of their students’ academic knowledge could help teachers recognize their student's abilities to succeed academically regardless of race or class.

## The Clinical Interview

## Explanation of the Clinical Interview

The clinical interview is a method for investigating the processes that support children's thinking. The clinical interview has been a part of studies about cognition since Piaget's experiments, but it has not been as closely examined or developed as more popular quantitative and qualitative methods (di Sessa, 2007). Ginsburg (1997) notes that despite being considered unscientific and methodologically deficient, it is being increasingly used in educational research. This is because it clarifies the thinking that accompanies children's problem solving attempts. Although standardized tests are excellent at providing actual scores as measures of achievement, they are not designed to expose the thinking behind the answers the students selected. Thus, while standardized tests measure learning outcomes or what children know, the clinical interview measures the processes by which students reveal their thinking or how they know (Ginsburg, 1997).

The main goal of the clinical interview is to find out about students' knowledge and how they solve problems. It is about discerning the true level of their ability in the concepts on which they are being examined (Nowak \& Gowin, 1984; Ginsburg, 1997). Following in Piaget's footsteps, during the clinical interview the interviewee is encouraged to reveal his or her innate ways of thinking about the issues or concepts. In order to do this, the interviewee has to comprehend the problems that are presented and
communicate this understanding to the interviewer (di Sessa, 2007). Through this process, clinical interviews provide insight about the specific strategies the interviewees used to make sense of the material on which they are being assessed (Nowak \& Gowin, 1984; Ginsburg, 1997; di Sessa, 2007).

Optimally, the clinical interview should be conducted in a comfortable place with minimal external distraction or interference (Ginsburg, 1997). The interviews are usually task-based (Goldin, 2000) and the interviewer has a variety of materials and manipulatives that are intended to facilitate the child's ability to answer the interview questions (Nowak \& Gowin, 1984; di Sessa, 2007). Nowak and Gowin (1984) explain that the items are helpful as they allow students to demonstrate their understanding and knowledge while precluding temptations on the part of the clinical interviewer to teach rather than assess the concepts. The length of the interview varies according to how much the child knows about the concepts but interviews should not go longer than thirty minutes for elementary school children in the lower grades (Nowak \& Gowin, 1984).

The clinical interview typically begins with the interviewer presenting a mathematical task to the child. The interviewer encourages the child to verbalize his or her problem solving efforts while making careful observations (Ginsburg, 1997). The participants interact with the interviewer and the tasks are therefore important elements of these observations (Goldin, 2000). As di Sessa (2007) shares, during the clinical interview the views or answers of the interviewee are not challenged, and the interview is conducted at a rate that allows for reflection. The latter usually occurs when the interviewee is asked to interpret, discuss, or affirm his or her responses (di Sessa, 2007).

The interviews are usually videotaped or recorded for analytic purposes (di Sessa, 2007; Ginsburg, 1997; Goldin, 2000).

Ginsburg (1997) states that as the child verbally shares his or her thinking, the interviewer makes hypotheses. These are derived from cognitive theories and illuminate the supporting strategies and prior knowledge that the child uses during problem solving. In order to test these hypotheses, additional questions that are tailor-made for the particular child are posed. Responses to these questions provide clues about student thinking as well as their approaches to and understanding of particular concepts. The interviews allow students to show the true extent of their knowledge about specific topics (Ginsburg, 1997).

When probing questions are used to ascertain the depth of students' conceptual understanding, the clinical interviewer needs to be able to react immediately to the child's response (Ginsburg, 1997). Goldin (2000) explains that this is because the clinical interview may proceed in an unanticipated direction and the clinical interviewer needs to be prepared for such contingencies. He or she needs to be able to provide "hints, related problems in sequence, retrospective questions, or other interventions" (p. 519). In order to maintain this flexibility and responsiveness while interviewing, the clinical interviewer must constantly monitor the child’s affective state (Ginsburg, 1997). Goldin (2000) says that the interviewer therefore needs to observe and accurately interpret verbal and nonverbal behaviors by the participants. These inferences allow the interviewer to learn more about "the mathematical thinking, learning, and/or problem solving of the subjects" (p. 519).

Using my professional experience, I will describe a scenario in which a clinical interviewer would need to deduce the reasons behind a student's answer. For example, if during a clinical interview a child solved a subtraction problem incorrectly, the interviewer would use his or her deductions to try to determine the reasons for the error. These deliberations could be influenced by possibilities like the following: Perhaps the child subtracted the smaller digit from the larger, regardless of its location in the problem. The child could possibly have had difficulty subtracting with zeros. Or maybe his or her mastery of basic subtraction facts was weak. The subsequent interactions with the student would be aimed at ascertaining which, if any, of these possibilities were responsible for the student's answer. The interviewer's interest goes beyond the answer itself to explore the thinking and reasoning that support it.

Thus, task-based interviews emphasize problem-solving processes and decrease the focus on identifying correct or incorrect answers (Goldin, 2000) as the clinical interview is about discerning the interviewee's ideas (di Sessa, 2007). Attention should be paid to the many ways students arrive at answers and not only on their actual responses as talking about their problem solving efforts provides them with the chance to share their mathematical reasoning (Fraivillig, Murphy, and Fuson, 1999). The clinical interview permits and fosters these displays of students' innate intellectual abilities (Ginsburg, 1997). Therefore, the clinical interview is one assessment option that could be used to facilitate mathematical reasoning and communication. As they solve mathematical problems during the clinical interviews, students' need to communicate their mathematical thinking or reasoning. Practice of this sort could help children hone
their mathematical literacy skills which are essential for them to be successful on today's ubiquitous accountability tests.

To conclude, the clinical interview is a qualitative research method used to explore children's thinking. In an individualized setting, children are presented with tasks and their problem solving efforts are observed. Emphasis is placed more on the process that leads to a solution rather than on the solution itself. Close attention is given to the verbal and non-verbal cues of the students, and their answers are probed but not challenged by the interviewer. The latter makes hypotheses about the children's knowledge that are explored through further questioning. The practical aspects of the clinical interview are supported by specific theories from the pioneering work of three famous scientists. These notable thinkers - Freud, Piaget, and Vygotsky, were the architects of the clinical interview method, and the cognitive ideas that anchor this research method were the result of their investigations (Ginsburg, 1997). They were the researchers who paved the way for using the clinical interview to engage children in a variety of activities that revealed their cognitive capabilities.

## Theoretical Validation for the Clinical Interview

The overarching goal of Freud's, Piaget's, and Vygotsky's explorations was to discover more about children's cognitive processes. They developed theories based on their observations of children. Freud's contribution was his belief in using observable data to make inferences and explanations about the hidden mental activities that created them. He was interested in the covert reasons for overt, observable behavior (Freud, 1938). "Freud, of course, was greatly concerned with the way primary processes
interfered with cognitive functioning. He concerned himself with the peculiar logic of dreams and neuroses, which is not unlike the conceptual thinking of young children" (Evans, 1973, p. 98). As previously mentioned, during the clinical interview, the interviewer makes deductions and inferences about the mental processes of the interviewee. These inferences are based on the answers (the observable data) that are given while the interviewee is engaged in solving a variety of problems. The data (the responses or answers) are used to provide insight about the covert or inner mental workings of the interviewee. These inner mental workings are the processes he or she used while engaging in problem solving. Freud was interested in children’s covert mental cogitations, but not as they appeared in problem-solving endeavors. He was concerned with how these cognitive processes affected children's dreams and neuroses. Like Freud, Piaget focused on the processes that affected children's cognition. However, unlike Freud, Piaget’s area of exploration was the strategies children used as they figured out answers to questions. Nonetheless, despite their different perspectives, they both examined procedures that enhanced cognition and those that derailed it. Based on his observations of children's cognitive abilities, Piaget formulated several theories that are still influential in education today.

Evans (1973) writes about the ideas that Piaget shared with him during an extensive interview. Piaget had formulated stages of cognitive development and believed that children developed differently during these phases according to their individual traits. He felt that all mental processes had to be constructed and this belief influenced the constructivist aspects of the clinical interview. Piaget was mistrustful of intelligence
and performance based assessments, and explained that this stemmed from his work with trying to standardize intelligence tests. As he proceeded with this endeavor, he became increasingly intrigued not only by a child's incorrect answer but also by the reasons that caused it. He therefore began to explore the factors that caused children to misunderstand certain concepts. This proved to be an endeavor that he found infinitely more fascinating than developing standardized tests (Evans, 1973).

This difference in emphasis (Piaget's vis-a-vis that of standardized testing) is significant for the clinical interview. It became integral to the clinical interview that sought to explore students’ thinking. The clinical interview method as pioneered by Piaget facilitated the exploration of children's correct and incorrect answers in a way that standardized tests could not (Ginsburg, 1997). Piaget was criticized for his unorthodox research techniques, but defended his methods by saying that to truly discover the differences between his cognitive stages he had to abandon predetermined experimental formats and follow his loosely formed hypotheses where they led (Evans, 1973). Piaget told Evans (1973) that this strategy was aligned with his belief that children should be allowed to develop at their own level and be afforded the opportunity to solve problems that fostered their abilities. He believed that this could be accomplished by using a judicious mix of direction and freedom (Evans, 1973). This belief influenced the clinical interview as its current format exemplifies this combination of structure and flexibility.

Piaget's House of Cards experiment reveals other elements that are still part of the clinical interview method as used today. In this experiment, Piaget provided the child with concrete materials (a pile of playing cards) and assigned the open-ended task of
using them to build a house. If the child was unfamiliar with the task, Piaget helped him or her get started. Then he allowed the child to work with little or no interference. If the child made several unsuccessful attempts, Piaget would ask questions to ascertain the child's thinking through the process. Finally, he analyzed his observations and described what he had learned about the child's cognitive development in a summary (Piaget, 1978). This methodology reflects the following elements that are still used in clinical interviews today. They are: providing the child with manipulatives; presenting the child with specific tasks; giving the child help if needed to get started; asking the child probing questions to reveal the child's thinking and reasoning; analyzing the observations of the child's cognitive processes; and summarizing the information about the child's cognitive structures and processes.

This approach epitomizes the pre-eminent aim of the clinical interview ascertaining the mental cogitations of the child regardless of how it may be represented (Ginsburg, 1997). During an experiment with a child, Piaget demonstrated the importance of this perspective. He noticed that a child stated that a shape that he, the child, had previously designated as a square was no longer a square. To try to understand this shift in belief, Piaget initiated another experiment to explore the reasons behind this occurrence more closely (Piaget, 1969). This type of exploration is a procedural precursor for the clinical interview. If the child's logic is obscure to the adult, the adult needs to discover the reasons for the schism in understanding. In this situation, the interviewer is required to ascertain what the child knows, doesn't know, or simply approaches from a different viewpoint (Ginsburg, 1997)

As he shared with Evans (1973) early in his work of exploring the cognitive structures and processes of children, Piaget thought that he was interrupting them too much thus preventing them from truly showing their independent knowledge. Piaget (1969) describes how he remedied that situation by allowing the children to use concrete materials during his interviews, and in this way he limited the amount of adult talking during the interview. Children were thus able to use the materials to solve the questions and in so doing revealed their cognitive processes without intrusive adult interference (Piaget, 1969).

In restricting the level of adult interference, and focusing on what children could do independently, Piaget's theories differed from those of Vygotsky, the third member of the theoretical trio. In trying to learn more about the cognitive developmental process in children, Vygotsky thought that this interest should take into consideration not only what children could do independently, but also what they could do with the help of an adult. He called this difference between independent work and work that was supported by or done in collaboration with adults, the Zone of Proximal Development (Vygotsky, 1978).

Like Freud and Piaget, Vygotsky was very interested not only in children's cognitive actions, but also in the underlying processes that supported these actions. Vygotsky (1978) believed that some of these procedures were especially evident when children encountered problems that were just slightly above their ability at a given point in time. When confronted with these situations, based on the level of their prior knowledge, they had choices. They could opt to solve the problem independently without recourse to any resource. They could use materials or manipulatives, ask for help
from the questioner, or simply verbalize their hypotheses to the object itself (Vygotsky, 1978). For example, in solving a problem during a clinical interview, children could implement the developmental strategy of drawing upon past experiences to help them solve something in the present. They could remember a mathematical procedure that they had learned in the classroom and use it to solve the current task. They could also make conjectures about how they would solve it in the future. That is, although they may have solved the question using a particular method, they could hypothesize about different ways a similar query could be figured out (Vygotsky, 1978). For Vygotsky, the goal was to distinguish between children's zones of actual and proximal development. Vygotsky believed that what the child could do in the present with adult help, he or she would be able to do independently at a later date (Vygotsky, 1978).

In sum, the covert and overt processes underlying children's cognition as well as their attempts at metacognition were of great interest to Freud, Piaget, and Vygotsky. Their theories in these areas have been pivotal in many of the elements of the clinical interview. For example, children are expected to engage actively during the problem solving activities with which they are presented during clinical interviews. However, being actively engaged includes reflection about the questions, and this is a necessary step towards arriving at solutions. The problems they attempt to solve are of varying levels of challenge and difficulty, so sometimes the children will be able to solve the problems independently, while at other times, they will need some assistance from the adult interviewer.

In essence, the theories that the early researchers held in common influenced the method and the purpose of the clinical interview to a considerable degree. In terms of method, the clinical interview has adopted the practice of posing different types of problems for children to solve and providing them with tools and help if necessary. In terms of purpose, its goal was and still remains to determine what children know and don't know. Freud, Piaget, and Vygotsky, uniquely yet complementarily, were instrumental in making the clinical interview a viable and useful tool for ascertaining what children know and don't know about a variety of concepts. In tribute to its value in this regard, the clinical interview has been used increasingly in educational research, particularly in mathematics education. Some examples of this research are explained in the following section.

## Empirical Support for Using the Clinical Interview in Educational Research

In recent empirical studies, from 1997 to 2010, clinical interviews were used to explore students' knowledge of various mathematical topics. For many of these researchers, the stated rationale for using clinical interviews was that they facilitated investigation of thinking processes. Students could be encouraged to elaborate on their statements and judgments, and this in turn allowed the researchers to make salient inferences about the students' underlying intellectual processes. Below are descriptions of studies that used the clinical interview. I focus specifically on the similarities and differences among these studies and the current study and discuss how these characteristics supported my decision to use clinical interview methodology in this study. Similarities

The most obvious feature of compatibility among the studies is that they all used task based clinical interviews as the primary methods of data collection. However, more specific similarities exist. A study conducted by Huntley et al (2007) shared the following factors with this dissertation. In addition to using task based clinical interviews, these researchers used some of the same sources in their data analysis discussion. Specifically, the theories of adaptability and versatility (Sfard and Linchevski, 1994) were utilized in both studies as was the rubric created by Malone et al. (1980). In creating the tasks for the clinical interviews, both studies used other assessments. For example, in selecting items for the clinical interview protocol, I used items from the NJPASS. Huntley et al. (2007) used items from previous NAEP and TIMSS (Trends in International Math and Science Study) assessments in their clinical interviews. They too had three primary sources of data of which the transcribed transcripts and student sheets were integral to the process of data analysis. These data sources provided the detailed description that was essential to analysis. Similarly, participants' answers were scored for answers given before probes were used, and codes were used to indicate responses that were received before and after probes. The effect of probes on participants' ability to solve the items was crucial to the findings of both studies. Other similarities with the (Huntley et al., 2007) study and the current study are not as critical but serve to underscore the methodological possibilities of the clinical interview. Their sample was on a small scale, and they too benefited from having access to schools from which to select their participants. Finally, the use of clinical interviews in the studies facilitated the gathering of information that went beyond students' scores in both studies.

Ginsburg and Pappas (2004) incorporated several corresponding elements in their study. Some of these were simply procedural. Examples include informing the children at the beginning of the clinical interview that they would be videotaped while they answered some questions. Additionally, the clinical interviewers managed the tasks by using a protocol as a template. Further, each task was read or paraphrased as frequently as necessary in order to allow the child to complete the items. Finally, one of the authors of the study was always present during the clinical interviews. From a more methodological standpoint, during data analysis in both studies, the zone of proximal development (Vygotsky, 1978) was calculated by finding the difference between the children's independent answers and the answers they achieved with adult help. The two studies - (Huntley et al., 2007) and (Ginsburg and Pappas, 2004) - were the most representative of the similarities that existed among the studies. There were many differences among the studies but they were in categories in which such disparities would be expected. While the differences seem trivial when viewed independently, when viewed as a group, they serve to underscore the many methodological possibilities that can be utilized in clinical interview research.

## Differences

Most of the differences dealt with sample type, size, and composition in terms of ethnicity, race, gender, academic ability, and educational level (Dunphy, 2006; Groth, 2005; Singh, 2000; Huntley et al., 2007; Inoue, 2005; Empson \& Turner, 2006). As with other research methods, the clinical interview provides the researcher with great flexibility in tailoring the sample to match the specific needs of the inquiry. However,
for purposes of comparison with the current study, the most notable difference was the use of dynamic assessment in some of the clinical interviews in the study Charles and Nason (2000) conducted. This was an important aberration because while I deliberately chose the clinical interview method over dynamic assessment, these researchers combined the two methods by including limited teaching episodes with the clinical interviews. The fact that Charles and Nason (2000) chose to combine these two complementary assessment methods in their study serves to delineate a key differentiating factor between the two methods. Using dynamic assessment meant that the children were exposed to specific instructional moments during the clinical interview. In order to be able to include an instructional element in their study, Charles and Nason (2000) had to include dynamic assessment, as teaching students is not a part of the usual procedure in the clinical interview.

## Significance of Similarities and Differences

The above descriptions of similarities and differences among these diverse studies and the current study highlight that the clinical interview is an effective and increasingly popular qualitative research method. It facilitates the exploration of a plethora of mathematical topics. The clinical interview can be used with participants who range from preschool to college and who are of various ethnicities, races, and gender. Its flexibility is not limited to topic and participant selection, but the method allows the researcher to choose the quantity and duration of the sessions, and how to establish interrater reliability. The ability of the clinical interview to reveal more about how students think has led to its growing use in empirical research and these studies support my
decision to use it in this current study. An important constant among these studies is that they all obtained deeper insight into students’ mathematical thinking by using clinical interviews. This factor further sustained my decision to use the clinical interview in my study as opposed to another method, namely, dynamic assessment.

## Dynamic Assessment

Clinical interviews should not be confused with dynamic assessment despite some overt similarities. Haywood and Lidz (2007) write that dynamic assessment has been described as "a subset of interactive assessments that includes deliberate and planned meditational teaching and the assessment of the effects of that teaching on subsequent performance" (p. 2). An essential feature of dynamic assessment is that the ultimate goal of the assessor is to improve the performance of the person who is being assessed (Haywood \& Lidz, 2007). Further, Elliott (2003) asserts that the zone of proximal development (Vygotsky, 1978) is the crux of the dynamic assessment method and states that educational psychologists seized on dynamic assessment as a way of addressing their dissatisfaction with aspects of intelligence testing (Elliott, 2003). Thus, dynamic assessment shares with the clinical interview method a theoretical mentor in the person of Lev Vygotsky, and the psychologists’ disenchantment with intelligence testing is reminiscent of Piaget's mistrust of that type of assessment. However, there are important differences between the clinical interview and dynamic assessment.

Grigororenko and Sternberg (1998) explain that dynamic assessment involves a combination of assessment and intervention. The examiner gives feedback to the examinee until he or she either succeeds or fails at the task, and in this way assessment is
aligned with teaching. Psychologists use dynamic assessment to test the examinee's ability to reuse skills and knowledge that he or she had learned in previous dynamic assessment scenarios (Grigorenko \& Sternberg, 1998). The tester determines how much help to provide the testee based on the latter's efforts as well as on his or her ability to change (Elliott, 2003).

Consequently, in addition to questioning and evaluation, intervention and change are also important aspects of dynamic assessment. The most prevalent description of dynamic assessment is "active intervention by examiners and assessment of the examinees' response to the intervention" (Haywood \& Lidz, 2007, p. 1). Dynamic assessment is used not only to test but also to teach and develop cognitive skills, and to measure children's subsequent cognitive development (Grigorenko \& Sternberg, 1998). While children may learn certain skills during a clinical interview, that is not a primary focus of the assessment. This is in contrast with dynamic assessment where the emphasis is on teaching cognitive skills and then assessing how well they have been learned.

Like the clinical interview, dynamic assessment is not intended to replace other types of assessment. Haywood and Lidz (2007) note that dynamic assessment is complementary to other commonly used methods of evaluation such as observations. It is particularly useful when it is implemented with other types of assessment such as standardized testing and clinical interviews. Combining dynamic assessment with these other assessment methods allows the psychologist to find answers to questions that these methods alone can't answer and enhances the information gained from them. In this way, a benefit of dynamic assessment is that "it can add information about present and
potential performance that is not readily (or even at all) obtainable from other sources" (Haywood \& Lidz, 2007, p. 2). This type of assessment is often a better predictor of future educational output (Elliott, 2003). Many experts in the field of dynamic assessment believe that it is particularly beneficial in situations where standardized test scores are poor, where there are a variety of language issues, or where there is cultural dissonance (Haywood \& Lidz, 2007). While these characteristics may seem identical to those of the clinical interview, the key difference remains the disparity in their intent and methodology. Dynamic assessment teaches strategies to improve cognitive skills while the clinical interview's emphasis is not on instruction but on understanding children's thinking. Dynamic assessment often works best in conjunction with the clinical interview, not as a replacement for the latter.

Haywood and Lidz (2007) state that unlike standardized assessments, dynamic assessment is not neutral. Practitioners deliberately manipulate both the test and the testee so that the latter can be moved along on a continuum of learning outcomes. In this way dynamic assessment is intended to explore the cognitive processes that fuel learning. It is focused on collaboratively developing strategies that help individuals overcome barriers to academic competence. Therefore, a critical factor of dynamic assessment is to identify "the cognitive and metacognitive processes that need improvement, the tactics that lead to enhancement of such processes, and the differential responsiveness of individuals to intervention designed to enhance their development and application" (Haywood \& Lidz, 2007, p. 26).

In sum, Haywood and Lidz (2007) note that dynamic assessment tries to make informed suppositions about the challenges that people may encounter in showing their intelligence and tries to find ways to improve their academic and social aptitudes. It tries to identify the cognitive and metacognitive processes that need improvement as well as the strategies that enhance them. It recognizes that standardized intelligence tests are not always accurate and are often unable to recognize the potential for learning (Haywood \& Lidz, 2007). These principles underscore its differences with the clinical interview. Dynamic Assessment and the Clinical Interview

Dynamic assessment is targeted towards clinical and school psychologists and not towards regular education teachers. This is important because the teachers do not have the technical training required to administer dynamic assessment sessions. Since my study aims to see what teachers can learn from the clinical interview, dynamic assessment would not be helpful in this case as it would be more likely to be used by a child study team and not by regular educators. School psychologists use dynamic assessment to determine the kinds of modification in school programming that are required to help learning disabled and other disadvantaged students succeed. Regular classroom teachers would use the clinical interview to better understand how their students were thinking as they solved problems. Whereas with dynamic assessment intervention and assessment are inseparably linked, in the clinical interview, despite the fact that the interviewer can use probes, there is no active teaching of cognitive strategies. The child will not be evaluated in order to determine how much his or her cognitive functions were ameliorated by previous interventions.

The premise of dynamic assessment is to evaluate, intervene, and change. The clinical interview is designed to allow the interviewer and the interviewee to see how much knowledge the latter really knows about a topic. It is about data gathering and gaining insight into underlying cognitive processes and strategies. Dynamic assessment focuses on using modifications to induce learning and to change the cognitive functions of the learner. In essence, dynamic assessment is used to determine the level of intelligence of the learner in order to determine special education or other program modifications. It makes inferences about inhibitions to intelligence and seeks to improve academic and social learning. The clinical interview focuses on ascertaining the actual depth of academic knowledge in particular subject areas, primarily mathematics.

Clinical interviews are preferable to dynamic assessment in this study because of its goals. This study focuses on understanding the differences in performance on mathematical concepts as evidenced by standardized tests that deal with the products of learning. This is in contrast with clinical interviews that delve into the processes of learning. This study's primary goal is not about improving the cognitive functions of the potential participants, but is rather about gaining insight into their mathematical understanding. Additionally, the other participants are primarily regular education teachers and not clinical or school psychologists. The clinical interview is therefore more suitable for supporting the research goals of my study.

## Conclusion

Empirical evidence indicates that for a variety of reasons, African American students from diverse socio-economic backgrounds generally do not perform as well as

White peers on standardized tests. With the introduction of state mandated tests across the nation in response to NCLB requirements, there is increasing pressure on schools and teachers to ensure that students are successful on these assessments. Studies that try to discern the reasons for the success of some African American males on standardized tests could be very beneficial to these efforts. Educators could use this discourse to effect changes in their pedagogical practices. Teachers are critically important to the success of their students on standardized assessments. The curriculum that they use with children affects the latter's ability to answer questions accurately. The caliber of teachers and the level of their experience and effectiveness are also very important factors affecting students' success.

Despite the differences in their academic abilities, the students in this study were exposed to a broad curriculum. They were beneficiaries of the expertise of highly qualified teachers who had many years of teaching experience at the time of the study. The educators were accustomed to reflecting on their pedagogical praxis and found the information that the clinical interview provided interesting and useful. They welcomed the complementary perspective of their students' abilities. The constructivist elements of the clinical interview method allowed them to gain better insight into not only what the students knew, but also about how they engaged in problem solving. In this way, the clinical interview was a better methodological choice than dynamic assessment for this study. This literature review provided the theoretical background of this study. The next chapter outlines the various methodological procedures that were implemented in the study.

## CHAPTER THREE

Methodology

This qualitative study used the clinical interview to explore the depth of mathematical knowledge of African American students. It was also an investigation to ascertain what teachers might learn about the mathematical knowledge of selected students, and whether or not the information generated by the clinical interviews would change their views about the students’ math abilities. The focus of the study was on what new, complementary information could be gathered by using the clinical interview. The specific research questions were: 1) Will this complementary assessment show that students interviewed actually know more about mathematics than they are able to reveal on the standardized test? 2) What might teachers learn about their students’ mathematical ability, and how might this knowledge affect their views of those students?

The information from the transcripts of the clinical and teacher interviews facilitated data analysis and the quest for answers to the study's questions. Finding similarities, differences, themes and patterns in the data also evolved from this process. Clinical interview methodology allowed me to link students with their teachers and to see how these connections affected the findings.

## Participants

The participants were seven second-grade African-American male students enrolled in two elementary schools in a suburban district in New Jersey. In both schools, the second grade cohort of African American boys was quite small this school year,
including only these seven boys - the minimum number proposed for the sample. The student sample was homogeneous in terms of ethnicity and grade level. The students ranged in age from seven years and eight months to seven years and eleven months (Appendix A, p. 313). The other participants in the study were the students’ five homeroom teachers as well as the special education teacher of one of the students in the study (Appendix B, pp. 314-316). These six teachers were White (non-Hispanic).

The participants were selected for the study by using convenience sampling (Patton, 2002, p. 241). The sample was selected because it could provide in depth information about a particular group. It was also selected because the two schools were geographically near each other thereby facilitating the scheduling of the clinical and teacher interviews. Since less travel time was needed, I was able to maximize the duration of the interviews. Finally, the sample was chosen because I knew, and was known by, all of the adult participants as well as by some of the students and their parents. I had hoped that this mutual familiarity would make them willing to participate in and help with the study, and this proved to be the case.

In order to observe ethical and confidentiality requirements, all participants were given pseudonyms that were used throughout the study. In addition, when the teachers viewed the DVDs of their students' clinical interviews, they did so in my presence in order to maintain the students’ confidentiality.

## Research Setting

The study was conducted in E and B Schools in the Sligo school district. (The names of the school district, teachers, and students have been changed to respect the
privacy of the participants.) The Sligo school district is located in north central New Jersey about 27 miles west of New York City. The town is in a suburban setting and has a diverse population of approximately 30,000 . Sligo is a regional school district that serves approximately 5, 500 students from Preschool through Grade 12. The district has five elementary schools (Preschool-Grade 4), two middle schools (Grades 5-8), and one comprehensive high school (Grades 9-12). This information is based on the district's website, which for reasons of anonymity cannot be identified.

During the 2010-11 school year, the district enrolled a total of 426 students in the second grade. Of these, 172 were White males, 146 White females, 17 Black males, 17 Black females, 10 Hispanic males, 11 Hispanic females, 18 Asian males, 16 Asian females, 1 Native American male, 0 Native American females, 1 Hawaiian male, 0 Hawaiian females, 8 males of two or more races, and 9 females of two or more races (New Jersey Department of Education, 2010-2011).

E School serves preschool through fourth grade students and had an enrollment of 420 students. English was spoken at home by $97.1 \%$ of the students, with other languages being Cantonese (0.7\%), Albanian (0.5\%), Spanish (0.5\%) and others 1.2\%. The total number of students in Grade 2 was 83, including five of the seven African American boys in the study. The student/faculty ratio was 13.9, and the percentage of teachers with a bachelor's degree was $48.4 \%$, while the percentage of those with a master's degree was 51.6\%. There were no suspensions or expulsions for the 2009-2010 school year. The attendance rate for Grade Two was $96.3 \%$, and the student mobility rate (students leaving or coming during the school year) was $3.8 \%$. The percentage of
students with disabilities was 14.8\% (New Jersey Department of Education - 20092010).

B School enrolled 407 students, from preschool level through fourth grade. English was spoken at home by $92.9 \%$ of the students, with other languages being Mandarin (1.2\%), Hindi (0.7\%), Russian (0.7\%), and others 4.4\%. A total of 67 second graders were enrolled in this school, including two African American boys who participated in the study. The student/faculty ratio was 12.4 . The percentage of faculty with bachelor's degrees was $34.4 \%$, those with master's degrees were $62.5 \%$, while those having doctoral degrees were $3.1 \%$. There were no suspensions or expulsions for the 2009-2010 school year. The attendance rate for Grade Two was $96.2 \%$, the student mobility rate was $5.9 \%$, and the percentage of students with disabilities was 23.1\% (New Jersey Department of Education - 2009-2010).

## Data Sources

## School Report Card

Descriptive statistical data was collected from the schools’ New Jersey School Report Card (New Jersey Department of Education, 2009-2010). This document included demographic data on the Sligo school district such as the total number of students in the grade level, the ratio of faculty to student, ethnic and linguistic composition, attendance and mobility rates, standardized test scores for grades three and above, as well as expulsion and suspension rates. These data helped to paint a clearer picture of the research setting.

## Mathematical Concepts for Teacher Interview \#1 Sheet

Prior to conducting the study, I piloted the two teacher interviews with a teacher who was not part of the study. She found it challenging to make an accurate assessment of her student's performance during the pilot for the first teacher interview because she had not had the opportunity to review the student's work and record of progress prior to our meeting. Based on this experience, I decided to give the participating teachers a sheet listing the mathematical concepts (Appendix C, p. 317) to complete prior to the first teacher interview. I told them that during the first teacher interview I would review their evaluations of their students' competency in those concepts. To do these evaluations they used a rating scale with which they were all familiar.

## Teacher Survey \#1

The first teacher interview (Appendix D, pp. 318-320) asked each teacher to indicate the level of proficiency she had assigned her student for each of the categories listed on the Mathematical Concepts sheet that had been provided. These proficiency ratings were carried out with the use of a scale from the Chicago Math Program used district-wide. The scale is based on ascertaining whether a student is at a secure, developing, or beginning stage of understanding of a particular concept.

The Everyday Math Program, published by the University of Chicago, presents concepts in a spiraling manner, revisiting them at different times in a grade level as well as in the grade levels before and after. As a result of the circular nature of the program, teachers use the scale to ascertain the levels of difficulty of the skills or concepts they are
teaching at different times throughout the school year. There are three levels of challenge that are common throughout the program regardless of the grade level.

The three levels are secure, developing, and beginning. The same skill can have varying levels in different grades. For example, in first grade, double-digit addition with renaming would be classified as a beginning skill, while in second grade it would be classified as a developing skill, and in fourth it would be classified as a secure skill.

The student who has a thorough understanding of a concept or skill is considered to have attained a secure level. A student who has some familiarity with the concept but has not totally mastered it is classified at the developing level. The student who is just learning a concept for the first time is classified at the beginning level.

## Post Interview Sheet

The second teacher interview (Appendix E, pp. 321-322) was used after the teacher participant had viewed the clinical interview of her student. The questions on the sheet were used as prompts to elicit what they had learned about their students' performances from watching the clinical interviews. During the discussions, teachers shared whether their students' proficiency levels were the same as they had previously indicated or if they were different.

## Teacher \& Clinical Interviewer Notes

During the interviews, the teachers shared information about their students. Usually they wrote these notes themselves, but sometimes I wrote them as the teachers spoke. The notes covered material that had not been addressed on either the teacher surveys from the first teacher interviews or the sheet used for the second teacher
interviews. This information provided insight about the students' performances during data analysis.

The clinical interviewer made a few notes during the clinical interviews with students. The notes were simply observational. In them, she commented on the manner in which some of the students approached certain problems, or made hypotheses about the reasons for a student's answer. These notes were referenced during inductive coding as I searched for themes, patterns, and relationships among the data.

## Clinical Interview Protocol

The clinical interview protocol (Appendix F, pp. 323-328) consisted of items that evaluated some of the same concepts as those presented on the New Jersey Proficiency Assessment of State Standards (NJPASS). The problems on the protocol were similar but not identical in wording and content. The sheet with the pictures for the symmetry item was given to the students when that question was read to them. The sheet was collected upon the completion of each clinical interview.

## NJPASS Scores

The NJPASS scores were presented in two different formats. One format was the Individual Performance Profile. This profile presented each student's individual scores in Mathematics and Language Arts. The mathematics section was divided into five broad categories: number sense, operations, and properties; measurement; spatial sense and geometry; data analysis, probability, and discrete mathematics; and patterns and algebra. This section gave the total number of points possible on the NJPASS, the student's score
and the percentage of the total that it represented. Finally, the student's performance level on the NJPASS was given.

The second format of the NJPASS scores was the Item Response Record. This record presented the individual items that were included in the broader mathematical categories of the Individual Performance Profile. For example, under the category of number sense, operations, and properties, some of the individual items examined subtopics such as: models for representing simple fractions; operations with money; and place value, counting, and grouping. This record also indicated whether the item was multiple-choice or open-ended. A correct response was indicated by a check mark, while an incorrect response was indicated with an alphabetical notation. The Item Response Record facilitated individual item comparison with the items on the clinical interview protocol.

## Procedure

## Teacher Interview \#1

On April 29, 2011, I conducted the first interviews with all six teachers. I met with the teachers individually in meeting rooms adjoining the offices in the two schools. I interviewed the teacher at B School first, and then moved to E School. The interview was scheduled during each teacher's preparation period. The clinical interviewer videotaped all the teacher interviews. These tapes were then formatted on to DVDs.

At the beginning of the interview, I reiterated the purpose of the study and what I hoped to accomplish during the interview. The teachers had come prepared with the

Mathematical Concepts for Teacher Interview \# 1 sheet (Appendix C, p. 317) that I had given to them prior to the interview.

The teachers were given copies of Teacher Survey \#1 (Appendix D, pp. 318-320) that had the same concepts listed as Appendix C (p. 317). This survey had columns labeled secure, developing, and beginning. The teachers chose the levels they thought their student exhibited for each of the concepts by checking the appropriate column. As the teachers indicated these levels, they often gave reasons for their choices beyond gauging the students' levels of proficiency for the different mathematics concepts.

On the survey, I asked the teachers to indicate the strategies that the students used during computation. Specifically I asked if the students completed basic algorithms mentally, or if they required other materials (fingers, paper and pencil, manipulatives) to help them. I also queried whether these children explained their thinking orally, in writing, using diagrams, or by using a combination of these methods. Additionally, I asked what resources the teachers used to arrive at their evaluations. This was to ensure that evaluative judgments were supported by actual student work, such as workbook pages and tests, as well as teacher data from conferences and observations. At the end of these first teacher interviews, I collected the Mathematics Concepts for Teacher Interview \#1 sheets and the teacher surveys.

## Clinical Interviews

On May 2-3, 2011, each of the seven student participants was interviewed individually in rooms next to the offices of their respective schools. Despite the fact that business was being conducted as usual in these offices, the students ignored any
extraneous sounds. The clinical interviews were conducted after the NJPASS so as to limit any carry-over effects of the clinical interviews to the standardized test. The clinical interviews were done at pre-determined times during the school day. We worked around the children's schedules to be minimally disruptive of their school activities. While parents had explained the Assent Form (Appendix L, pp. 353-356) to them, I reiterated its contents prior to starting the interviews and had the students sign the forms.

An experienced clinical interviewer conducted the clinical interviews. She is a former elementary school teacher who taught at the elementary school level for over 30 years, including second grade. Until her recent retirement, she was the K-5 Math and Science Supervisor for the Sligo school district. The individual selected to conduct the clinical interviews had been part of two empirical studies done in the target district that used clinical interviews, serving as the clinical interviewer for both. She co-authored a study with the district's current Assistant Superintendent and Dr. Herbert Ginsburg. The study investigated using the clinical interview as a component of lesson study. Since she had done all the clinical interviews in the district to date, I valued her input with reference to the protocol items. Based on her recommendation, I piloted the protocol with a student not involved with the study.

Students received a clinical interview protocol (Appendix F, pp. 323-328) on which they recorded their answers with a pencil. The protocol had fourteen items with a total of thirty-one questions. The items were open-ended as the students were not given multiple choices from which to select an answer. However, many of these items only required them to write a single number or number model, not sentences or paragraphs.

All the items on the clinical interview protocol were read to the students, and in many cases reread as needed during the clinical interview. The necessary manipulatives were easily accessible on the table at which they sat.

The clinical interviews all lasted between 41 and 45 minutes, longer than had been originally anticipated. The students took more time on two of the tasks than the pilot clinical interviewee had done. Only the clinical interviewer and I were present during the clinical interviews that I videotaped. Upon entering the room, the students were told that the videotaping was to ensure that what they told us would not be forgotten. The videotapes of the clinical interviews were formatted on to individual DVDs to make it easier for the teachers to watch their students, to maintain student confidentiality, and to facilitate verbatim transcription and subsequent analysis.

The decision to conduct only one clinical interview was based on the following factors. In reviewing other studies using clinical interviews for data collection, I found many that used only one interview regardless of grade level (Cooper 1998; Empson \& Turner, 2006; Ginsburg, Mast \& Snow, 2009; Inoue, 2005). A study by Dunphy (2006) included two interviews, but this was the exception. These studies, using a variety of age groups, found that one interview generated enough information for their research goals to be met. The young age of the participants in this study made one interview preferable because they had taken the NJPASS in May. I was concerned that more than one interview under these circumstances would simply be too much evaluation for them, and that they would be unable or unwilling to put in the requisite effort as they participated in the clinical interviews.

## Teacher Interview \#2

On May 4-5, 2011, I conducted the second teacher interviews. I arranged for a substitute teacher to cover their classes while they participated in the interview. My goal was not to deprive teacher participants of their preparation periods during the school day or to take their time after school. Consequently, the teachers did not lose any of their professional or personal time to meet with me. The second set of teacher interviews was also videotaped and subsequently formatted on to DVDs. In addition to the teachers, only the clinical interviewer and I were present, so confidentiality was maintained. Our roles were reversed in that while I taped during the clinical interviews, the clinical interviewer taped the teacher interviews that I conducted. Being familiar with the teachers and the grade level, she also sometimes asked questions or made comments. All the teacher responses from the two teacher interviews were collected within the same time frame, two weeks, which was important for data comparison purposes.

During these second teacher interviews, the teachers watched the DVDs of their students' clinical interviews. I gave them guidelines on what to keep in mind as they watched (see Appendix G, pp. 329-330). I also encouraged them to comment on and take notes about anything that caught their interest as the DVD of their individual student's clinical interview was played. Thus, the second interview with the teachers was less scripted. The discussion flowed more freely as the teachers were able to discuss topics more randomly.

Unlike the first teacher interviews, the teachers were not apprised of the questions that would be asked although they did know they would be watching their students'
clinical interviews. They were provided with copies of the clinical interview protocol so that they could follow along with ease. This allowed them to make comparisons with the items on the NJPASS. After they finished viewing their student's DVD, I questioned them about what they had seen. I noted these answers and collected any notes they had made. The teacher interviews ascertained the teachers' evaluations of their students' mathematical abilities at two points in time - before and after the clinical interviews.

At the end of the week I asked the teachers to send home a DVD copy of the clinical interviews with each student, so their parents could see what had transpired during the interviews. These copies were visual confirmation that the clinical interviews had been conducted properly and ethically, and gave the parents insight into how their sons had performed.

## Data Analysis

## Organizing the Data

Verbatim transcriptions of the clinical and teacher interviews provided what Patton (2002) called "thick description" (p. 437). The research questions guided the development of the findings. The results were chronologically organized according to when the data collection was conducted. They started with the first teacher interviews, were followed by the clinical interviews, and ended with the second teacher interviews.

In preparing the findings, I examined the transcripts looking for data that would be useful in answering the study's research questions. The main foci of the analyses of the transcripts were the ratings the teachers gave their students, how the students performed on the clinical interviews compared to the ratings, and whether or not the
teachers maintained their opinions of their students' academic abilities. Examples and details from the transcripts that supported the data were included. These samples allowed me to identify common themes and patterns among the data.

After I finished examining the individual responses of the students and teachers, I compared them. These analyses highlighted the similarities and differences among the students' performances as well as the teacher interviews. There was a comparison of the students' problem solving strategies as well as an examination about how their performances in the clinical interviews were affected by probes. Patterns and themes that emerged from data analysis were examined and any connection to the research questions was explained. Information from the teachers' feedback was connected to their particular student as well as to their colleagues.

## Unit of Analysis

The unit of analysis used during coding was a speaking turn by one of the participants. These speaking turns also included one or more idea units. Speaking turns and their embedded idea units were represented by individual words, clauses, complete sentences, and symbols. The speaking turns were coded as a whole. Their individual idea units were not coded separately. One example is: "...because it’s the smallest there." This is one speaking turn with one idea unit represented as a clause. It was coded as a correct answer. Another example is: "Doing well, but developing because I didn’t feel there was enough time to have it become a secure skill." This is one speaking turn with three idea units (doing well; developing; didn't feel there was enough time to have it become a secure skill) represented as phrases. It was coded to indicate that the teacher
had assigned a rating of developing to a student. Specific units of analysis for each coding category are outlined in Table 1.

## Table 1

Specific Units of Analysis for Each Coding Category

| Coding Category | Unit of Analysis |
| :--- | :--- |
| Correct/no probe | Sentence |
| Correct/with probe | Individual Word |
| Correct/with rating scale | Sentence |
| Incorrect/no probe | Phrase |
| Incorrect/with probe | Sentence |
| Incorrect/with rating scale | Phrase |
| Rubric | Symbol |
| Cognitive Strategy | Sentence |
| Zone of Proximal Development | Individual Word |
| Themes | Sentence |
| Versatility | Symbol |
| Adaptability | Sentence |
| Conceptual Categories | Individual Word |
| Teacher's Responses | Phrase |
| Math Rating Levels | Individual Word |

## Coding Categories

The coding categories, Table 2, emerged from the data, had constructivist influences, and were derived from assessment materials. The categories such as "correct/no probe" emerged from the data during open coding as I read through the transcripts from the clinical and teacher interviews. Some - the rubric and the Zone of Proximal Development - were influenced by the constructivist theories of Piaget (1969, 1978), Vygotsky (1978).

Table 2
Common Coding Categories
$\left.\left.\begin{array}{|l|l|l|}\hline \text { Category } & \text { Description } & \text { Examples } \\ \hline \text { Correct/no probe } & \begin{array}{l}\text { Student got the item correct } \\ \text { with no adult intervention. }\end{array} & \begin{array}{l}-9-3=6 \\ -6 Q \text { 's, 4D's, 1 nickel }\end{array} \\ \hline \text { Correct/with probe } & \begin{array}{l}\text { Student got the item correct } \\ \text { with adult intervention. }\end{array} & \begin{array}{l}\text { Student answered "Trent" to } \\ \text { a question asking for the } \\ \text { greatest amount. After a } \\ \text { probe the student correctly } \\ \text { answered "7." }\end{array} \\ \hline \text { Correct/rating scale } & \begin{array}{l}\text { Student got the item correct } \\ \text { and had been rated at } \\ \text { secure, developing, or } \\ \text { beginning levels by their } \\ \text { teachers. }\end{array} & \begin{array}{l}\text {-Student circled ring, flower, } \\ \text { and bib on the symmetry } \\ \text { sheet. This answer was } \\ \text { correct, and the student had } \\ \text { been given a secure rating. }\end{array} \\ & & \begin{array}{l}\text {-"I did it backwards. } \\ 3+6=9 . " ~ T h e ~ s t u d e n t ~ h a d ~ \\ \text { solved the problem correctly } \\ \text { and had been given a } \\ \text { developing rating. }\end{array} \\ \hline \text { Incorrect/no probe } & \begin{array}{ll}\text { Incorrect/with probe } & \begin{array}{l}\text { Student got the item } \\ \text { incorrect despite adult }\end{array} \\ \text {-"Brownies, oatmeal raisin } \\ \text { cookies, chocolate chip } \\ \text { cookie." The student got the } \\ \text { problem correct and had } \\ \text { been given a rating of } \\ \text { beginning. }\end{array} \\ \text { incorrect after working } \\ \text { independently. }\end{array} \quad \begin{array}{l}\text { "This one is like small, and } \\ \text { this one is like a little }\end{array}\right\} \begin{array}{l}\text { "100 more." This answer } \\ \text { was not a reasonable answer } \\ \text { in the context of the numbers } \\ \text { in the item that were 139-89. }\end{array}\right\}$
\(\left.$$
\begin{array}{|l|l|l|}\hline & \text { intervention. } & \begin{array}{l}\text { bit...inch." } \\
\text { "An inch. Ok. And is a } \\
\text { centimeter the same as an } \\
\text { inch?" } \\
\text { "Um, no." }\end{array}
$$ <br>

"No? Ok. So...which one\end{array}\right\}\)| In you think is the |
| :--- |
| centimeter?" |

$\left.\begin{array}{|l|l|l|}\hline & & \begin{array}{c}\text { House. 2+4=6+3 is } \\ \text { 9." The student } \\ \text { missed the key words } \\ \text { in the directions that } \\ \text { asked for 'as many } \\ \text { rides as possible' and } \\ \text { gave three rides } \\ \text { instead of four as his } \\ \text { answer. }\end{array} \\ \hline \text { Incorrect/rating scale } & \begin{array}{l}\text { The student got the problem } \\ \text { incorrect and had been rated } \\ \text { at secure, developing, or } \\ \text { beginning levels by } \\ \text { teachers }\end{array} & \begin{array}{l}\text {-"Haunted House, Mirror } \\ \text { House, and Goblin Ride" - } \\ \text { the student got this item } \\ \text { incorrect despite receiving a } \\ \text { rating of secure. }\end{array} \\ \hline \text { Rubric } & \begin{array}{l}\text { ""6/9" - the student got the } \\ \text { item incorrect and had a } \\ \text { rating of developing. }\end{array} \\ \text {-"8:30" - the student got the } \\ \text { item incorrect and had a } \\ \text { rating of beginning. }\end{array}\right\}$
$\left.\begin{array}{|l|l|l|}\hline & & \begin{array}{l}\text { - "I don't know the answer } \\ \text { to this one." The student } \\ \text { was unable to begin the } \\ \text { problem. }\end{array} \\ \hline \text { Cognitive strategy } & \begin{array}{l}\text { Student used the strategies } \\ \text { listed below to solve the } \\ \text { items. They were either } \\ \text { used singly or in a } \\ \text { combination: } \\ \text {-Mental Math } \\ \text {-Manipulatives } \\ \text {-Paper and Pencil } \\ \text {-Fingers } \\ \text { The most common } \\ \text { cognitive strategy used by } \\ \text { the participants was mental } \\ \text { math. } \\ \text { Manipulatives were } \\ \text { required for the item based } \\ \text { on the coins, so all students }\end{array} & \begin{array}{l}\text { 3+6 is 9, 9+4 is 13, 13+7 is } \\ \text { 20 + 5 is 25 +4 is 24, is 29." } \\ \text { In explaining how he arrived } \\ \text { at his answer the student } \\ \text { mentally computed the } \\ \text { solution. }\end{array} \\ \text { had to use them. } \\ \text { For the remaining items, } \\ \text { manipulatives were used } \\ \text { based on student preference } \\ \text { and/or need. } \\ \text { grid as he tried to solve the } \\ \text { item } \\ \text {-Paper and Pencil - "Can I } \\ \text { write the number model } \\ \text { down?" The student } \\ \text { requested permission to } \\ \text { write down the number } \\ \text { model in the process of } \\ \text { solving the item. } \\ \text {-Fingers - "6." The student } \\ \text { got the answer six by } \\ \text { counting on his fingers. }\end{array}\right\}$
$\left.\begin{array}{|l|l|l|}\hline & \begin{array}{l}\text { intervention. } \\ \\ \hline \text { The Zone of Proximal } \\ \text { Development (ZPD) was } \\ \text { the difference between the } \\ \text { two scores. }\end{array} & \begin{array}{l}\text { Versatility refers to the } \\ \text { collection of tools that a } \\ \text { student has available to } \\ \text { solve a problem and the } \\ \text { ability to use those tools. } \\ \text { An example would be a } \\ \text { student's ability to } \\ \text { represent and solve a } \\ \text { problem both symbolically } \\ \text { and graphically. }\end{array}\end{array} \begin{array}{l}\text { Philip’s answer for one of } \\ \text { the ways to make \$1.95 was: } \\ \text { Q Q Q Q - Q Q N N N D } \\ \text { D D. }\end{array}\right\}$

## Clinical Interview Categories

In coding the clinical interviews, I first concentrated on whether the items were correct or incorrect. I then used the rubric (Table 3) to evaluate the quality of the students' answers. If a student got a problem completely correct he was given a four, the highest score on the rubric. Incorrect answers were checked to determine the nature of the errors based on the definitions on the rubric. If the student almost got the problem
correct but had made minor errors, his score was three. If the student showed that he had been working towards a reasonable solution, but then made major mistakes or misinterpreted a part of the question, the score was two. When a student attempted to solve a problem but was soon unable to continue, the score was one. A student who was completely unable to solve an item or who handed in an answer that did not match the context of the question received a score of zero.

Table 3
Problem Solving Rubric

| 4 - Completion | An appropriate method is applied to yield a <br> valid solution. |
| :--- | :--- |
| 3 - Result | The problem is very nearly solved; minor <br> errors produce an invalid final solution. |
| 2 - Substance | Sufficient details demonstrate that the <br> student has proceeded toward a rational <br> solution, but major errors or <br> misinterpretations obstruct the correct <br> solution process. |
| 1 - Approach | The student approaches the problem with <br> meaningful work, indicating some <br> understanding of the problem, but an early <br> impasse is reached. |
| 0 - Non-commencement | The student is unable to begin the problem <br> or hands in work that is meaningless. |

Problem-solving rubric (Malone et al. 1980, p. 209.

After coding the clinical interview protocols (the sheets on which the students had actually written their answers), I coded the clinical interview transcripts (the typed records of their clinical interviews). First, I linked correct or incorrect answers with the rating scale levels that the teachers had assigned the students in teacher interview one.

The codes indicated whether or not the students had been rated secure, developing, or beginning in the concepts being examined in the items. These rating levels were linked to the previous categories of correct or incorrect. For example, one student could receive "correct-secure," while another could get "incorrect-secure." These codes were written directly on the transcripts.

Then I coded if the answers, correct or incorrect, were made without probes or with probes (Table 2). Two examples of this coding are "correct-no probe" and "incorrect w/probe." I added a code to indicate an item that was solved incorrectly due to directional miscues or misunderstandings - "incorrect-direction." For example, a direction to one of the items asked for a number model to show how the item should be solved. The direction did not ask the students to actually solve the subtraction problem in the item. If the students tried to do the actual subtraction, and did not give the number model as requested, this directional miscue would be the reason they got the item incorrect. Next, I indicated whether the student had solved the question using mental math, fingers, paper and pencil, manipulatives, or a combination.

The category that measured the Zone of Proximal Development (Vygotsky, 1978) was an attempt to measure the effectiveness of adult intervention in the context of the study. I assigned scores from the rubric (Table 3) to incorrect responses that the students had originally made before the clinical interviewer used probes. These scores ranged between zero and three on the rubric. If the student subsequently got the problem correct as a result of probes, I re-scored the item assigning it the top score of four. I labeled the difference between the original score and the score after probes as representing the Zone
of Proximal Development (Vygotsky, 1978). It showed what the students did independently and what they did with adult intervention.

Theme Identification \& Further Analysis
Themes, for example, misconceptions and valid propositions, resulted from focused analysis as I looked for repeated ideas and patterns in the data. In Table 4, I explain the interpretations of the various themes. In Table 5, I present actual examples of the themes from the students' transcripts. The themes were analyzed for additional insights about the students’ performances. Using information from the clinical interview transcripts I wrote these themes on individual sheets (Table 5). After identifying the themes, patterns that emerged from the data were the boys' use of time, the language of the questions on the protocol versus that of the NJPASS, the success or lack of success with probes, directional misunderstandings, and omissions while reading some questions by the clinical interviewer.

## Table 4

Themes and Explanations

| Themes | Explanations |
| :--- | :--- |
| Misconceptions | Faulty or incomplete knowledge that leads <br> to erroneous answer |
| Valid Proposition | Arguments, strategies or solutions that the <br> student suggests or uses that make sense in <br> the context of a particular problem |
| Unanticipated Responses | Answers given that were surprising <br> because the thinking they revealed was <br> unexpected |
| Patterns | Elements that occur repeatedly for more <br> than one student |
| Time | Group of patterns that highlight an <br> important idea |
| Cognitive Structures | How long did the participant take, and was <br> this time reflective of the students' <br> knowledge of the mathematical concept <br> being evaluated? |
| Manipulatives | Did the student use fingers, paper and <br> pencil, manipulatives or mental <br> computation? |
| Probes | What manipulatives did the student choose <br> and how significant was the choice for the <br> students’ thinking? | | Probes are the questions that the clinical |
| :--- |
| interviewer uses to help a student realize an |
| error or arrive at a correct answer. |, 

## Table 5

## Theme Comment Sheet

| Themes | Comments |
| :--- | :--- |
| Misconceptions | Jake started counting by fives on the clock <br> from the number 6. |
| Valid Propositions | Kevin - "Symmetrical means looks the <br> same on both sides." |
| Unanticipated Responses | Oliver - "That's the only way," in response <br> to the spinner problem. There was another <br> way. |
| Patterns | All the students misunderstood the <br> language of two particular items. This <br> misunderstanding became a pattern since it <br> was repetitive. |
| Motif | During the data analysis, challenges with <br> language were recurring and the frequency <br> of its recurrence led it to become a motif. |
| Time | Philip took the longest time on first section <br> of clinical interview. After 20 minutes he <br> was still on first page of clinical interview <br> protocol. |
| Cognitive Structures | Martin - used mental math; paper and <br> pencil. |
| Manipulatives | David - used coins but didn't use number <br> grid |
| Probes | Martin - sometimes probes didn't help him <br> to get correct answers e.g. coordinate grid <br> and bake sale items. |

In addition to identifying themes, while analyzing the students’ answers I
considered two other factors. These factors were versatility and adaptability (Sfard \&
Linchevski, 1994). Versatility refers to the collection of tools a student has available to
solve a problem and the ability to use those tools. An example of versatility would be a student's ability to represent and solve a problem both symbolically and graphically. For instance, in some of the items on the clinical interview protocol the children had to write number models. For another item they were required to write coordinates. Adaptability refers to the student's ability to select and use tools that are well matched to the problem. For example, depending on a student's mastery of basic facts, and the level of Bloom's Taxonomy (1956) at which he was operating, he could choose do his computation by using mental math or could opt to use a number grid.

The students’ answers were examined to determine the level of Bloom's Taxonomy (1956) at which they were functioning as they solved or attempted to solve the items. Their verbal explanations revealed their problem solving strategies and provided invaluable insight into their thinking. Their answers were also checked to explore the difference in effectiveness of probes among the students.

## NJPASS Categories

When the NJPASS scores were returned, the Individual Performance Profiles, (Appendix I, pp. 338-341), were examined to determine the students' levels of proficiency. These profiles showed how the students performed in each of the broad conceptual categories on the NJPASS. The mathematics section of the NJPASS had a total of 40 points. The score scales divided test takers into three categories - advanced, proficient, and basic - and the category depended on the actual score attained on the test. Out of a total possible raw score of 40 that could be earned on the mathematics section,
the advanced proficient level ranged from 34-40, the proficient level included scores in the range of 23-33, and the basic level encompassed scores that fell between 0-22.

Item by item analysis was made possible by using the Item Response Record that showed how the students had performed on each item in the conceptual categories on the NJPASS. I used this Item Response Record to compare the students’ performances on corresponding items on the clinical interview protocol.

Teacher Response Categories
Prior to the clinical interviews, the teachers were asked to give ratings of their student's abilities for specific mathematical concepts. They were also asked to give the student's cognitive strategies as well as the sources they had used to make their evaluations (Table 6).

## Table 6

Coding Sheet for Teacher Interview \#1
Participant’s Identifier: Rae for Martin

$\left.$| Conceptual <br> Categories | Rating Scale | Cognitive <br> Structures | Evaluation Sources |
| :--- | :--- | :--- | :--- |
| Part/Whole | Developing | Martin used <br> number grid | Workbook pages |
| Measurement | Secure | He did not use <br> mental math | Open-ended <br> questions |
| Money | Developing | He did not use <br> fingers | Pre-tests |
| Probability | Secure | He did not use <br> paper \& pencil | End of unit tests |
| Elapsed Time | Beginning |  | Math games |
| Telling Time | Developing |  | Teacher created <br> assessments |
| Patterns | Secure | Teacher/Student <br> conferences |  |
| Symmetry | Developing/Beginning |  | Teacher/Student <br> observations |
| Graphs, Charts, <br> Grids |  | Plastic <br> communicators on <br> which student write <br> answers |  |
| Place Value to 100s | Secure | Beginning of lesson <br> review |  |
| + \& Facts | + Developing |  |  |
| - Beginning |  |  |  |$\quad$| Checking homework |
| :--- |
| sheets | \right\rvert\, | Word Problems |
| :--- |
| Developing |

After viewing the clinical interviews the teachers were asked if their evaluations from the first interview were supported or negated by the performance that they had just viewed, and if there had been any surprises. I coded their responses on the corresponding coding sheet (Table 7).

## Table 7

Coding Sheet for Teacher Interview \#2
Participant Identifier: Dahlia for Jake

| Conceptual Categories | Evaluation Supported | Evaluation Negated | Surprised at Answer |
| :--- | :--- | :--- | :--- |
| Word Problems | Yes |  | No |
| Measurement | Yes |  | No |
| Money | Yes |  | No |
| Probability | Yes | No |  |
| Elapsed Time | Student did better |  | No |
| Telling Time | N/A |  |  |
| Patterns | Student did better |  | No |
| Symmetry | Yes | No |  |
| Graphs, Charts, <br> Grids | Yes | Yes |  |
| Place Value to 100s | Yes |  | No |
| + \& - Facts | Student did better |  | No |
| Part/Whole or <br> Fractions |  | Student did not do <br> as well | Yes |

The coding categories that were used in Table 7 are explained below. Table 8 gives examples of the teachers' responses as they described whether their evaluations were supported or negated. It gives the rating levels that were assigned to each conceptual category, and it lists these categories.

## Table 8

Teacher Coding Categories

Teachers' Responses to the Clinical Interviews

Chicago Math Rating Levels

The teachers' responses were examined to ascertain whether or not their evaluations were supported or negated, and if the students' responses were surprising.
"Supported, I think.
My prior evaluations were supported. Yes." This teacher felt that her evaluations had been substantiated by her student's performance on the clinical interview.

No teacher said that her evaluations had been negated.
"The coins. I know that they have a tough time with coins, but that was eye- opening." This teacher was surprised at her student's performance in this area.

These were the levels from the Chicago Math program that the teachers used to evaluate their students’ mathematic abilities in certain concepts.

These were the concepts on which the teachers evaluated their students using the Chicago Math rating scale.

These concepts were also the ones that were assessed on the clinical interview protocol.

Secure rating - the student has a thorough understanding of a concept or skill.

Developing rating - the student has some familiarity with the concept but has not totally mastered it.

Beginning rating - the student is just learning a concept.

The concepts were:
fractions - can children accurately describe the parts and wholes of shapes by correctly identifying numerators and denominators;
measurement - knowing the difference between cms and
inches;
money - identifying and computing with coins;
probability -using spinners;
elapsed time - measurement of another sort; patterns - being able to identify whether the numbers in a group were increasing or decreasing and by how much;
symmetry - identifying which figures on a sheet were symmetrical;
bar graphs - being able to answer questions that are based on a bar graph; grids knowing how to find locations on a coordinate grid by giving two numbers from each of the axes in the correct order,
charts - selecting the correct rides based on the information in the chart;
place value - correctly
identifying place values up to 100;
word problems - solving word problems that require varying solutions;
using basic + and - facts to solve the items.

In keeping with Patton (2002) I presented the data in descriptive yet neutral detail. I tried not to make any judgments about what occurred, but simply attempted to make what happened during the clinical and teacher interviews clear. My goal was to keep any personal bias in check.

## Credibility and Reliability

The following methodological choices were used to achieve credibility and reliability. The teachers were asked the same questions during the first interview using the survey. The questions on the clinical interview protocol were derived from a standardized test, and all the students were presented with the same items. The clinical interviewer was the same for all the students. Data was collected from three disparate sources including standardized test scores to obtain a broad yet accurate perspective. Data was accurately transcribed data from video recordings of the clinical and teacher interviews. The rubric was developed and validated by Malone et al, (1980). The goal was to have a scoring scale that showed the varying levels of progress utilized by students with disparate ability levels during problem solving. Malone et al, (1980) found that the rubric even when implemented with multiple scorers achieved a high level of consistency among them. The authors noted two types of consistency. First, there was within-marker consistency where a specific scorer marked certain answers with persistent uniformity. Second, there was between-marker consistency that entailed multiple scorers marking specially identified responses in similar ways. The rating scale that the teachers used to make their evaluations of the students' mathematic abilities was derived from another outside source, the Chicago Math Program. All other codes used in the study, for
example the idea units, were clearly identified and explained in tables. At two different times during the study, I got participant verification (Patton, 2002, p. 560) from the teachers in the study.

To achieve inter-rater reliability, I used two different raters. The first, a math educator, independently recoded a sample of the data using some of the coding categories I had devised for the study. The sample included four clinical interview protocols, each representing one of four students I randomly selected for the inter-reliability task. Fortytwo percent of the data was examined on the first attempt. After her independent coding was completed, I compared my coding with hers to determine the level of inter-rater reliability. The percentages of similarity for the coding done in the first attempt are outlined below:

## Rater One

| Code | Percentage of Similarity |
| :--- | :---: |
| Correct/No Probe | $74 \%$ |
| Correct/with Probe | $74 \%$ |
| Incorrect/No Probe | $49 \%$ |
| Incorrect/with Probe | $47 \%$ |
| Rubric | $61 \%$ |

These differences in scoring can be attributed to the fact that I had not provided the first rater with information that would have made her scoring more consistent with mine. Specifically, I had not given her desired answers or possible student solution methods. In addition, the numbering scheme for the sub-questions was not clearly identified. After I had shared the levels of similarity in coding with the first rater along
with my analysis about the source of the differences, she indicated that if she had that insight prior to coding the protocols, our scores would have been better aligned.

After this experience I used a second rater, a lecturer at a Mid-western university, to score the categories that had not been evaluated by the first rater. Fifty-eight percent of the data was examined on this second attempt. First, the second rater and I discussed the categories and what they meant. In this way we were able to use the same understandings of the coding categories as we scored the protocols. We tested this correlation by coding a protocol that had not been in the sample given to the first rater. After discussing our scoring we identified and resolved minor discrepancies between our scores. Then the second rater independently coded two of the protocols that had already been evaluated by the first rater. Once again the second rater and I discussed our scores and determined that there was a very high degree of similarity among them. The percentages of similarity for the codes with the second rater are presented as follows:

Rater Two

| Code | Percentage of Similarity |
| :--- | :---: |
| Correct/Rating Scale | $97 \%$ |
| Incorrect/Rating Scale | $94 \%$ |
| Incorrect/Directions | $89 \%$ |
| Cognitive Strategy | $94 \%$ |
| Zone of Proximal Development | $100 \%$ |
| Versatility | $93 \%$ |
| Adaptability | $95 \%$ |

The differences in the percentages of similarity between the two raters illustrates the advantages of doing a practice run with the second rater prior to her conducting independent coding as there is a much higher rate of similarity.

In addition to using two raters, after transcribing both sets of teacher interviews, I emailed them to the teacher participants to verify that the transcripts accurately reflected what had transpired during the interviews. Findings of the study were also shared with these participants and their reactions were included in the discussion section. Consistent with Patton (2002, p. 560), such triangulation of qualitative data sources was achieved when the adult participants were able to respond to the data of the study. The study's findings are presented in the next chapter.

## CHAPTER FOUR

## Findings

The goals of this study were twofold. First, the study aimed to discover if a complementary assessment would allow a group of African American boys to demonstrate more mathematical competency than on the standardized test. Second, the study examined whether the information the teachers gained about their students' mathematical abilities altered their opinions of the students’ proficiency in mathematics. The answers to these two paths of inquiry are presented in the pages of this chapter. The findings are presented from the constructivist perspective. As the literature review revealed, the study was framed around the theoretical foundation laid by (Freud, 1938; Piaget, 1969, 1978; and Vygotsky, 1978). Their emphasis on fathoming children's mental processes as they solve problems is reflected in the results in this chapter.

The findings of the study are presented so as to illustrate how the students constructed their answers on both the clinical interviews and the NJPASS. Since there was not a one to one correlation among the items on both assessments, included in this introduction is a brief explanation of how I arranged the comparison of the students' scores. The first section identifies the particular assessment on which each student performed best; presents the teachers' ratings in the conceptual categories; shows indepth examples from the clinical interviews; provides a summary of the students’ NJPASS scores; shares the teachers’ comments about their students' clinical interview performance; and summarizes the changes in the teachers' attitudes in relation to the
information they had received from watching the clinical interviews. In the second section, common trends and patterns for several of the students are presented. The third section addresses individual or unique discoveries about the boys that were pertinent to the results. The chapter ends with a discussion and interpretation of the findings as they related to both assessments.

## Explanation of Item Comparison

The clinical interview protocol had 14 items with a total of 31 possible points and was administered in one session. The NJPASS assessment had 30 items for a total of 30 possible points and was administered over a period of two sessions. In order to accurately compare the performances of the students on these assessments, the items on the NJPASS that best corresponded with a particular item on the clinical interview were selected. Occasionally two items on the NJPASS were the best match for one item on the clinical interview. Ultimately, I compared 20 NJPASS items worth a total of 30 points with the items on the clinical interview. A description of the selected items as well as of the students' scores on the clinical interview items and the corresponding NJPASS items is available (Appendix H, pp. 331-337).

The results of the students' performances on the clinical interview items and on the NJPASS corresponding items provided the answers to the first research question. First, the clinical interview protocols were scored. The responses for each item by the individual students were noted. The NJPASS Item Response Records gave the answers for each item on that assessment. The students’ scores on the corresponding NJPASS items were then compared to the boys' scores on the clinical interviews.

The scores from the clinical interview and the analogous NJPASS items reveal that six of the students - Kevin, Sam, Oliver, David, Jake, and Martin - performed better on the comparable NJPASS test items than on the complementary assessment or the clinical interview items. The other student - Philip - performed better on the clinical interview items than on the standardized test items. It should be emphasized that in comparing the scores of the items from both assessments, I used specific items from the NJPASS, and did not use all the items in the math segment. While six of the boys earned better scores for the corresponding standardized test items, there was a 50 percent difference between the highest and lowest scores for the NJPASS items. There was a 35 percent span between the highest and lowest scores for the clinical interview. The scores are outlined in the table below.

Table 9
Students' Scores

| STUDENT | CLINICAL | NJPASS SCORE ON | DIFFERENCES | NJPASS SCORES |
| :---: | :---: | :---: | :---: | :---: |
|  | INTERVIEW | CORRESPONDING | IN | FOR ALL ITEMS |
|  | SCORE | ITEMS | PERCENTAGES | ON ACTUAL |
|  |  |  | BETWEEN THE | ASSESSMENT |
|  |  |  | TWO SCORES |  |
| KEVIN | 22/31-71\% | 24/30-80\% | +9\% | 33/40-83\% - |
|  |  |  |  | Proficient |
| SAM | 23/31-74\% | 24/30-80\% | +6\% | 29/40-73\% - |
|  |  |  |  | Proficient |
| OLIVER | 20/31-65\% | 29/30-97\% | +32\% | 38/40-95\% - |
|  |  |  |  | Advanced |
| DAVID | 15/31-48\% | 20/30-60\% | +12\% | 28/40-70\% - |
|  |  |  |  | Proficient |
| JAKE | 13/31-42\% | 14/30-47\% | +5\% | 18/40-45\% - |
|  |  |  |  | Basic |
| MARTIN | 18/31-58\% | 18/30-60\% | +2\% | 22/40-55\% - |
|  |  |  |  | Basic |
| PHILIP | 24/31-77\% | 19/30-63\% | -14\% | 24/40-60\% - |
|  |  |  |  | Proficient |

Data from the first teacher interviews, the clinical interviews, the NJPASS scores, and post teacher interview comments about each student's clinical interview performance are presented in the following sections. The data in this first section are organized to paint a comprehensive picture of each individual student. The information from the first teacher interviews reflects how the teachers rated the students before the clinical interviews were administered. In addition to their evaluations, some teachers shared pertinent information about the students' work habits and home environments. The clinical interview segment indicates whether or not the students were successful on the items. I specify whether their responses matched the teachers' ratings as the accuracy of the teachers' evaluations affects the results of the second research question. Detailed
descriptions of the students' individual responses are reserved for those items that illustrate a pertinent finding for the current study. Next, the NJPASS scores on the items that corresponded to the clinical interview items are presented. In the post teacher interviews that follow, the teachers' observations as they viewed the DVDs with the clinical interviews are shared. These teacher comments specifically address the students’ responses on the clinical interviews and the NJPASS. They do not represent answers to the second research question as these are provided at the end of the first section.

## Summaries of Performance

## Kevin

## First teacher interview.

Kevin attended B School and was seven years and eight months old when the study was conducted. His teacher, Liane, had taught for a total of 35 years, 15 of which had been spent teaching second grade. She had a Masters degree in teaching. In her first teacher interview, Liane used the Everyday Math rating scale (from the University of Chicago math program) to rank Kevin on the conceptual categories that would be explored during the clinical interview. Kevin was rated at the secure level in all but two of the categories, and in those he was classified as developing. With reference to his preferred cognitive strategies, Liane indicated that Kevin used much mental math and paper and pencil. She indicated that he only used manipulatives if the lessons or problems required their usage.

During that first interview, Liane also provided more information about Kevin's mathematical strengths and weaknesses and the factors that potentially affected those areas. She felt that based on his work in all the academic subjects, he was most proficient
in mathematics. Liane also thought that he was better at rote learning than at solving open-ended items. This was an important observation given the results of the current study where Kevin performed better on the NJPASS with its primarily multiple choice format.

The clinical interviewer asked Liane to account for the reasons why she had evaluated Kevin so highly on the conceptual categories given his propensity for rote learning. The reason for her inquiry stems from the nature of the Chicago Math Program with its emphasis on open-ended problem solving. Liane credited his success to two disparate factors - parental support and a facility with memorization.

Liane: So I understand what you're saying, but, uh, a lot of the ... what helps him is knowing all his addition facts, his subtraction facts, he knows his multiplication facts, he knows his ... because that's what his dad has worked with him at home, therefore a lot of that transfers into his reasoning. But he is much better at rote ... he’s just a whiz at memorizing things (TI1, lines 237-243).

Liane's positive opinions and expectations of Kevin were based on concrete examples of his work such as workbook pages and tests, as well as teacher observations and conferences. The following exchange occurred after she had rated Kevin highly in most categories and succinctly captures her view of Kevin's academic ability.

Principal Investigator: Wow. Ok. Very good.
Liane: Hmm, hmm...good student (TI1, lines 135-137).
This opinion of Kevin's mathematical ability was not only substantiated by Liane's documentation, but it was also supported by his work during his clinical interview. A detailed description of this performance follows in the next section.

## Kevin

## Clinical interview.

As one watched the recording of Kevin's clinical interview, his concentration and focus throughout the assessment were clearly evident. Kevin answered both the first item - a word problem that explored the students’ knowledge of basic subtraction facts, and the second item that evaluated measurement - accurately and without the need for probes. His responses in these areas corroborated Liane's secure ratings for these conceptual categories. However, Kevin’s response to the third item necessitated adult intervention on the part of the clinical interviewer. It is illustrative of an instance during the clinical interviews when probes were instrumental in allowing a student to correct a previously incorrect answer.

Clinical Interviewer: If your tee ball practice began at 3:30 and lasted for 45 minutes, what time would it end?
(Kevin uses clock manipulative to figure out the answer).
Kevin: 10:03
Clinical Interviewer: Umm ... do you ... how did you come up with 10:03? Kevin: I counted, I counted, I counted 5, 10, 15, 20, 25 (stopped talking and computed mentally)
Clinical Interviewer: There is a problem with the hands on this clock. Let's put it back to 3 , let's put it back to $3: 35,3: 30$. Now, just move the minute hand. I'll hold the hour hand still because it's sliding with the hand that you've got ... Kevin: 4:15
Clinical Interviewer: 4:15. Ok. Could you write that on your paper? And, ah, what did you count to figure that out? How did you count to get that answer?
Kevin: Fives
Clinical Interviewer: You counted by fives?
Kevin: (Nods yes) (CI, lines 41-68).
This interaction reveals how important it is for the clinical interviewer to be attentive to the interviewee so that the correct type of probe can be provided if necessary. The clinical interviewer had been closely observing Kevin as he tried to manipulate the
hands of the clock and had noticed that he had not been able to move the hands accurately. By offering to hold the hour hand, and by having Kevin move the minute hand, she gave Kevin the opportunity to show that he did know how to solve the elapsed time item. For additional confirmation, after he had successfully found the answer, she asked him to explain how he had arrived at his solution. Kevin was able to say that he counted by fives. Liane had noted in the first interview that when solving problems of elapsed time, Kevin sometimes experienced challenges with five-minute intervals. "... when it's five minute intervals like 2:55, uh, 8:50, he has a little difficulty" (TI1, lines 8788). This was not the case this time, but his difficulty with manipulating the hands of the clock substantiated Liane’s rating of developing in this conceptual category. For the fourth item, Kevin responded in a manner that definitively supported the secure rating that he had received for the conceptual category of patterns.

By discerning and completing the numerical pattern almost immediately, Kevin displayed his mastery of rote processes and mental math. He was also able to explain what he had noticed about the pattern. The fifth item was a task that required the students to use coins to represent a specific sum in two different ways. Kevin computed this problem mentally and accurately although it took him a few minutes. Again, this question showcased Kevin's proficiency with mental math, and his skill in this area was enhanced by his secure knowledge of his basic facts. Kevin’s facility with rote memorization was not as important in the next conceptual category that was explored.

Probability was the second of the two categories in which Kevin had been rated at being at a developing level. He had no problem completing the task, but he was reticent
about explaining how he had decided how to color the spinners differently while using the same colors. Kevin's initial silence reveals one of the potential pitfalls of the clinical interview method. The method reveals children's thinking, but this revelation is only possible if the interviewees are willing (and able) to explain their thinking.

Kevin: (Works on coloring first spinner according to the directions). Clinical Interviewer: Ok. Can you do a second spinner that would have equal amounts of red, yellow, blue, and green, but doesn't look like this one?
Kevin: (Turns second spinner around for a couple of seconds, then starts coloring it in).
Clinical Interviewer: Very nice. And how did you decide to do that?
Kevin: (Remains silent)
Clinical Interviewer: How did you make them different? What did you do?
Kevin: (Turns spinners but says nothing)
Clinical Interviewer: You did make them different. How did you ... what did you do?
Kevin: The blue was over here and the yellow was over here (said as he points to the spinners).
Clinical Interviewer: Ok. So you switched where you positioned the colors?
Kevin: Uh, (shook his head saying no). I put the red over here, the yellow over here, the yellow here, and the green over here.
Clinical Interviewer: Ok. Thank you (CI, lines 312-336).
The clinical interviewer's persistence resulted in Kevin's eventual response, and her perseverance exemplifies the critical importance of the interviewer's ability to adjust to the interviewee's behavior during the clinical interview. Despite his reluctance to reply to questions about this item, Kevin showed no hesitation in answering the next set of inquiries from the clinical interviewer.

In completing the next item on the bar graph, Kevin correctly answered the subquestions that he was asked. There were four sub-questions, but the clinical interviewer inadvertently omitted the third sub-question. He needed no probes and his responses again validated Liane’s rating of secure for this category. Additionally, Kevin’s acuity
with mental math was further displayed as he added several numbers totaling 29 without using pencil and paper or the number grid.

Clinical Interviewer: How many books were read in all?
Kevin: (Computes mentally. Then writes 29).
Clinical Interviewer: And how did you figure that out?
Kevin: ... because $3+6$ is 9
Clinical Interviewer: So you added all the different numbers?
Kevin: (Nods yes)
Clinical Interviewer: Good. Umm ...
Kevin: $9+4$ is $13+7$ is $20+5$ is $25+4$ is 24, is 29 (CI, lines $156-170$ ).
Easily transitioning from bar graphs to symmetry, Kevin continued to reveal his mathematical competence. He easily completed both parts of this item, and thus measured up to Liane's evaluation. Kevin was less successful on item nine that explored his knowledge of coordinate grids. He knew that two numbers were necessary for identifying the coordinates, but unfortunately, put them in the incorrect order. This item provides an example of where a probe was unsuccessful in helping the student realize that he had made an error.

Clinical Interviewer: If I told you that the rectangle was R on the grid, what coordinates could you put to show me where the rectangle was?
Kevin: $4 \ldots 4$, 1
Clinical Interviewer: 4, 1. And how did you decide that?
Kevin: Because 4 is right here and 1 at the bottom
Clinical Interviewer: And, and is that the order that you (should) put them in? Kevin: (Nods yes) CI, lines 238-252).

For this item, Kevin was unable to use the probe to realize that he had erred in putting the coordinates in the correct order. As a result, he did not meet Liane's rating of secure in this category in contrast with his performance on the next item.

In the word problem that involved computation with money, Kevin easily made the correct choices by once again utilizing mental math. Item eleven was another word
problem, but due to a directional miscue, Kevin was not able to solve the problem accurately. He did not register that the directions required him to write a number sentence that showed how to solve the problem. Instead, Kevin tried to solve the item by computing mentally. However, this was a multi-digit subtraction problem that required renaming, so his efforts were not only unnecessary, they were also unsuccessful. Not understanding all the nuances of the directions again proved to be Kevin's downfall on the next item, number twelve that featured a chart. He chose tickets that totaled the requisite amount, but didn't realize that he had to use the greatest number of tickets that was possible. The clinical interviewer recognized that Kevin had focused solely on choosing tickets that totaled nine and not on getting as many as possible. However, based on her observations of his body language, she decided against using probes to get Kevin to reconsider his answer. Consequently, for items eleven and twelve, Kevin's performance was not consistent with Liane's evaluation in these categories. Nonetheless, the next item allowed Kevin to demonstrate his generally good knowledge of mathematical concepts.

Item thirteen examined Kevin's understanding of place value to the hundreds place, and he had no difficulty whatsoever in answering the two parts of the question. He clearly met Liane's rating of secure in this category. The last item examined his understanding of fractions or parts and wholes. Kevin quickly and correctly answered the first part of the item. His initial answer to the second part of the item reflected a conflict between what the students were accustomed to answering in their mathematics curriculum as contrasted with how the question was phrased on the clinical interview. As

I subsequently discovered in Liane's post interview, the students were accustomed to finding the fraction of the figure that was shaded. Kevin initially interpreted the question that way and gave the fraction for which that would have been an accurate answer. After a probe, he was easily able to give the fraction of the un-shaded portion of the figure.

Clinical Interviewer: ... ah... did you write the fraction to show how many were shaded or did you write the fraction to show how many were un-shaded? Kevin: Shaded
Principal Investigator: So could you write for us how many were un-shaded now?
You did shaded, no... that's fine. That's great.
Clinical Interviewer: Could you write ...
Principal Investigator: ...how many were un-shaded now. Write the fraction ... Clinical Interviewer: You could just write it down here (CI, lines 385-398).

Kevin's clinical interview performance depicted a student with very competent math skills who appeared confident in his mathematical abilities. This supports Liane's opinions that were shared in the first teacher interview.

Liane: Uh, he's more ... much more confident in math than in other areas Principal Investigator: Ok
Liane: So, other areas that he needs to work on. And math is his strength so he, he volunteers more often. He feels good which means he is more confident (TI1, lines 208-213).

Liane's evaluations were verified by his accurate answers to most of the items and by his masterful command of mental computation. The coordinate grid problem was the only item that Kevin got incorrect due to incomplete conceptual knowledge. For the other items on which he was unsuccessful, the errors can be attributed to a misunderstanding of the language of the directions. Despite these lapses, Kevin's performance on the clinical interview revealed that he had a very good grasp of most of the conceptual categories that were evaluated. Even so, as good as his work was on the
clinical interview, his performance on the corresponding NJPASS items proved to be better. The details of his answers on the standardized assessment as well as selected comparisons with clinical interview items are provided in the following section.

Kevin
NJPASS results.
Kevin's NJPASS scores both on the specific items that matched the clinical interview items and on the actual assessment were quite good. As was seen in Table 9, Kevin got 24 of the 30 items that corresponded with the clinical interview items correct. He got 33 of the 40 items on the actual assessment correct. The latter score earned him the level of Proficient on the complete NJPASS assessment. In addition to these general scores, a second set of scores provided more details about the types of questions in which Kevin experienced success and failure. Unlike the clinical interview that provided insight into the thinking that led to the students' answers, the scores on the item response record merely indicated which items the student had gotten incorrect or correct. They did not present the teacher or analyst with the opportunity to discern the reasons behind the students' answer choices.

## NJPASS Item Response Record

The NJPASS Item Response Record provided details about Kevin’s performance on each item on that assessment. The scores were divided into the following categories: Number Sense, Operations, and Properties; Measurement; Spatial Sense and Geometry; Data Analysis, Probability, and Discrete Mathematics; and Patterns and Algebra. Each
category concluded with an open-ended item. In the sections below I indicate how Kevin performed on the relevant NJPASS items.

Number sense, operations, and properties.
In the category of number sense, operations, and properties, Kevin got two items in that section incorrect. The first of these items was one of two that corresponded with item fourteen on the clinical interview. It showed three figures and asked the children to identify the one that was three-quarters shaded. The second item that Kevin got incorrect in this section evaluated children's ability to estimate using money. The item stated that someone had a certain amount of money (under a dollar), and had spent some of it. It then asked the students to estimate the amount of money that remained. They were given three choices. In addition, the open-ended item in this category of the NJPASS also explored the students' abilities to determine money values. Kevin earned the maximum three points for this item. An analysis of Kevin's responses to the various items dealing with money on both assessments suggests that Kevin's difficulty with the first incorrect item on the NJPASS was not with determining money values, but was with its requirement to estimate. This supposition is predicated on the fact that Kevin had successfully completed all the other items in which money had been the conceptual category. This was the case on both the NJPASS and the clinical interview regardless of whether the items were multiple choice or open-ended. Kevin's mixed success in this category was continued in his performance on the measurement section of the NJPASS.

## Measurement.

In this category, Kevin got the clock problem, item nine, incorrect. On the NJPASS assessment, the pictures of three clocks showed three different times from which the student had to identify the correct time. The importance of these three times is that they only varied by five-minute intervals. Kevin's difficulty with time in this item is supportive of Liane's evaluation of his facility with clocks in the first teacher interview.

Liane: There are two different ... the report card this time says five-minute intervals. Kevin knows his half hours, his fifteen ... quarter hours, fifteen minutes after the hour but when it's five minute intervals like $2: 55$, uh, $8: 50$, he has a little difficulty, so I put down ... I put developing (TI1, lines 86-95).

Despite getting item nine incorrect on the NJPASS, three problems later, Kevin accurately solved the open-ended item for this category that also evaluated elapsed time.

He earned the maximum score for this item. Therefore, Kevin’s challenges and successes with the items that explored elapsed time proved that he was definitely working at a developing level in this area. His responses to the next category of spatial sense and geometry showed that he was much more accomplished and secure in those concepts.

Spatial sense and geometry.
In this third category on the NJPASS Item Response Record, Kevin got all the items correct. I will only discuss those that had comparable clinical interview items. The first item was one of the two that corresponded with number fourteen on the clinical interview. It was a multiple choice fraction item that requested that the students identify how many squares in a figure were shaded. Item seventeen was the NJPASS item that corresponded to the coordinate grid problem on the clinical interview. On this NJPASS item, there was a map of rides at a carnival. There were numbers on either axis and the
children had to identify the location of only one of the rides from three multiple choices that used across and up or down terminology along with the pertinent numbers from both axes. The NJPASS item on symmetry was similar to the one on the clinical interview in almost every aspect. Neither item presented any complications for Kevin. This was not the case for the next category where his answers to some of the items were inaccurate.

Data analysis, probability, and discrete mathematics.
In this category, Kevin got two items incorrect - the bar graph item and the openended chart item. The bar graph item on the NJPASS was a single multiple-choice item. Surprisingly, given the fact that it was multiple-choice, Kevin was unsuccessful on this item. I did not have enough information from the item response record to ascertain if Kevin had experienced the same difficulty with the problem as he had on the clinical interview. Specifically, he had not followed the directions and therefore had not found as many different rides as both items requested. Complying with instructions was not as problematic for Kevin in the following section that explored his knowledge of patterns and algebra.

## Patterns and algebra.

In this final category, Kevin was successful on all the items. Item two corresponded to item one on the clinical interview and explored basic addition and subtraction facts. Item twenty-seven was the second of two items that corresponded to the elapsed time item on the clinical interview. Item twenty- seven was also open-ended, and once again Kevin computed the answer accurately.

To conclude, Kevin's performance on the NJPASS items again revealed a student who overall had a good understanding of the mathematical concepts that were evaluated in the various categories. Kevin experienced the most success in the following categories: spatial sense and geometry and patterns and algebra. He got all the necessary items correct in these sections. In the other sections - number sense, operations, and properties; measurement; data analysis, probability, and discrete mathematics, Kevin made mistakes on a few items. The item response record did not give details about the reasons Kevin got them incorrect. Estimating with money, elapsed time with five-minute intervals, and understanding directions proved to be the issues that were most problematic for Kevin. Despite Liane's statement that he did better on items that required rote memorization rather than on open-ended items, of the five open-ended items on the NJPASS assessment, Kevin got the maximum points on four of them. In some instances he also did better on an open-ended clinical interview item than on a multiple-choice NJPASS item. This performance gives credence to Liane's explanation that Kevin's skill with rote work gave him a measure of automaticity that allowed him to successfully answer open-ended items. Also, since the district's mathematics curriculum was based on the open-ended format of the Everyday Math program, this exposure could have influenced Kevin's aptitude in this sphere. Overall, Kevin's responses to the NJPASS items were very similar to his work during the clinical interview. He performed well on both assessments. Liane voiced this opinion after she had viewed Kevin's clinical interview. In the post interview, she paid close attention as he proceeded through the items on that assessment.

Kevin
Post teacher interview.
During the second or post teacher interview, Liane observed Kevin’s performance intently. Her comments during this interview included reflections not only about Kevin's performance, but also included curricular comparisons that were based on the differences in presentation of the items in the Chicago Math program, on the clinical interview, and on the NJPASS. Soon after the beginning of the post interview, the clinical interviewer mentioned that Kevin had taken a lot of time to work through the problems on the protocol. Liane's response indicated that some of that delay could conceivably be attributed to Kevin's occasional difficulty in interpreting language.

Liane: Our friend (I had not informed Liane of Kevin's pseudonym, and she used this term to maintain confidentiality) normally does take a lot of time thinking, processing. Sometimes I think ..., sometimes I think interpreting the language, sometimes I think also (TI2, lines 34-36).

This statement about Kevin's periodic difficulty with interpreting language is important given his misunderstanding of directions in two of the items on the clinical interview. Liane went on to add that Kevin sometimes found directions confusing especially when they were given in a whole group context.

Liane: You know I'm finding too he's working better one on one. He's a bright student, but that's a good point to make. You know, in a larger class, sometimes he gets lost in the directions, in the whole ...competitiveness, you know ... (TI2, lines 289-291).

So twice in the post teacher interview, Liane mentioned Kevin's difficulty with interpreting language and directions. This issue with language was also applicable to the
multi-digit subtraction item. Kevin did not register what the directions required and as a result did not arrive at the correct answer.

As the interview progressed, Liane made comments that revealed similarities in the students' responses to the corresponding items on both the clinical interview and the NJPASS. She made the following observations about the multi-digit problem. Liane stated that the multi-digit with subtraction item had been difficult for the students on the NJPASS as well. " This was, there was something similar to that (on the NJPASS) and that was difficult" (TI2, lines 492-493). In refutation of the perceived challenge, Kevin got that problem correct on NJPASS. As previously mentioned this was not the case on the clinical interview.

However, despite the fact that Kevin did not arrive at the correct answer for this item on the clinical interview, his approach to solving the question revealed that he was using an interesting cognitive attribute. The item had three digit numbers and required borrowing and renaming, yet Kevin attempted to solve it mentally even though he had access to the number grid and paper and pencil.

Liane: When he's looking up it's almost like he's visualizing ...
Principal Investigator: Yes. I think he is. Yeah, that's good.
Liane: He's trying it on the grid, but mentally he's trying to figure out... (TI2, lines 607-611).

This observation suggests that Kevin's prowess with mental computation was not solely reliant upon his mastery of basic facts but also depended upon his ability to visualize tools such as the number grid. The following describes Kevin's actions after the item had been read to him.

Kevin: (Thinking, doing work mentally)
(Clinical Interviewer gives him a number grid after a while. He uses it but is still computing mentally. Writes something down, then goes back to the grid. Goes back to computing mentally. Erases what he wrote previously. Writes an answer after taking his time to work on the problem).
Clinical Interviewer: How did you do ... decide on that answer?
Kevin: I counted 10, 20, 30, 31, 32, 33, 34, 35, 36, 37, 30... 10, 20, 30, 31, 32, $33,34,35,36,37,38,39 \ldots$ (he stops and erases and changes his answer to 39) (CI, lines 411-424).

From this description, it would appear that although the number grid ended at one hundred, Kevin seemed to have been able to visualize what it would look like if the numbers were increased. In this way, although, the item did not require an actual solution, Kevin's answer, although incorrect, was very close to what the answer would have been. Considering that this item had three digit numbers and also required subtraction with renaming, this was excellent mental computation on Kevin's part.

The multi-digit word problem had multiple answer choices on the NJPASS and none on the clinical interview. Antithetically, the chart items on both assessments were identical except for the names of the rides and the number of tickets that were required. This similarity in presentation could have contributed to the analogous nature of Kevin's responses to the corresponding chart items. He got the chart items wrong on both assessments.

After reading Kevin's response to the chart item on the NJPASS assessment as
she walked around the classroom, Liane had realized that he had gotten it incorrect.
Liane: He did a very different interpretation on the NJPASS for this.
Principal Investigator: He didn't get it right?
Liane: He interpreted it completely differently, so it's interesting to see how he will do on something similar. I remember looking, and Oh! Where did he get that from? There weren't enough, there weren't enough tickets, so he interpreted it as adding up all, there was one with rides, and how many tickets the rides were, and which rides could you use if you could only use eleven tickets?

Principal Investigator: That's what I based this on...
Liane: And he said that there weren't enough tickets to go on all the rides (TI2, lines 495-509).

Liane's explanation of how Kevin had misinterpreted the question on the NJPASS illustrated his difficulty in comprehending the language of some directions. She had mentioned this challenge in the first teacher interview. He experienced a similar problem with the instructions of the chart item on the clinical interview. There, he did not say there were an insufficient number of tickets. Instead, Kevin focused on the total number of nine and quickly found three rides that made that amount.

Clinical Interviewer: He has exactly nine tickets. How can he use them so that he can go on as many rides as possible?
Kevin: (Looking into space and thinking)
Clinical Interviewer: You could just circle the ones that he's going to use.
Kevin: (Circles tickets that total nine, but that were not as many rides as possible) (CI, lines 356-362).

He ignored the words that indicated that he had to choose the most rides that could be obtained using only nine tickets. Unfortunately, as a result of his focus on choosing rides that totaled nine, Kevin's facility with mental math may have added to his misinterpretation of the directions. He quickly identified the rides that totaled nine and in doing so overlooked the real focus of the item.

After this, the discussion during the post interview shifted. After reflecting on Kevin's responses, Liane's made observations pertaining to the differences between the presentation of the items on the clinical interview, the NJPASS, and the Everyday Math curriculum. For example, Liane noted that on the Chicago Math curriculum the students were more accustomed to finding the shaded portion of a figure rather than of the unshaded part. "Usually when they did fractions, it was how much was shaded was the
fraction" (TI2, line 565). On the clinical interview, the children were required to find the un-shaded portion of the figure. In addition to the fraction item, the probability item also had a different element.

Liane found the way that probability had been explored in the clinical interview interesting. It varied from the approach in the Everyday Math curriculum and the NJPASS assessment where the presentations were similar to each other.

Principal Investigator: On the test, do they have to make their own spinners? Liane: No. The spinners are there with and arrow pointed and it will say what is the probability of the spinner landing on a ... red ...
Principal Investigator: Right
Liane: ... or what's the probability of the spinner landing on the number five, and the spinner did not have any number five on it ...
Principal Investigator: Right
Liane: So, your choices were, you know ...
Prinicipal Investigator: Yeah, right ...
Liane: Always, never ...
Principal Investigator: That's what they had, well ...
Liane: So this is interesting to see a different way ... presented in a different way. (TI2, lines 351-371).

She also felt that this aspect of the probability item on the clinical interview made it more challenging than for the Everyday Math curriculum and on the NJPASS assessment.

Liane: Oh, it’s four colors, it's four colors. So eight, and two of each.
Principal Investigator: So first he did two of each, then you could do one, one, one, one.
Liane: That's a difficult concept ... (TI2, lines 450-454)
Liane's observations about differences in presentation were limited to these two concepts in her post interview about Kevin. After this discussion, Liane ended the interview by stating that Kevin's responses had been as she had anticipated. She noted that his preferred method of computation had indeed been mental math. That pleased her as she felt it showed a greater level of proficiency. "Which is good, great that he could
do that ... that's a higher level" (TI2, line 678). She reiterated that he had attained to the task better on a one to one basis than he would in a larger group.

Liane: He attained to the task more efficiently in a one to one setting than in a large group, I think he always feels like he's pressured you know to get things done, where as you saw it took him quite a while, but he wasn't as pressured to finish it as when he's in the classroom, so I ...there weren't surprises, I think he did really well, but I think he did better probably on a one to one than in a large group setting (TI2, lines 693-697).

This summation is an accurate portrayal of Kevin's performance not only on the clinical interview, but also on the NJPASS items. He was efficient and did very well. Additionally, his performance helped to illustrate the importance of the clinical interviewer's ability to solicit answers from interviewees, showed instances when probes worked and when they didn't, and raised the importance of language interpretation to the findings of the study. As it turned out, language - in directions and as a means of explaining one's thinking - proved to be an important factor as well for the next student, Sam.

## Sam

First teacher interview.
Sam was seven years and eleven months old at the time of the study. Like Kevin, he attended B School, and Liane was also his teacher. She gave him ratings that varied between secure and developing. The four conceptual categories in which Sam was rated as developing were measurement, probability, elapsed time and word problems. Liane made the following statement about Sam’s ability with measurement. "Measurement, uh, S is developing. He is coming along nicely. But he is developing..." (TI1, lines 300301). Like Kevin, Sam got a rating of developing for the probability concept. Liane
explained that the topic was not covered much in the Everyday Math program and she had only been reviewing it for a week prior to the testing. For the elapsed time problem, Liane felt that like Kevin, Sam had issues with five-minute intervals. "Developing - he is fine with the basic times but the five minute intervals..." (TI1, lines 317-318).

In addition to evaluating Sam's ability in the conceptual categories, Liane gave her views about his preferred computational strategies. She stated that Sam benefited from using mental math, fingers, paper and pencil, and manipulatives. To explain his thinking Liane felt that Sam preferred a combination of oral and written explanations, and diagrams or pictures. "I think combination because he has a little bit of a problem following directions so ... all methods benefit" (TI1, lines 361-366). As she had done with Kevin, Liane supported her evaluations with evidence from sources such as workbook pages, open-ended questions, pretests, end-of-unit tests, math games, teachercreated assessments, teacher-student conferences, and teacher-student observations.

When asked if she had anything else to add about Sam's mathematical ability, Liane indicated that Sam had much mathematical ability. However, she felt that his challenges with following directions often derailed his efforts.

Liane: Umm, basically that a lot of Sam's math ability is there, but I don't feel... he does have a difficulty following directions as I said before ...
Principal Investigator: Ok, has the ability but has trouble following directions. Liane: Hmm, hmm (yes) and once ... sometimes you have to go over it again with him ... Let's go over ... and go over to him sometimes and say let's go over what we just talked about and he'll be like ...ooh!! Ok!
Principal Investigator: So he needs repetition.
Liane: Oral, written, right ... repetition ...
Principal Investigator: ... of directions?
Liane: Yes
Principal Investigator: Ok.
Liane: Sometimes individually

Principal Investigator: Ok. Does he benefit from like group work, or small group work, or refreshing?
Liane: Hmm, hmm (yes). On occasion he does go for small group work in the back of the room. I have in class support during that time. Even though there are other children that are specified, if he is also having trouble with a skill, um, he is moved to the back.
Principal Investigator: Right. Ok. He benefits from that?
Liane: Hmm, hmm (yes) (TI1, lines 382-413).
Liane's interview ended with her agreeing with the clinical interviewer that on a scale of novice to expert, Sam was somewhere in the middle. "Hmm, hmm (yes). Right in the middle, and has some secure skills, nice secure skills also" (TI1, lines 429-430). Liane’s realistic yet positive opinion of Sam's mathematical skills was evident during this first teacher interview. She had based her ratings on a substantial number of resources, and her evaluation of Sam's mathematical ability was also verified by his overall performance during the clinical interview. In the next section, Sam's responses to the items on the clinical interview protocol are presented and discussed.

## Sam

## Clinical interview.

Sam's demeanor and body language was noticeably different from Kevin's from the very beginning of his clinical interview. He had his head propped up on his hands prior to the clinical interviewer's introductory explanation and before he was even asked the first question. Additionally, in response to her query as to whether or not he liked math, Sam merely nodded in affirmation. However, these behaviors did not diminish or detract from his ability to share his mathematical knowledge. He gave the correct answer for the first item almost instantaneously after computing mentally, and wrote the required
number sentence before being asked by the clinical interviewer. Sam also was friendlier to the clinical interviewer at the end of this first item.

Clinical Interviewer: And if I asked you to make a number sentence...oh, you already did it. $9-3=6$. Very nice. Thank you. Umm, did you read my mind or did you read ahead?
Sam: (Smiles and says): I read ahead (CI, lines 34-41).
Sam experienced no difficulty with the measurement item and easily gave the correct answer and explanation. Although he had received a rating of developing for measurement, for this particular item he performed at a secure level. For the third item on elapsed time, Sam's mental computation led him to an answer that was incorrect by only five minutes. A probe by the clinical interviewer allowed Sam to discover and rectify his error.

Sam: 4:20
Clinical Interviewer: And how did you decide 4:20?
Sam: Umm, because if it was $3: 30$ and you put 45,40 minutes to it, it would equal, umm, 4:10, and if you put the five next to it, it would equal 20.
Clinical Interviewer: I see. So tell me that again. If you put, ah, 40 minutes to it, it would equal what?
Sam: Umm, ...4:10
Clinical Interviewer: 4:10. Ok. 4:10. And another ...Ok. That's 40 minutes and then what were you going to do next?
Sam: Umm, ...add five minutes
Clinical Interviewer: Add five minutes... so what would that be then?
Sam: Umm, 4:15
Clinical Interviewer: 4:15. Would you like to write that down? (CI, lines 70-97)
This probe not only allowed Sam to correct his mistake, it also highlighted his prowess with mental math. He did not need to use the clock manipulative to determine the answer. However, the fact that Sam's answer was five minutes off substantiates Liane's statement that he had challenges with those intervals when finding elapsed time. In this area he was definitely operating at a developing level.

Sam easily understood the pattern item. His prompt response and reasoning matched Liane's rating of secure in this category. For the coin word problem, Sam computed mentally. In order to get some insight into his mental cogitations, the clinical interviewer invited him to tell what he has done.

Clinical Interviewer: You want to say what you're thinking out loud so I can hear?
Sam: Umm...
Clinical Interviewer: I could hear you whispering 25, 50...
Sam: (Places the coin he has counted out in front of the clinical interviewer. He still doesn't verbalize what he's thinking).
Clinical Interviewer: (Looks at them) What's this? Can you tell me what this is?
Sam: Umm... 6, 6 quarters, and two, and two dimes, and ... a nickel
Clinical Interviewer: Can you count this for me?
Sam: $25,50,75$, a dollar, a dollar twenty, a dollar twenty-five, a dollar fifty, a dollar sixty, a dollar seventy, a dollar eighty
Clinical Interviewer: What's this? Five cents ... a dollar seventy, a dollar ...
(and pointes to last coin that he has grouped together)
Sam: (Looks at coins) 75, 85, 95 (and reaches for coins in the group of coins that he has not used in his group and places coin in).
Clinical Interviewer: 95. Ok. So you've used...
Sam: Four dimes
Clinical Interviewer: Ok. What if we put...how many quarters? So we could write Q ... Good.
Sam: (Writes 6Q)
Clinical Interviewer: ... and how many dimes?
Sam: Four. ...and a nickel (CI, lines 132-168).
By questioning Sam and encouraging him to share his thinking the clinical interviewer again afforded him the opportunity to fix a computational error and get the first part of the item correct. Unfortunately, despite spending several minutes, he ended up arranging the coins in the same way for the second part of the item, thus getting it incorrect.

Sam experienced more success with the item on probability. He colored both spinners differently. However, when the clinical interviewer tried to get him to explain how he had made his choices, he did not (or could not) comply with the request.

Clinical Interviewer: Sam, I have a question. How did you decide how many to color of each color?
Sam: Umm ... (makes no further response as stares at his spinner).
Clinical Interviewer: I saw you stop and look and count or something, then you shook your head. You nodded. Do you know what you were thinking? You did a good job.
That's ok. You don't have to tell me (CI, lines 276-283).
Sam was either unwilling or unable to explain his thinking despite the efforts of the clinical interviewer. This underscores the challenges that the clinical interviewer faces in trying to reveal the interviewee's thinking. Nonetheless, even though Sam did not or could not explain his strategy, he successfully completed the item. This was not the case with the first sub-question for the bar graph item that came next on the protocol.

Sam chose two names instead of one, and in this instance he was unable to use the probe to recognize where his reasoning had gone awry.

Clinical Interviewer: Ok, so you came up with a couple of names. You came up with Trent and Andrew. And how did you decide that?
Sam: Umm ... because Andrew was only one book away from Trent, and so I put both of them.
Clinical Interviewer: I see. How many books did Tyesha read?
Sam: Three
Clinical Interviewer: And how many books did Andrew read?
Sam: Six
Clinical Interviewer: And how many books did Trent read?
Sam: Seven
Clinical Interviewer: Seven. Umm... anything else you want to tell me about that?
Sam: (Shakes his head indicating no) (CI, 314-334).
For this item, the probe did not help Sam to realize that Andrew was the correct choice. Given his aptitude with basic facts, the reason for his error here is not easily inferred. In spite of this lapse, his talent for mental math was revealed in the next sub-question as he added numbers totaling twenty-nine. In contrast with his lack of details for the
probability item, Sam was able to explain that in solving this problem he had started with the higher numbers and then had worked his way down through the smaller numbers. This kind of insight into children's thinking is what makes the clinical interview so useful.

Clinical Interviewer: And, what did you do to find that answer?
Sam: Umm ... I added the higher number first. I added seven first, then I added umm...then I added six and that equals thirteen, and then I added four and that equals 17, and then I umm... I added four more and that equals 21 ...
Clinical Interviewer: Uh, huh. I think somewhere you probably added five in also. 'Cause did you go from the biggest number and added going down? Sam: (Nods yes) (CI, lines 346-358).

Although Sam missed some numbers during his explanation, his response provided important information about a mental math strategy that he had used successfully. A probe for the third sub-question allowed Sam to rectify his error there, and he was able to complete the final sub-question without the need for adult intervention. So for the bar graph item, probes had mixed success. His responses suggest that Sam's performance on the bar graph item was more representative of a developing rather than of the secure rating that he had received from Liane in this conceptual category.

His teacher had also rated him as secure in symmetry, but in completing that item, while Sam identified the three figures that were symmetrical, he also chose two that were not. However, he was able to describe what it meant for a figure to be symmetrical "Umm ... because I looked if they were half and half" (CI, line 407). His response shows that Sam understood what elements were needed for a figure to be classified as symmetrical and supports Liane's rating of secure in this area. His ability to succinctly
explain the key aspect of symmetry makes his inclusion of the two non-symmetrical figures puzzling. Sam's performance on the coordinate grid item was not as puzzling. Here, he also clearly knew what was needed, but inadvertently reversed the numbers in a couple of the coordinates. Thus, he was able to look at the grid and correctly identify the coordinate pairs for four of the six locations. Despite the reversals of the numbers in the two pairs of coordinates, his performance on this challenging item is in agreement with Liane's rating of secure in this category. He had to look at both axes and choose the appropriate pairs.

The next item, a word problem with money, also required him to select from among choices to meet a particular specification. Sam computed mentally and quickly wrote down the appropriate answers. No adult intervention was needed here and Sam surpassed Liane's rating of developing on this item. For the next item, the multi-digit word problem, Sam was not as successful. This was due in part to the fact that in reading the item to him, the clinical interviewer omitted the part of the question that indicated that a number model was all that was required. However, Sam reread the question silently to himself. It was after this rereading that he attempted to find the actual solution for the item. This indicates that he did not understand the directions of the item and therefore attempted to find the actual answer. Although he tried hard, including rereading the instructions a second time, he was ultimately unable to arrive at an accurate actual solution.

Misunderstanding directions also contributed to Sam's incorrect answer to the chart item that followed next in the clinical interview protocol. Like Kevin, Sam
immediately focused on using the total of nine tickets, to the detriment of realizing that he needed to go on as many rides as possible. As the clinical interview finished reading the item, Sam immediately responded as follows:

> Sam: Ooh!! I can see what ...
> Clinical Interviewer: Ok, what can you see?
> Sam: Umm... Four plus three equals seven plus two equals nine (CI, lines 497-501).

Sam's proficiency with mental math caused him to leap to an inaccurate assumption here. Since he was able to instantaneously find rides that totaled nine, he did not pay sufficient attention to the other part of the directions. His response therefore did not meet the secure level that he had been assigned in this category, but it does again illustrate his misunderstanding of the directions.

Place value was the next conceptual category and Sam's secure rating in this area was fully validated. He easily completed both parts of the item without adult intervention. The final item on fractions proved to be more complicated. Sam came tantalizingly close to the correct answers after probes. Promisingly, in the first part of the item, he shaded in nine of the fifteen rectangles as had been requested.

Clinical Interviewer: So, how many squares, they are not really squares, how many rectangles are left?
Sam: Six
Clinical Interviewer: Six. Could you write how many are left as a fraction?
Sam: (Wrote 6/9)
Clinical Interviewer: And how did you decide that?
Sam: Umm ... because there are six left ... and ...
Clinical Interviewer: There are six left ...
Sam: Because there are six left, and, and you already used nine
Clinical Interviewer: There are six left, and you used nine. Ahh ... how many were there in all?
Sam: Fifteen

Clinical Interviewer: Fifteen. So would you write your fraction the same way or would you write it differently?
Sam: Differently
Clinical Interviewer: Ok. Go ahead
Sam: (Wrote 9/15)
Clinical Interviewer: Ok. So does that show how many were shaded or how many were un-shaded?
Sam: How many were shaded
Clinical Interviewer: Ok. And how would you write how many were un-shaded? Sam: (Wrote 0/6)
Clinical Interviewer: Ok (CI, lines 544-580).
The probes revealed that Sam knew how to find the shaded portion of the figure. However, despite being able to use the accurate numerator and denominator for the shaded portion of the figure, he got confused when asked to find the fraction for the unshaded portion. The Everyday Math curriculum mainly asked the students to find the shaded portion of figures, so Sam's facility there is understandable and explains why Liane gave him a rating of secure in this category.

Overall, based on his clinical interview performance, Sam provided the supporting evidence for Liane's assertions in the first teacher interview that Sam was right in the middle. She felt that he was developing in general, but that he had some secure skills. For some items, probes were instrumental in allowing Sam to figure out the correct answer, while in other cases such as the fraction item, they were ineffective. Sam's explanation of his mental math strategy on the bar graph item provided useful insight into his mathematical reasoning and highlighted this benefit of the clinical interview method. As with Kevin, language interpretation proved faulty with two items on the protocol. Additionally, Sam did not explain his thinking for the probability item and that counters his ability to explain his computational strategy for the bar graph sub-
question. Finally, Sam's responses on the clinical interview belied his initial body language that suggested disinterest in the task at hand. Sam worked diligently on all items on the clinical interview and availed himself of adult intervention whenever he was able. I was not present while Sam took the NJPASS assessment, but based on his scores there I can infer that he worked conscientiously there as well.

Sam
NJPASS results.
Sam's NJPASS scores for the items that corresponded with the clinical interview items were 24 out of 30 or 80 percent of them correct. On the items on the entire mathematics section of the NJPASS, Sam earned a level of Proficient as he got 29 out of 40 items or 73 percent correct. The Item Response Record provides more details about these cumulative scores.

## NJPASS Item Response Record

As I explained previously, the Item Response Record divides the students’ scores into five different conceptual categories. Each item on the NJPASS belongs to a particular category and each of the five sections ends with an open-ended item. In discussing Sam's performance, I once again focus only on items that corresponded with clinical interview items in the various categories.

Number sense, operations, and properties.
In this category, Sam got one of the two corresponding fraction items incorrect. This was a multiple-choice item that required the student to choose the figure that was three-quarters shaded. He did not get the total amount of points for the open-ended item that matched the coin item on the clinical interview. He earned two out of a maximum of
three points. There are no details as to why he did not earn the maximum points for this question. He was more successful in his answers for the next conceptual category of measurement.

## Measurement.

Sam got all the corresponding items correct in this section of the assessment. This included the open-ended item that matched the clinical interview word problem about money. Sam earned full points for this item. The other item was a multiple-choice item that matched the yarn item on the clinical interview. So in this section, measuring in centimeters and calculating with money provided no obstacles for Sam. He also did not find the items in the next section unduly problematic.

Spatial sense and geometry.
The second corresponding fraction item proved to be no challenge for Sam in this section, and he got it correct. He also accurately answered the coordinate grid multiplechoice question. That is consistent with his success on the open-ended version on the clinical interview. On the open-ended symmetry question that completed this section of the NJPASS, Sam again earned two of the three possible points. Not getting the full points for this item raises the question as to whether he had selected correct and incorrect figures as he had done on the clinical interview. The information on the item response record again provided no details from which to make a determination about the reason for the score. Knowledge of the particulars of his work was not as necessary in the next section due to Sam's proficiency with the multiple-choice items.

Data analysis, probability, and discrete mathematics.

Sam answered all the corresponding multiple-choice questions accurately in this category. These items included the probability items that used spinners as well as the two bar graph items. For the open-ended item, Sam earned one of the three possible points. This question was congruent in design to the chart item on the clinical interview, and raises the possibility that he could have made similar errors with the directions on both items. A mistake in computation caused Sam to get an inaccurate answer for one of the multiple-choice items in the final conceptual category on the NJPASS.

## Patterns and algebra.

Given his facility with basic facts it was surprising that Sam got the corresponding basic fact item incorrect on this section of the NJPASS. However, he got the more challenging elapsed time problem correct. For the open-ended item, Sam earned the maximum three points. This question examined students' knowledge of extended patterns, and he had also experienced no challenges with the similar item on the clinical interview.

In sum, Sam's performance on the NJPASS was illustrative of a student who was at the developing level in most of the categories. Measurement was the category in which Sam performed best on the corresponding items as he got all of them correct. In terms of open-ended items, Sam was inconsistent. While he earned full points in two of the categories, in the other three he did not. Recalling Sam's inability to explain how he had decided on the colors for the spinners on the clinical interview, the question is raised if Sam experienced similar challenges in unpacking his thinking on these NJPASS openended items. Also, his challenges with the chart items suggest that some of the difficulty
with these items could be attributed to incorrectly understanding directions. As Liane had indicated during the first teacher interview, Sam experienced some difficulty with comprehending instructions. Sam's performance on the NJPASS items shows a student who had a good grasp of the mathematical concepts, but who would benefit from more practice and experience in some areas in order to attain mastery. Like Kevin, Sam did well on both the NJPASS items and the clinical interview. When Liane watched his clinical interview, she not only commented on his performance, but she also made comparisons between both assessments.

## Sam

## Post teacher interview.

Liane's pride and pleasure in Sam's mathematical ability were patently obvious as she watched his performance on the clinical interview. She nodded her head, and her face reflected her satisfaction as she watched him effortlessly complete the first two items. As the interview progressed to the third item on elapsed time, Liane noted that during the NJPASS they were not allowed to use clock manipulatives. Number grids were the only manipulatives that were permitted. Antithetically, during the clinical interviews the use of a wide array of manipulatives was encouraged as a means of getting children to share their thinking while they solve problems.

While watching the clinical interview, since Sam was taking a considerable amount of time to complete the item, Liane shared that in the Chicago Math curriculum, the concept of clocks had not come around again before the NJPASS in keeping with the spiraling nature of the program. They had occasional examples in the math boxes, but no explicit lessons on the concept were given. In preparing the review for the NJPASS she
realized that more would need to be done with calculating elapsed time as the concept had not been formally presented to the children in some weeks. This information was important as it confirmed that Liane was not teaching to the test, but was instead following the district-mandated curriculum. Review was done in areas where it was felt that the program might not have adequately prepared the students.

Liane: And the clocks especially, clocks have not come around again in that full circle, so when we were looking to review for NJPASS we were like, oh my gosh, we have to do some more with clocks, because it just didn't come around...
Clinical Interviewer: It wasn't in the math boxes?
Liane: Oh, yes. I do the math boxes too but we just felt like formally... I use that as a math message in the morning, along with the journal, and we go over it, and we do, but it's one out of six problems...
Principal Investigator: It could get lost in the shuffle...
Clinical Interviewer: Right
Liane: Yes, where the other units, the units in math, most of that was covered. The number stories, the measurement... measurement was actually right before this (TI2, lines 802-817).

While this discussion was occurring, Sam was still working on the coin item.
Liane again took the opportunity to make further comments.
Liane: He, umm ... I am just watching his body language too, because you had that down. Yawns a lot, yawns so much in school. I don't know about his bedtime ... you know it just seems that more than Kevin ... seems tired Principal Investigator: Body language more disaffected, tired ...
Liane: Yes. Just see it a lot in class. You're talking and he yawns (doing the action) and he's right in front, so I see him (laughs) (TI2, lines 839-846).

Her words made it clear that Sam's posture at the beginning of the clinical interview was not unusual for him. They helped to explain why despite his outward demeanor Sam worked diligently during the clinical interview. Liane's reflections as she watched Sam's performance also helped to shed light on his lack of explanation for his strategy during the probability item. It was noted that although Sam clearly had a strategy for completing
the spinners differently he was unable to articulate it. In response, Liane stated, "That’s not unexpected either ... with Sam" (TI2, lines 948-952). This indicated that Sam had experienced difficulty with explaining his thinking not only on the clinical interview but also in the classroom. As she watched Sam complete the final item on fractions in the clinical interview, Liane reiterated that the students were accustomed to working with shaded parts of the figures, which could have contributed to Sam's challenges with the second part of this item. Despite these concerns, Liane was pleased with Sam's as well as with Kevin's performances on the clinical interviews. "Yeah, they did good. Nice" (TI2, line 1269).

To conclude, Sam's performances on the clinical interview and the NJPASS assessment exemplified a student with developing skills who needed reinforcement in certain areas. Probes varied in effectiveness depending on the level of Sam’s prior knowledge as well as on his understanding of the directions. Sam's inability to explain his thinking during the probability item illustrated the importance in the clinical interviews of interviewees being able to explain their reasoning. It reinforced the importance of language as a theme. His performance also raised the issue of differing levels of difficulty between open-ended versus multiple-choice items. Each student had differing levels of success with both types of questions as Kevin's and Sam’s performances show. Oliver, whose work is discussed next, added to the variation of the replies to the items due to his nuanced thinking on some of the items.

## Oliver

First teacher interview.
Oliver attended E School and was seven years and eight months old when the study was conducted. Danielle was his teacher and she had taught for a total of 15 years, nine of which had been spent in second grade. She had a Master's degree in reading plus thirty credits in administration and supervision. Using the Everyday Math rating scale with which she had been provided, Danielle gave Oliver the following ratings in the requisite conceptual categories. He was rated at the secure level in all categories except money (making change), elapsed time, bar graphs, grids, and charts, and word problems. In these categories he was evaluated as being at a developing level. During the interview, Danielle stated that while completing addition and subtraction algorithms Oliver used mental math and paper and pencil. He used manipulatives when provided with them, but he did not really need them. He did not use his fingers. In explaining his thinking Danielle indicated that Oliver used a combination of methods including oral and written explanations as well as diagrams and pictures. She too supported her evaluations with resources such as workbook pages, open-ended questions, pre-tests, end-of-unit tests, math games, teacher-created assessments, teacher/student conferences, and teacher/student observations.

When asked if she wanted to share anything else about Oliver's mathematic ability, Danielle's response was positive.

Principal Investigator: Anything else about Oliver ... that you'd like us to know about his mathematical skills, ability, his approach to math, or anything ... Danielle: Just that he is confident. He is a confident student. And he umm... he likes to explain how he got his answers.
Principal Investigator: Ok.

Danielle: So he speaks very low, very softly ... but he likes to share how he got his solutions. If we ask how did you get it, how did you get it, he likes to solve, he likes to share how he gets his solutions.
Principal Investigator: Does he think out of the box? Are there some times that his solutions are ...
Danielle: He , he can think out of the box
Clinical Interviewer: When he shares is he generally right...
Danielle: Yes
Clinical Interviewer: Or sometimes does this thinking go awry?
Danielle: No. He's pretty much on target when it comes to solutions (TI1, lines 314-345).

This conversation is noteworthy because in his clinical interview Oliver gave unexpected responses to some of the items on the protocol. While they were unanticipated, and were definitely examples of thinking outside the box, they were not always accurate. Also, given his propensity for explaining his thinking, the clinical interview method would appear to have been tailor made for him. I had anticipated that he would have done extremely well on its tasks. However, as the scores indicate, his performance on the NJPASS surpassed not only his work on the clinical interview but also the scores of the other students in the study. Oliver's ability to come up with unusual yet accurate solutions appeared to have contributed to his success on that assessment. This is in contrast to his performance on the clinical interview where his performance was not as stellar.

## Oliver

## Clinical interview.

Oliver's penchant for solving problems slightly differently from the customary or expected way was displayed from the start of the clinical interview. After easily dispatching the first item using mental math, when asked to write his answer, he stated that he had done it backwards. "I did it backwards" (CI, line 19). The problem had
required subtraction, but Oliver wrote the number model using addition. Instead of writing nine minus three equals six, Oliver put down three plus six equals nine. He was the first of two interviewees who answered in this manner. His response supported Danielle's belief that he was secure in his basic facts and also demonstrated his use of mental computation.

The second item on measurement provided no challenges for Oliver, nor did the third on elapsed time. However, the promptness of his response to the latter was quite surprising, partially because he had received a rating of developing in this category, and partly because of the high caliber of mental computational ability that his answer revealed.

Clinical Interviewer: Um ... the next question says your tee ball practice lasts forty-five minutes, and it begins at 3:30. What time does your practice end? Oliver: Umm... 4:15? (He answered almost immediately). Clinical Interviewer: 4:15. How did you do that so quickly? Oliver: Um, 30 minutes is 4 because, I mean, 30 plus 30 is 60 so that equals o'clock and then 30 plus 15 equals 45 so I added 15 to the 30 and got 4:15. Clinical Interviewer: Ok. Could you write 4:15 down there for me? I liked the way you explained your thinking. It was very helpful. Thank you (CI, lines 3748).

Additionally, the clarity of his explanation substantiated Danielle's belief that he not only enjoyed sharing how he arrived at solutions but had also become quite proficient in its execution.

The pattern item was not problematic for Oliver and his response was befitting to his secure rating. The coin item necessitated that the students use coins to make the same amount in two different ways was marginally more challenging for him. The clinical interviewer used probes to help him catch minor computational errors for both parts of
the item. Nonetheless, he was ultimately successful in finding two separate approaches that allowed him to get the required amount. Oliver was the third interviewee, and given the ease with which the two previous students had completed the probability item, his response to the item was startling. He quickly colored the first spinner with the requisite colors. Then when he was subsequently asked to use equal amounts of the same colors on a second spinner but in a different pattern from the first spinner, Oliver was unable to comply with the request.

Oliver: (He looks at spinner, and counts. Then...) That's the only way. Clinical Interviewer: You think that's the only way you can make it?
Oliver: Um, hmm (meaning yes).
Clinical Interviewer: Ok. How many sections is it divided into?
Oliver: Two
Clinical Interviewer: Each color is divided into two?
Oliver: (Nods yes)
Clinical Interviewer: Ok. How many sections is the whole spinner divided into?
Oliver: Eight
Clinical Interviewer: Ok. So is there a way that you could arrange the colors differently? So that you would have an equal amount, but it wouldn't necessarily look like that?
Oliver: (Thinks about it, then again counts the spaces on the spinner. He puts his hands on his head after looking at it for a long time).
Clinical Interviewer: Is there another way to show equal amounts of red, yellow, green, and blue that doesn't have that same arrangement?
Oliver: I don't think so. (He still has both hands on the top of his head).
Clinical Interviewer: Ok (CI, lines 140-172).
As she proceeded to try to get Oliver to realize that there was indeed another way, the clinical interviewer first had to discern that Oliver's reply to the first probe was incorrect. The spinners had been divided into eight and not two. Oliver had interpreted her query to refer to the divisions of the colors and not the actual spinner. Clarifying that distinction allowed her the opportunity to pose the question about the sections again. This time, Oliver was able to answer correctly. However, he was still unable to arrive at a solution.

The body language with his hand on his head appeared to simultaneously denote both deep thinking as well as puzzlement.

Prior to a probe from the clinical interviewer, Oliver's inability to color the second spinner differently was followed by yet another surprising answer.

Clinical Interviewer: Who read twice as many books as Tyesha?
Oliver: (He is thinking, not responding, with his hand on his jaw). Uh, no one?
Clinical Interviewer: No one. Can you see how many books Tyesha read?
Oliver: Yeah.
Clinical Interviewer: Can you check the amount of books the other people read?
Oliver: (Does not respond)
Clinical Interviewer: And what does twice as many mean?
Oliver: Hmm, half of something I think
Clinical Interviewer: Half of something? Ah, which one would be half, Tyesha or the other person? You want paper to write on? (She hands him paper).
Oliver: Uh, huh (meaning yes). (He seems unsure about what to write). Andrew?
Clinical Interviewer: Ok. And how did you decide on Andrew?
Oliver: Because $6 \ldots$ and 3 is equal to 6 (CI, lines 178-210).
The unexpected response that Oliver made in response to the question is an example of how the clinical interview can follow unexpected trajectories based on the nature of an interviewee's answer. The probes provided the clinical interviewer with the opportunity to have Oliver re-examine the numbers again. Her interventions also revealed that he did not know what the key term 'twice as many' meant. Nonetheless, Oliver managed to arrive at the right answer although his explanation was unclear. He easily solved the remaining sub-questions for this item. In doing so he once again revealed his superior mental math skills and used addition to explain the solution for a subtraction problem.

Symmetry was the concept for the next item on the protocol. Oliver was only able to choose two of the three symmetrical items despite a probe to get him to recheck
his solution. When asked to name another item that was symmetrical, Oliver selected a crayon. Curious about his answer the clinical interviewer probed further.

Clinical Interviewer: A crayon. Would that be symmetrical? How would you, how would you cut ... how, where would you draw the line of symmetry on a crayon?
Oliver: (Takes a pencil and puts it vertically over the crayon so that it divides it in two with the point at one end and the blunt edge at the other end).
Clinical Interviewer: Ok. And would it be symmetrical?
Oliver: (Peering intently at his pencil on the crayon) Ah, I think so ...I think it would be symmetrical.
Clinical Interviewer: You think it would be?
Oliver: (He nods his head in affirmation).
Clinical Interviewer: Um ... so if the crayon would be symmetrical, why wouldn't this be symmetrical? (She points to the picture of an apple that is the first picture in the group).
Oliver: There is no leaf there.
Clinical Interviewer: I see. Ok. And how about this one? (She points to a mitt). Oliver: This one doesn't have an end.
Clinical Interviewer: And how about this one? (She points to the pig).
Oliver: There is no ear on the back.
Clinical Interviewer: Ok. And the crayon is symmetrical? (She places the crayon in front of him, next to the non-symmetrical pictures).
Oliver: Yes (CI, lines 273-314).
Although Oliver could clearly see why the other figures were not symmetrical, even when the clinical interviewer placed the crayon with the pencil indicating the line of symmetry that he had chosen, Oliver was unable to see the similarities among the figures and thus recognize his error.

For the coordinate grid item, Oliver initially seemed unable to see that he had made a mistake by giving only one number of the coordinate pair for the first location. Eventually after probes by the clinical interviewer he realized that he needed two numbers, but he unfortunately reversed the order of the pairs. Using mental math he solved the next item, a word problem that involved adding money, easily. While reading
the multi-digit subtraction with renaming item, the clinical interviewer accidentally omitted the section of the directions that said that only a number model was required. Using mental math, the number grid, and paper and pencil Oliver made a valiant if unsuccessful attempt to find the actual solution.

In reading the next item based on the chart, the clinical interviewer emphasized that there needed to be as many different rides as possible. Oliver clarified for himself that there were nine tickets. Then as Kevin and Sam had done before him, he too chose rides that totaled nine, but that did not meet the other criteria. With the place value item, Oliver again gave an answer that was unpredicted due to his secure rating in this concept. When the clinical interviewer read the item, after completing the first part of the item, Oliver seemed uncomprehending of what the directions for the second part meant.

Clinical Interviewer: Could you write another number that would be in the hundreds?
Oliver: (Sighs, still standing. Looks puzzled).
Principal Investigator: It would have hundreds, tens, and ones like that number, but you make it up.
Oliver: (Still doing nothing). I don't get it.
Clinical Interviewer: What did you say ... you don't get it? Um ... well what was this number?
Oliver: Eight
Clinical Interviewer: The whole ...
Oliver: 857
Clinical Interviewer: Ok. So can you make up another number that would be in the hundreds? It doesn't have to use those digits but can use any digits that you want. Just any big number like that.
Oliver: (Writes)
Clinical Interviewer: You wrote 959. That's a nice number. Is in the hundreds? Oliver: Yup.
Clinical Interviewer: Yup, it is. Good (CI, lines 491-521).
His response showed that he had challenges with understanding the language of some directions on the protocol. However, adult intervention again enabled him to give an
appropriate solution. Probes were also instrumental with helping Oliver successfully answer the final item on the clinical interview protocol.

Oliver got the first part of the item correct. However, when asked to give the remainder of the figure as a fraction, he gave the fraction for the part that he had shaded. As was previously mentioned, the students were accustomed to finding the shaded rather than the un-shaded portion of figures. When a probe helped him to realize that he had given the fraction for the shaded part of the figure, Oliver was easily able to determine the correct answer.

In summarizing Oliver's performance during his clinical interview, there are certain adjectives that come immediately to mind. They are excellent, surprising, and illustrative. Oliver displayed excellent mental math computation as exemplified by his answer to the elapsed time item. Many of his answers were surprising such as his responses to the probability and place value items. Using addition to explain subtraction was unusual in comparison with the other students in the study. Finally, his performance showed how necessary it is for the clinical interviewer to be prepared for any deviation that occurs during the clinical interview. Despite not being completely accurate, Oliver's responses were thoughtful and interesting. Although I am unable to comment on the nature of his responses on the NJPASS, another adjective comes to mind with reference to his performance there - superb. An examination of his performance on the corresponding NJPASS items supported this opinion and details are presented below.

## Oliver

NJPASS results.
Oliver's NJPASS scores on both the corresponding and actual items were outstanding. He earned 29 out of 30 items correct for the corresponding items on the clinical interview, and 38 out of 40 items correct on the actual items on the NJPASS assessment. As a result, Oliver achieved the level of advanced on this assessment. He was the only student in the study to attain this standard. Since Oliver got only one corresponding and two actual items incorrect on the NJPASS, an analysis of his item response record will be commensurately brief.

## NJPASS Item Response Record

In the Number Sense, Operations, and Properties category, Oliver got all the multiple-choice items correct. However, he did not get the full points for the open-ended item that corresponded with the coin item on the clinical interview. He earned two of the three possible points. An analysis of Oliver’s performance on the Measurement, Spatial Sense and Geometry, Data Analysis, Probability, and Discrete Mathematics sections reveals that he got all the items in those categories correct. In patterns and algebra, Oliver got an elapsed time multiple-choice item incorrect.

In sum, in contrast with his performance on the clinical interview and with some of the ratings from Danielle, Oliver performed at a superior level on the NJPASS assessment. While his responses on the clinical interview required frequent probes and provided interesting nuances with regards to comprehension of directions and completion of items, Oliver clearly had no challenges with the NJPASS items. His NJPASS scores are indicative of a student with exemplary mathematical knowledge. His performance on
the clinical interview showed a student who had good but not exceptional math skills. Danielle's ratings would support this evaluation since they included areas in which Oliver was considered to be in the developing category. In the first teacher interview Danielle had indicated that Oliver was a confident math student. The NJPASS scores confirm this opinion more so than his responses on the clinical interview. Danielle's comments about his work on the clinical interview will be presented in the next section.

## Oliver

Post teacher interview.
Danielle was pleased by Oliver's performance on the first items as evidenced by her delighted laughter as she watched him answering the questions. When the DVD progressed to the coin item, Danielle commented that the wording on the clinical interview was different from what Oliver was accustomed to hearing. The clinical interview asked the students to list two different ways to show the amount. Danielle stated that the children were more used to being asked to show two different ways. This minor difference in wording was not problematic for Oliver, but the issue of the language of directions is relevant given the challenges Oliver experienced with some items during the clinical interview.

In discussing the probability item, Danielle indicated that the Chicago Math program did not address the concept, and she had just practiced with spinners and different activities the week before the NJPASS. Her revelation is further confirmation that the students were exposed to a balanced curriculum that adhered to district guidelines and that the boys had not been subjected to a narrowed curriculum. Since Oliver took a long time to complete this item, Danielle commented that Oliver was one of her slower
workers. This is reminiscent of Liane’s observation about Kevin especially since both boys were the same age.

On the bar graph item, Danielle was proud of Oliver's ability to use addition to explain a subtraction problem. For the symmetry item, Danielle noted that Oliver had ignored the point on the crayon when he asserted that it was symmetrical. With the coordinate grid item, Danielle mentioned the First Move chess program. The second graders had just started working with this addition to the curricula that year.

Commenting on his performance, she stated that she had hoped that the program would have helped him to get the coordinates correct. "Since we’ve been doing chess so much I thought he would have figured that out 'cause whenever we do chess ... it's the letter and $\ldots$ it's H, $5 \ldots$ D, 3... Yeah. But I don't know if we use the word coordinates with them" (TI2, lines 546-547; 552-553).

Danielle seemed surprised that Oliver had solved the money word problem mentally although she had indicated in the first teacher interview that he did use mental math. Her response was also unanticipated given Oliver's indisputable command of mental math on both the clinical interview and the NJPASS. For the multi-digit subtraction word problem, Danielle indicated that even on the corresponding NJPASS items, the students had found the problem challenging.

Principal Investigator: On the NJPASS it was multiple choice?
Danielle: Yeah, and they still had trouble.
Principal Investigator: And they still had trouble. Ok. That is important for me to know.
Danielle: I was just going to say that was a tough one for them (TI2, lines 586-593).

Danielle may have noticed other students in the class experiencing difficulty with this item, but Oliver's scores show that it posed no challenge for him. Danielle’s reflections and comments about Oliver's performance on the actual items during the clinical interview ended at this point. Like Liane, Danielle had two students who participated in the study. A description of the second student's ratings and his work on both assessments is presented below.

## David

## First teacher interview.

David attended E School and was seven years and eleven months old at the time of the study. Danielle thought that David had achieved a level of secure in the conceptual categories of counting money, telling time, symmetry, and basic facts. He was assigned the rating of developing for fractions, measurement, probability, patterns, bar graphs, charts, grids, and place value. She evaluated him as working at a beginning level in making change, finding elapsed time, and in word problems. During the interview, Danielle indicated that David did not use mental math, used his fingers, used paper and pencil if asked, and used the number grid and counters. He preferred oral explanations and only used diagrams when prompted. Danielle used the same resources to support her evaluation of David's mathematical ability that she had used with Oliver.

When asked if there was anything else that she wanted to share about David's progress in mathematics, Danielle stated that he was very inconsistent.

Danielle: Umm ... the only thing with David is that he's very inconsistent. If there's something I think he has completely as a secure skill, the next day he would totally surprise me and like go the other way.
Principal Investigator: Ok.

Danielle: And vice versa. Sometimes when I think ... when he doesn't get things at the beginning ... he does (TI1, lines 168-175).

Danielle then gave an account of her observations of David during the NJPASS assessment that corroborated her evaluation of him as being inconsistent.

Danielle: I can't figure it out at all. It's totally ... We do math pretty much at the same time every day... Uh...I've had him for two years so this is my second time with him, and uh, I've seen tremendous progress, but sometimes ... like just looking during his standardized test this week I was very pleasantly surprised when I happened to be peeking over thinking he would have so much trouble with a particular thing and he was right on, and something I really thought he was gonna do ... You know his ... he, he would have not issues, my heart sank (TI1, lines 194-200).

In addition to her concerns about David’s inconsistent retention of mathematic content, Danielle also noted that she needed to reword directions for him.

Danielle: Very often I need to reword directions, and if he is having trouble with a skill, if I reword and just change the way I say it to him, sometimes he just like gets it, he just needs to be ... (TI1, lines 219-221).

This example concluded Danielle's first interview about David. The description of his clinical interview follows.

## David

## Clinical interview.

In answering the first item, David used an addition number model to represent his answer as had Oliver before him. When asked why he had chosen to write that he explained, "Because six and three equals nine" (CI, line 20). Despite getting a rating of beginning for basic facts, David had no difficulty arriving at the correct answer and in giving a valid explanation of his choices. Similarly, the second item on measurement presented no challenges for him. Unfortunately, this was not the same for the third item on elapsed time. In keeping with being at a beginning level in this conceptual category,

David was unable to arrive at a reasonable answer. Based on his answer, the clinical interviewer realized that David did not have the requisite prior knowledge that could be assisted by probing, and so she accepted his response without further comment.

Clinical Interviewer: Umm ... your tee ball practice lasts about 45 minutes and it begins at 3:30. What time would your practice end?
David: Umm ... (Scrunches up his face, looks at clock that is next to his paper). Clinical Interviewer: You can use the clock, you can move it, you can count on it, you can do whatever you want to do with it.
David: (Picks up clock and manipulates it. He is holding it down in his lap, so what he is doing is not visible. Then he holds it up in his hands on the table, but not flat, so we still can't see how he is moving the hands. Then he shows the clinical interviewer the clock).
Clinical Interviewer: Ok. And what time did you come up with?
David: 8:30
Clinical Interviewer: Could you write that on your paper (CI, lines 38-54)?
David's body language and behavior, his scrunched up face and shielding his movements as he manipulated the hands of the clock, as well as the implausibility of his answer, illustrated his perplexity with the concept.

He fared somewhat better on the next item. He was able to complete the pattern and was able to describe how he had figured it out. Then something went awry. He was unable to say whether the pattern was increasing or decreasing even with a probe that asked if the pattern was going in a particular direction.

Clinical Interviewer: Can you tell me what number pattern you used to figure that out?
David: I counted by twos.
Clinical Interviewer: Ok. Could you just write that there? By two. Ok. And was there anything else that was important to know about that pattern?
David: Umm... (Looks at his paper. Then he shrugs his shoulder indicating that he doesn't know what else to say).
Clinical Interviewer: Is it going in any particular direction?
David: (Nods head).
Clinical Interviewer: What was the direction?
David: Umm ... it was skipping one.

Clinical Interviewer: It was skipping one? Ok. It was skipping a number.
Anything else?
David: (Shakes head indicating no).
Clinical Interviewer: Ok (CI, lines 63-89).
Despite the probes from the clinical interviewer including rewording the first probe to be more specific, David was unable to give the correct response. He also forgot his original answer of counting by twos and said the numbers were skipping by ones. The excerpt shows that he was indeed at a developing stage in this concept and thus was more prone to be confused even with reworded questions. This example also highlights the fact that probes are only effective if the students have enough knowledge to utilize them.

Although David got a rating of secure for counting money, he was not able to use coin manipulatives to successfully depict the same amount in two different ways. For the first way he used all the coins and exceeded the required amount. For the second way the coins he listed did not add up to the requisite amount. The clinical interviewer tried to get him to explain his answer, as this could have led to him recognizing his error, but he was not able to do that.

Clinical Interviewer: And how did you come up with that answer?
David: Umm ... because ... (stops and stares and says nothing else).
Clinical Interviewer: Not sure?
David: (Shakes his head indicating that he is not sure) (CI, lines 144-150).
David computed this problem mentally and used no paper and pencil. This is noteworthy as, notwithstanding the inaccuracy of the computation, his use of this cognitive strategy contradicts Danielle's assessment that he did not use mental math.

He was successful in completing the probability item and exceeded his rating of developing in this area. For the bar graph item, David had mixed success as he got two
of the four sub-questions correct. It was interesting to note that as a strategy, David did not use the actual numbers to calculate who had read twice as many books. Instead he compared the actual lengths of both bars. Mental math again featured in this item as David once again used this strategy, albeit unsuccessfully. Finally, probes helped him get one sub-question correct, but did not work with another sub-question.

The symmetry item allowed David to reveal his confidence in this area as he competently explained how to ascertain if a figure was symmetrical. He then completed both parts without difficulty. For the coordinate grid item, a probe revealed that David did not realize that two digits were required for each pair of coordinates. He was therefore unable to complete that item accurately. For the money word problem, the clinical interviewer repeated the directions after seeing David's hesitancy to continue his response. He then was able to find the correct answers. The incident illustrated Danielle's point that repeating directions helped David to understand them.

With reference to the multi-digit word problem, like his peers before him, David missed that only a number model was required. The clinical interviewer clearly read that a number sentence had to be written to show how to solve the problem. However, David's actions indicated that he thought he had to actually solve the problem. After a while, he simply guessed and told the clinical interviewer an improbable answer. "I'd put out 100 more" (CI, line 414). Once again, based on her observations, the clinical interviewer did not intervene with probes, but moved on to the next item. This was the chart item. Unbeknownst to David, in focusing on getting rides that totaled nine he made the same error as the other three boys. As a result, he got this item incorrect. However,
he also used mental math in finding its solution. David had no challenges with the place value item in spite of Danielle's note that he could be inconsistent in his work in this category. For the final category, fractions, David got the first part correct, but was unable to write the un-shaded section of the figure as a fraction.

His responses to the final item exemplified his performance on the clinical interview as a whole. David had mixed success although he worked hard throughout the entire clinical interview. Sometimes he would be able to solve the items accurately and easily, while for some concepts his responses were inadequate. Probes were not always effective with David and from time to time his lack of prior knowledge made their usage impractical. He occasionally experienced difficulty in expressing his thinking orally and if he was at a developing stage in a concept he would get confused and uncertain. Yet, he benefited from having directions re-read and used mental math in multiple items. In contrast with this uneven performance on the clinical interview, David's responses on the corresponding NJPASS items were much better. The next section provides specific information about David's performance on this assessment.

## David

NJPASS results.
David got 20 out the 30 corresponding NJPASS items correct and accurately completed 28 of the 40 actual items. This accomplishment earned him the level of Proficient on the entire NJPASS assessment. The NJPASS Item Response Record gives more details about David's performance in the five main conceptual categories.

Number sense, operations, and properties.
In this category, while getting all the other multiple-choice items correct, David got two of the three designated mental math problems incorrect. Missing two-thirds of these types of questions would indicate that David had not yet attained a consistent level of proficiency in using this cognitive strategy. Only one of these two questions had a corresponding item on the clinical interview, the money word problem or item ten. For the open-ended question at the end of the section, David earned two of the three possible points. This question corresponded with the coin item on the clinical interview where he had been less successful so this score was an improvement. Overall, David did well on this section. The analysis of his NJPASS performance now shifts to the category of measurement.

## Measurement.

The multiple-choice items in this section were answered correctly. However, David earned one point out of three for the open-ended item. As this item corresponded with the elapsed time problem on the clinical interview, and given his rating of beginning in this concept, it is encouraging that he was able to score one point. David's improved performance was especially evident in the next category.

Spatial sense and geometry.
David got all items correct in this section. This included the open-ended question that explored symmetry. His excellent work in this area was unfortunately not duplicated in the next category.

## Data analysis, probability, and discrete mathematics.

The multiple-choice bar graph items proved to be challenging for David and he got both of them incorrect. Additionally he was not able to answer the open-ended chart item and earned none of the three possible items. His responses to the next and final category are very similar to his work in this one.

## Patterns and algebra.

David got two multiple-choice items incorrect in this section as well. One of them corresponded with the elapsed time item on the clinical interview. The open-ended item matched the pattern item on the interview, and contrary to his response on that item, David was able to earn the full points on the NJPASS.

As the above analysis shows, David's accomplishment on the NJPASS items was a great improvement over his score on the clinical interview. Although certain areas such as mental math, elapsed time, and graphs continued to be challenging for him, David was able to provide more accurate answers to the NJPASS items than he had been able to do on the other assessment. The difference in scores between the assessments serves as an example of the inconsistency of David's mathematical ability. It underscores Danielle's evaluation in the first teacher interview. Her observations about David's clinical interview follow.

David
Post teacher interview.
Danielle watched David work through the protocol as avidly as she had watched Oliver's achievements. However, she had less to say about David's performance. For the initial items, she just watched David on the DVD without commenting. When it
came to the coin problem, she shook her head when she realized that for the first part of his answer he had just put all the coins together. Later she confirmed that he had experienced difficulty with the bar graph item based on her observations during the clinical interview. She stated that the fact that the graph went horizontally and not vertically was very different for the students as in the math books the graphs were generally vertical. However, she agreed that the direction of the bars was not the main reason that David struggled with the item. Danielle noticed that for the multi-digit item he never really started. "I don’t think he has any idea of how to start" TI2, line 202).

David's clinical interview performance did not elicit many comments from Danielle except the above and she made no more observations about what she had seen. Liane and Danielle had two students each. The next student, Jake, conversely had two teachers. They were Dahlia, his homeroom teacher, and Becca, his special education teacher. Dahlia's first interview where she gave her evaluations for Jake is described in the next section. It will be followed by Becca's first interview, Jake's clinical interview, and Dahlia's and Becca's post interviews.

Jake
First teacher interview -Dahlia.
Jake was seven years and ten months when the study was conducted. He had an individual educational plan (IEP). His homeroom teacher, Dahlia, had taught for a total of 23 years, eight of which had been in the second grade. She had a Master's degree in education. In her first teacher interview, Dahlia thought that Jake was secure in the conceptual categories of symmetry, bar graphs, grids, and charts, and in addition facts up to twelve. She evaluated him as being developing in fractions, measurement where she
noted that he confused centimeters and inches, money (knowing the values of coins), patterns (counting forwards), subtraction facts, and place value. For money (making change), elapsed time (confusion with five minute intervals), patterns (counting backwards), and word problems (reading challenges) she believed he worked at a beginning level. For the category of probability she rated him on a continuum of beginning to developing.

Dahlia's positive opinions of Jake were evident from her comments in this first interview.

Dahlia: What I love about my student is that he loves to, he loves math, so he's willing to do anything and everything. So, he is not, not a difficult child by any means to work with. He wants to learn, he wants to do all the stuff (TI1, lines 74-76).

In addition to this opinion, Dahlia shared that Jake needed repetition and refreshing of previously taught concepts in order to successfully complete math questions. Dahlia believed that much of Jake's challenges in mathematics resulted from his struggles with reading. He had difficulty reading word problems independently, but once he was helped to understand what was required he could determine the necessary mathematical operations to find the solution. Manipulatives were very helpful to Jake's problem solving efforts, but these tactile-kinesthetic benefits were not enough to compensate for his difficulty in processing information. As a result he needed extra time in completing assignments. Dahlia also shared that Jake often focused more on the mechanics of writing rather than on the content of what he wanted to say. Written explanations were thus more challenging for him as he expended much energy in making sure each letter and number hit the lines perfectly.

Dahlia: Extra...Oh, totally extra time. Extra time and uh, he's very, I guess, careful. We have a slant board, he has some fine motor issues so with the writing, uh, we have him, you know. So he's very slow and careful in making his letters to hit the lines, and he's...his numbers, so he kind of gets, perseverates would be a good word too on that part (TI-1, lines 357-364).

Dahlia shared that Jake did not use mental math. However, he used his fingers, paper and pencil, and manipulatives. At this time, Dahlia clarified her comment about Jake's use of mental math.

Dahlia: Going back to mental math ... like if it's ...we play a game called Around the World where I will write a problem on the communicator and it will be you know between two people and umm... whoever can try to solve it because we said you want to try to get away as much as you can from using the number grid, use ... try to own it. So he can do some of the basic facts, uh, mentally (TI1, lines 405-409).

She further shared that Jake benefited from peer modeling with regards to improving his oral explanations. Dahlia encouraged him to share his problem solving strategies by following the examples of his cohorts. The practice of articulating his answers in the classroom enhanced his ability to explain his thinking during the clinical interview.

The resources Dahlia used to support her evaluations were similar to those used by the previous teachers, although she placed less emphasis on workbook pages and pretests. When asked if she had any other information about Jake's mathematic abilities, Dahlia shared that mathematics was his favorite subject. Jake self-advocated and was honest about admitting that he needed help. She felt he wanted to learn and put much effort into his work. As a result, he was learning more concepts. Jake was persistent, committed and was not lazy. "He doesn't want to give up. He doesn't take the easy way out. He’s not of that less is best philosophy, you know ..." (TI1, lines 536-537). Dahlia
began and ended her first interview by voicing affirmative views of Jake's mathematical efforts. These perspectives were similar to those held by Becca, Jake's special education teacher.

Jake
First teacher interview - Becca.
Becca was Jake's special education teacher. She had taught for a total of 16 years, with eleven of them being in her current position. She had a Bachelor's degree in education plus 20 additional credits. In rating Jake's mathematical abilities in the various conceptual categories, Becca felt that he was closer to the secure level in place value and in his basic facts to ten.

Becca: Place value, this one. Uh, secure, developing, but closer to secure, and again that was something that was done in small group. Uh, a lot of modeling, using communicators, and putting the digits in the columns (TI1, lines 111-113).

She felt he was at a developing stage in the areas of fractions, measurement, money, and patterns. Becca used a continuum to describe his rating in probability, symmetry, and elapsed time, placing him at a beginning/developing level in these categories. The following excerpt explains Becca's reasons for this dual classification for one of these categories.

Becca: Elapsed time, that was, uh, interesting. That was very beginning, but this past week we were divided into small groups and I saw the student do much better after two other students answered a question and the student caught right on (TI1, lines 52-54). And I had fear ... uh, but now I'm right in between, closer to developing (TI1, line 58).

Not only does this quotation reveal Becca's reasoning, it, along with the previous example, also supports Dahlia's belief that peer modeling was extremely helpful to Jake's mathematical progress.

In continuing with her evaluation, Becca assessed Jake as working at a beginning level in bar graphs, grids, and charts, in basic facts higher than ten, and in word problems. For the latter she indicated that he needed them read to him two or three times corroborating Dahlia's observations that he needed repetition in order to grasp concepts.

Becca: Word problems - that's the weakest part ... so word problems are definitely ...
Principal Investigator: Beginning
Becca: Beginning - must have them read to him. And sometimes it's pairs ... we put two together $\ldots$. and we do a lot of teamwork together and sometimes it's a group with an adult overseeing. So the adult will read, and we give strategies go back and look, what were your clues on what to do... So he can circle clue words and underline and then after that, two or three of being read, he'll come up with... (TI1, lines 169-178). He still requires either peer tutoring with that or the teacher modeling and saying to underline that part. Reading is difficult, and then he takes longer to just process the sentence (TI1, lines 195-197).

Becca's reference to Jake's processing challenges is reminiscent of Dahlia's
comments in her teacher interview. His disability was evident during the clinical interview and clearly impacted his performance there. The NJPASS scores also reflect his lack of success on that assessment. Having the items on both assessments read to him did not overcome Jake's challenges with processing information. However, despite his difficulty with processing material, Becca indicated that Jake handled himself very well.

Becca: Takes longer to process and will always say could you read that again. Advocates, advocates for himself. And yeah, he knows he needs wait time, and doesn't seem to get too frustrated or upset, and he comes through.
Principal Investigator: So he uses wait time.
Becca: Yes. He uses it and he comes up with his answer and a big grin and he just feels so good and that's the most wonderful thing. He works hard and he gets his answer, and it’s nice. So he make take a little longer, but I see him progressing, and he feels pride, but I do see with this student more modeling is needed. More on a daily basis to help him remember processes (TI1, lines 201211).

For example, reviewing homework directions helped him remember what to do when he did his homework at home. Becca also shared that Jake used manipulatives primarily the number grid and occasionally paper and pencil. She also indicated that he did not use mental math as much. Becca stated that Jake preferred to communicate orally as written explanations were harder for him. Again, this supports Dahlia's evaluation. However, when asked if she had anything to share about Jake's mathematical experiences, Becca alluded to a disability that Dahlia had not mentioned.

Becca: I see uh, visual, visual problems, so geometry ... shapes, creative patterns are difficult for him (TI1, lines 283-284). Using the measuring tool ... seeing those lines on the ruler, the half ... that was very difficult. I had to take a clear one and the little black lines I had to use a colored marker and then we blew it up, a nice page from the workbook and put it in a plastic protector and used different colors to highlight, and when it was this big, he said, "Ooh! Now I see." So the tiny scale on the ruler was very hard for him to see (TI1, lines 301-305).

Jake did not only have challenges with the ruler, but he also had to be specifically taught how to navigate the rows and columns on the number grid. Becca ended the interview by sharing more positive comments about Jake.

Becca: He is such a good student. He is always focused on the teacher. It's a wonderful way. He does not get distracted by this, or this, or this. So distraction has nothing to do with it. He, he’s working and taking a little bit longer to process the information (TI1, lines 391-394).

Jake's clinical interview showed his strengths and weaknesses. Its key elements such as reading the items aloud, giving as much time as needed to answer the questions, and allowing the use of a variety of manipulatives helped him to maximize his strengths as much as he was able. A summary of Jake's clinical interview is presented below.

## Jake

## Clinical interview.

The solution to the first item required Jake to complete a subtraction basic fact.
Both Dahlia and Becca had indicated that Jake had more difficulty with subtraction than with addition, but in this case, Jake easily found the solution. Dahlia's observation that he confused centimeters and inches, and Becca's description of Jake’s struggle with seeing the markings on the ruler were substantiated by his performance on the second item. While Jake knew that an inch was different from a centimeter, when asked to pick a string that measured a centimeter, he picked the one that was an inch long.

For the elapsed time item, Jake performed closer to Becca’s developing level than Dahlia's beginning level. Although he did not get the item correct, he was only five minutes off. A probe revealed the reason for his error, but he was unable to use it to see that he had made a mistake.

Jake: 3:30 ... 45 minutes?
Clinical Interviewer: Yeah, what would be 45 minutes after 3:30?
Jake: (Manipulates clock hands and quickly gets an answer). 4:10.
Clinical Interviewer: 4:10. Uh, can you show me how you counted that?
Jake: 5, 10, 15, 20, 25, 30, 35, 40, 45
Clinical Interviewer: Ok. So when you said the first five, you started on the number six when you said five?
Jake: (Nods yes).
Clinical Interviewer: Ok. Is that the way you should do it?
Jake: (Nods yes) (CI, lines 59-79).
Despite Dahlia's evaluation that he struggled with counting by fives, Jake had no problem doing that in this item. His mistake was in confusing where to begin counting by fives. He started on the six when he should have started on the seven. This item also
showed how Jake tried to address his processing problem. He repeated the key elements of the question to make sure that he had grasped what he had to do.

Jake had absolutely no difficulty in completing the pattern item and was able to describe that it was going backwards. This was good because Dahlia had said he was at a beginning level if patterns went backwards, while Becca had said he was developing in this area. For the coin word problem, Jake got one of the two parts correct. The directions were clarified for him and he also re-read the directions silently. This fit in with his teachers' observations that he often needed instructions to be read aloud to him multiple times. However, Jake's use of mental math to solve this item was surprising. Both his teachers had stated that he used little or no mental computation. This was not his only similarity to the previous students. Like them, Jake took a long time to complete the item.

For the probability concept, Jake had been evaluated at developing/developingbeginning levels. Jake colored the first spinner very slowly and carefully. The painstaking effort illustrated Dahlia's statements about Jake's perseveration with the mechanics of writing, in effect his fine motor skills. His self-advocacy and need for clarification were evident in this item. After coloring in the red, he checked his paper to see what color to do next.

Principal Investigator: You can choose the yellow, the blue, or the green next. Whichever one you want.
Jake: Any color?
Principal Investigator: Yeah, any one of the other colors. Yes
Jake: Like I can use green?
Principal Investigator: Hmm, hmm (yes) (CI, lines 213-221).

Again Jake tried to make sure that he understood what was being asked. Unfortunately, despite his efforts and those of the clinical interviewer, he still got confused about coloring the second spinner.

Clinical Interviewer: Ok. Now can you take this spinner and also have it come up with equal amounts of red, yellow, blue, and green, but not look exactly the same? It needs to look different.
Jake: (Colors the spinners only blue and green). Done.
Clinical Interviewer: Ok. So are those equal?
Jake: Hmm, hmm (meaning yes) (CI, lines 227-239).
It appears that Jake heard and only processed the last two colors and missed the remainder of the directions. When the clinical interviewer asked if the colors were equal, referring to the fact that there had to be four equal colors, Jake thought she was asking about the way he had colored the spinners. He had forgotten the original instructions. As a result, he got the second part of the item incorrect.

Jake was also not able to accurately solve three of the four sub-questions on the bar graph item. The one sub-question that he got correct was as a result of a probe. However, while trying to find the answers, Jake used more mental math. In addition, he used a good strategy of writing down the number model after trying to compute mentally. Then he also used his fingers. In this way he used strategies that were recognized by his teachers as well as mental math that was not identified as something he used. Contrary to his performance on the last two items, Jake excelled at the symmetry item and worked at the secure level indicated in Dahlia's evaluation. Not only did he identify the symmetrical figures, but he was also able to choose a symmetrical object that was not presented and accurately draw its line of symmetry. Jake's success on this concept was not repeated for the next item.

On the coordinate grid item, Jake gave only one number for each location.
However, a probe by the clinical interviewer elicited that he knew that there should be two numbers for each coordinate pair. He just didn't know how to find the second number in the pair.

Clinical Interviewer: Do coordinates sometimes have two numbers or just one? Jake: Huh?
Clinical Interviewer: Do coordinates that help us find things have two numbers or just one?
Jake: Two
Clinical Interviewer: Oh, so what would be a second number that would go with ...
Jake: (Silent)
Principal Investigator: You could go back to R and see what other number would go with the four.
Jake: (Still silent)
Clinical Interviewer: That's ok. If you think of it later we can go back to it (CI, lines 448-464).

While a probe didn't work in this item, adult intervention allowed Jake to solve the money word item. Unfortunately, like the other students, Jake could not find the solution to the problem. Similarly to his peers before him, Jake tried to find the actual solution by mental computation and by using paper and pencil. Then, he requested a number grid. Shortly afterwards he arrived at an answer that regrettably was incorrect as the directions had not been followed. Misunderstanding directions also caused Jake to get the chart item incorrect. This was again similar to his peers as he made the identical error. On the place value item, while Jake clarified what was required, this time he got both parts of the item correct without the need for probes.

Jake's response to the final item on the protocol was the most unique among the students. The figure had fifteen rectangles arranged as five columns with three rectangles
in each column. While the others saw the rectangles as separate, Jake did not see the figure that way. He saw it as five columns. So even though he colored the nine rectangles as asked, he thought of it as three columns. This was revealed when the clinical interviewer asked him to tell how many rectangles were left.

Clinical Interviewer: That asks how many squares (rectangles) are left?
Jake: Two
Clinical Interviewer: Two are left? Ok. Then it asks, it asks if you could write your answer as a fraction. How could you make that into a fraction?
Jake: I forgot what fraction means.
Clinical Interviewer: Uh, if I have a cookie, and I cut it in half and I give half to you and half to me. I get one half and you get one half. So do you know what would be a fraction of this?
Jake: Two
Clinical Interviewer: Two. In the cookie it would be a fraction of two parts? So in this, what would the fraction be for how many pieces are left?
Jake: Five
Clinical Interviewer: Ok
Jake: Two
Clinical Interviewer: Five over two? Ok. Could you write that? And how did decide five over two?
Jake: Because you have five of these and there is two, there is two left
Clinical Interviewer: Ok. I see how you got that (CI, lines 598-629).
This excerpt shows Jake's special perspective as he viewed this problem. His visual/spatial issues could have contributed to the way he viewed the figure.

Unfortunately, this was a case when viewing something in an unpredictable manner did not lead to an accurate resolution. Becca had been the one who had explained Jake's visual challenges. It was Dahlia who had mentioned his desire to do all that was required of him in math. This was exemplified when at the end of the clinical interview, even though he still didn't know how to find the second number in the coordinate grid pairs, he reminded the clinical interviewer that she had offered to go revisit the item if necessary. Jake's performance on the clinical interview was an accurate portrayal of his
mathematical knowledge and skills and dovetailed well with his teachers' evaluations. His teachers’ comments about his clinical interview performance in the second teacher interviews corroborated their earlier ratings.

## Jake

Post teacher interview - Dahlia.
Dahlia stated at the beginning of the interview that she couldn't address how he had done on the NJPASS because he had taken the assessment in the small group setting of the resource room as required by his Individual Educational Plan (IEP). Commenting on the fact that Jake had only been five minutes off in the elapsed time problem, Dahlia noted the source of his error but indicated that she was very pleased that he had the necessary foundations to do as well as he had. For the pattern item she noted that saying the pattern went backwards was his way of saying that he had subtracted. "To him backwards equals the subtraction" (TI2, line 72). For the coin item she noted that the children had recently done a similar item as part of their mathematics curriculum. Dahlia appreciated the fact that Jake had reread the question on his own. She also felt that the students had to complete a variety of things in order to solve the question.

Dahlia: ... I liked how he went back.
Principal Investigator: Hmm, hmm (yes). He reread the question.
Dahlia: So that is like a lot of different things...
Principal Investigator: What?
Dahlia: No, then they had to go back and write ... and make sure they knew the names of the coins (TI2, lines 88-97).

The different steps added to the complexity of the item and contributed to the fact that all the students took a long time to complete it. In addition, the coins used were plastic coins and Dahlia commented that she used real coins. Given Jake's visual issues, he probably
had more difficulty distinguishing the differences among the plastic coins. Dahlia commented that she taught the students how to organize money when they were counting it. As she watched him do mental math for this item, Dahlia thought that he was looking towards me for clues. However, he was actually using mental math to add up the values of the coins.

In Dahlia's opinion, the probability item was also more complicated than the similar items on the NJPASS.

Dahlia: This was on the NJPASS - well the probability, but they had spinners to look at with, you know, so they had that, so this is totally kicking it up a notch. Clinical Interviewer: Well, these spinners are done.
Principal Investigator: ...are done. They are already divided.
Dahlia: They are divided. Right, the ones on the test are. They have numbers on them. What are your chances of spinning a five? Are you less likely, not likely... Principal Investigator: Right
Dahlia: ... you know, so they had the ones to look at. Here is actually creating which is great (TI2, lines 212-225).
Principal Investigator: So you're saying that this spinner question is harder than the one on the NJPASS?
Dahlia: I am thinking it is.
Principal Investigator: Thank you.
Dahlia: I am thinking it is 'cause they just had to look at spinners, and umm... color in the bubble of basically what are the chances, you know ... so this one, here he is creating the spinner to show equal, you know, that it's ... and for Jake, the fine motor skills take him a little bit longer, the cutting, the writing of numbers, the coloring ... (TI2, lines 260-269).

Later, after Jake completed the second spinner using only blue and green, instead of the four colors as he had for the first spinner, Dahlia explained that his confusion was due to his challenges with processing.

Dahlia: Oh, oh!
Clinical Interviewer: It said red, yellow, blue and green ...
Dahlia: I heard you say that ...

Clinical Interviewer: But I ... when he colored so much of the blue, I thought what's going on, and I saw that he had colored half blue and half green, and I don't know whether it's the student or ...
Dahlia: It's him.
Clinical Interviewer: So he only picked up the last part of the phrasing, but I said blue and green last ...
Dahlia: Yeah ... but you also said make it look like... you know, don’t make it look like this one, so he didn't make it look like, (laughs), you know, make it look like the first one
Clinical Interviewer: And I think that was attributed to language more than anything, language processing.
Dahlia: But it's the processing that ...
Clinical Interviewer: Right
Dahlia: For him (TI2, lines 326-354).
This example demonstrated the challenges Jake had with language processing. It also is similar to some of the other students' inaccurate understanding of some of the directions on the protocol.

Having said that, Dahlia was pleased to realize that the bar graph item was worded like the NJPASS item, and again as Jake attempted to solve the sub-questions, she stated that she was satisfied if he knew how to solve the items. "Hmm ... I am just happy if he knows the process" (TI2, line 400). Jake did mental math to try to ascertain the total number of books that had been read. The clinical interviewer explained his strategy of finding the total of two of the numbers, and then adding the other numbers to that sum one at a time. Dahlia was proud of Jake for using that strategy. "That's good! He just kept building on that number" (TI2, line 439). She observed that Jake was always looking up at the sky, and that he did that when they played Around the World, the game that allowed children to practice their facts through mental computation. The excerpt below supports the notion that although Dahlia said that Jake only used mental math
during that game, he had extended his usage of this computational strategy to the clinical interview.

Dahlia: I like that ... always looking up to the sky. Look at him.
Principal Investigator: He's visualizing.
Dahlia: And that's what he does. We do our Around the World game just to help increase fact power, and uh, he, he does that (TI2, lines 451-456).

She wondered why he hadn't asked for the number grid although she was pleased that he was trying to solve the sub-question without using it. She wondered if he had simply forgotten that it was there. "But there's different rea ..., you could think why he didn't, he's doing it on his own, or he forgot that he had the number grid available, that it was even, you know" (TI2, lines 476-478). Regardless of the reasons why Jake decided not to use the number grid, the fact remains that he employed mental math as his cognitive strategy for this problem. Given his disability, this was an important accomplishment for him.

After this discussion about Jake's use of mental math, Dahlia shifted the conversation to query about the length of time that he had taken to finish the items on the protocol. She found it interesting to learn that he had fit in with the general time frame of the other students. As some of the other students were also slow workers, Jake’s disability did not put him at a disadvantage in this area. For the symmetry item, Dahlia noted that Jake's object, a football, was one that they had used when the class had been learning about the concept. She was unsure if he had remembered it from then as that had been done a long time prior to the clinical interview. "And that was one of our cards, but I'd be surprised if he remembered that 'cause that was a long time ago" (TI2, lines 538-539). In addition to using the symmetry cards to help with teaching that concept,
like Liane, Dahlia also used the chess program, First Move, to help the students learn the concept of coordinate grids.

Dahlia: The chess program has helped a lot with the plotting of grids. They know, umm ..., there's letters and numbers on the board, of course here both are numbers, but with chess it's letters and numbers. And that has you place the different pieces in different ... so that has been helpful (TI2, 549-552).

As she watched Jake only give one number in each pair as his answer, Dahlia noted that the clinical interviewer had tried to get Jake to see his error. "Rephrasing too, you know. You said it all different ways to try to ..." (TI2, line 573).

Moving on from the coordinate grid item, Dahlia also felt that the multi-digit word problem item on the NJPASS had been difficult for the students despite the fact that it had been multiple-choice. "This one was a killer on the test" (TI2, line 609). For the fraction item, Dahlia understood that Jake had seen the rectangles as columns. When the clinical interviewer stated that Jake hadn't wanted to give up, Dahlia concurred by saying, "And he doesn't, he doesn't give up. He really is a good little man. He wants to do, he wants to succeed, he wants to do his best" (TI2, lines 731-732).

Dahlia's second interview ended not only with these opinions, but also with her noting that Jake had used mental math as well as paper and pencil during the interview. Oral language was his preferred way to explain his thinking, and as the interview progressed, his body language had become increasingly more relaxed. Dahlia's comments about Jake's clinical interview showed the close attention she paid to his responses. Becca, his other teacher was also intrigued by Jake's work on that assessment.

Jake
Post teacher interview - Becca.
When Becca heard Jake's answer to the measurement item, her insightful comment was, "[He] didn’t hear his own words" (TI2, line 21). For the elapsed time item she was disappointed that he had gotten the problem incorrect, but was glad that he knew what to do. "Because we practiced. Awww... So he knew what to do" (TI2, line 37).

Becca was thrilled at the speed at which Jake was responding to the questions.
Becca: This is quick.
Principal Investigator: Hmm?
Becca: He is responding so quickly. I love it.
Principal Investigator: Hmm, hmm!
Becca: I think the quiet is helping him ...Nice! (This was in response to a good answer from Jake). So many students that call out the answers when they're not supposed to ... I have never seen him...
Principal Investigator: Speak up
Becca: I have never seen him be the first to respond. That was a highlight (TI2, lines 58-74).

Jake did not get many items correct on the clinical interview, and had mixed success with probes. However, one thing that Becca felt that the clinical interview had allowed him to do was to be the first to respond to an answer. When he got the first part of the coin item correct, Becca was very happy. "This is a big improvement since the beginning of the year" (TI2, lines 92-93). When she noted how a probe helped him to realize his error, she bemoaned the fact that such adult intervention was not possible on the NJPASS. "But we're not allowed to do that on the test. But this is ... what he needs" (TI2, lines 121125). It was also what they did in class for example when Jake would be asked to recheck an answer. Becca abandoned the topic of adult intervention to return to the unaccustomed speed at which Jake was working. She was really impressed by this.

Becca: I have to say I think the peace and quiet in this room is a big factor because he is acting so quickly. I love it. I don't see this quickness.
Principal Investigator: Hmm!
Becca: I do see him get things correct.
Principal Investigator: Oh, yeah.
Becca: But it takes him longer.
Principal Investigator: Hmm, hmm
Becca: And sometimes there is a lot of partner work, and he is not always the first partner to come up with the suggestions (TI2, lines 135-149).

Becca admitted to sometimes having "purposeful groups" (TI2, 158) so that he could get the opportunity to be first one to successfully answer a question. She reiterated her assessment of Jake’s self-advocacy and persistence. "He will ask a question, or he will say, could you please read that again. He always advocates. He is determined, he does not give up" (TI2, (lines 170-171). While persistence is a good trait, perseveration is not as desirable. Like Dahlia, Becca noted his compulsion with the mechanics of writing and the fact that this negatively impacted his time management. Therefore, she was pleasantly surprised that given his penchant for erasing, he had only erased once up to that point. Another positive that Becca noticed was Jake's ability to request additional information.

Jake frequently had trouble processing language. Sometimes he did not know some of the vocabulary in the directions such as 'twice as many' (see Appendix F, pp. 323-328). However, Becca noted that when he asked questions to clarify his understanding, his meaning was always very clear. "He prefers to ask, rather than read. But when he asks a question, it's always clear, the language is always clear" (TI2, lines 209-210). Becca was distressed that Jake had only processed the last two colors for the second spinner. She felt that she would have clarified the directions for him had he been
in the classroom. We noted that the clinical interviewer only realized what he had done when he was finished. While she was surprised at his answer for the probability item, she was not disconcerted by his overall lack of success on the bar graph item. " Doesn't surprise me. Visual is not his strength" (TI2, line 252). Thus, she was thrilled when he successfully identified the symmetrical figures. "Excellent! I am impressed. He drew the line of symmetry" (TI2, line 364-366).

Becca understood how Jake's unique visual perspective resulted in his seeing the fraction item as columns.
Becca: So he answered two looking at the columns instead of each box. Uh, huh. (She is not surprised when Jake says he has forgotten what a fraction is). This was difficult.
Principal Investigator: See he is doing it $-1,2,3,4,5, \ldots 5,2$. So he got five over two because he was seeing...
Becca: The columns that's here ... He came up with 2, 2 digits. Knew parts and total. He’s developing (TI2, lines 504-515).

She was impressed that he knew that he needed two digits. This was because "Fractions and measurement were absolutely his weakest in the class" (TI2, line 528). The interview moved from discussing what Jake answered to how he answered the items.

Like Dahlia, Becca, realized that Jake had used mental math as well as paper and
pencil. She also agreed that he preferred oral to written explanations.
Becca: I saw mostly mental math Principal Investigator: Yes
Becca: Mental math ... used his fingers. I was thrilled to see him ask for a grid because the grid is always available. He uses the grid in class. Mental first, then he goes to the grid (TI2, lines 573-579).

This observation was surprising, because in her first teacher interview, Becca had indicated that Jake used little or no mental computation. This concluded her comments
about Jake's clinical interview performance. A description of his work on the relevant NJPASS items follows.

Jake
NJPASS results.
Jake's NJPASS scores were a little better than his clinical interview performance. He got 14 out of 30 items that corresponded with the clinical interview items correct. He got 18 out of 40 items on the actual assessment correct. In so doing, he earned the level of Basic on the complete NJPASS assessment. An examination of the categories of the NJPASS Item Response Record provides more details about Jake's performance.

Number sense, operations, and properties.
In this category, Jake got several multiple-choice items incorrect. Four of them corresponded with items on the clinical interview protocol. These were the multi-digit word problem, the money word problem, the coin item, and the place value item. Additionally, he got zero points for the open-ended item that that explored students' understanding of coin values. Jake didn't fare very well in the measurement category either, which supports Becca's assertion that it was one of the areas in which he was less knowledgeable.

## Measurement.

Of the three multiple-choice items in this category, Jake got one correct. He earned one point for the open-ended item. This was the only item that had any matches to the clinical interview. It addressed the topic of elapsed time and given Jake's misconception about where to start counting on the clock his lack of success with the
item is unsurprising. Since Jake had visual perceptual challenges, one could have assumed that he would not have done well on the spatial sense and geometry category. Spatial sense and geometry.

Such a prediction would have been inaccurate. Despite his impairment in the area of visual acuity, Jake only got one item incorrect in this section. The item did not correspond with any on the clinical interview. He also earned two of the three possible points for the open-ended item. It examined symmetry and his success here showed that Jake was very competent in this area. His proficiency in spatial sense and geometry did not extend to the next category. Jake got only half of the items correct.

Data analysis, probability, and discrete mathematics.
These items corresponded with the bar graph and chart items respectively. The latter was the open-ended item, and Jake earned one of the possible points for it. This was not entirely surprising, as he had also experienced difficulty with these items on the clinical interview. For the final category on the item response record Jake displayed more mathematical competency.

## Patterns and algebra.

The only multiple-choice item for which he had found an incorrect solution was on one the items that matched the elapsed time problem on the clinical interview. Jake earned two points for the open-ended item that explored patterns. There are no details explaining why he didn't get the full three points. He had no difficulty with the almost identical item on the NJPASS. A possible reason for loss of the point may have been that
on the clinical interview he was able to explain the pattern orally. On the NJPASS he had to write it and his struggles in that area could have contributed to the loss of the point.

In sum, an adjective that would describe Jake's performances on both assessments would be consistency. His scores on both the clinical interview and the NJPASS were very similar. The analyses of his responses on the individual items show that he constantly erred on the same concepts. His work exemplified a student who had developing to beginning skills in most of the categories. His acquisition of skills and knowledge were negatively affected by his cognitive processing issues and by challenges with visual perception and fine motor skills. While Jake was the only classified special education student in the group, the next student, Martin, was the only other student who scored in the Basic range on the NJPASS.

## Martin

First teacher interview.
Martin attended E School and was seven years and nine months old when the study was conducted. Rae was his teacher. She had been teaching for thirty years, and twelve of them had been spent in second grade. Her highest degree was a Masters in Teaching. She had also earned thirty additional credits.

In her first teacher interview, Rae thought that Martin performed at the secure level in measurement, probability, patterns, symmetry, and place value. She assessed him at the developing level in fractions, basic addition facts, and in solving word problems. For bar graphs, grids, and charts she considered that he fluctuated between developing and beginning.

Rae: Uh, bar graphs, grids, and charts ... almost developing to beginning because what happens is, um, when we are working with a chart in particular, he doesn't ... if, if I haven't set up the scenario, and he has to do this on his own, he doesn't look at the title, doesn't look at what it says over here, and sometimes confuses himself. But again, it's sometimes. The work in general is often inconsistent. That's why I have some papers with me (TI1, lines 75-80).

For the concepts of elapsed time and basic subtraction facts she felt he worked at a beginning level. With his subtraction basic facts she also indicated that it was "hit or miss" (TI1, line 99).

She didn't believe that he used mental math, but instead preferred to use manipulatives such as the number grid. He didn't use his fingers or paper and pencil. Rae believed that he used pictures and diagrams to explain his thinking, but would sometimes give oral explanations if able to do so. He frequently told her he didn't know or didn't remember when asked to describe how he got an answer.

Rae: Because, some... sometimes when I ask him to tell me how he figured it out, his most favorite answer is, 'I don't know, I don't remember." But when I pursue him further, he will give me something, but it is also inconsistent, so favored answer is I don't ... well, let me put "I don't remember" first because that's what he always tells me - "I don’t remember" (TI1, lines 165-169).

In addition to the resources used by the previous teachers to make their evaluations, Rae included a beginning of the lesson review that asked the students to reflect on the previous day’s lesson. She also added checking their homework sheets to the list.

When asked about Martin's strengths and weaknesses in math, Rae stated that while Martin understood the concepts of the fact triangles that the children used to practice their basic facts, he had not memorized the actual facts that were associated with the triangles. In addition to these insights, Rae shared the following.

Rae: I find when I work with him, he's really just wanting to get an answer down. It's like look at it, put an answer down. And when we talk through it, and we talk about, you know, the algorithm or whatever, and he looks better, then he makes a better decision. I think the fact, I think some of his immaturity surfaces when he is left on his own because he often strikes me often as just let me get the task done. But not in the sense of the child that's just trying to do a careless job. That isn't what I mean. You know he appears to take pride in his work, but it's very important for him to just do it, get an answer, do it, get an answer, like little rote kinds of things (TI1, lines 259-266).

Rae's most important challenge with Martin's work in mathematics was his inconsistency. He would seem to know a concept securely, and then if it was revisited a few weeks later, he would either do as well as he had previously or he would do very poorly.

Rae: ...In his workbook, you know, when you presented things to the children and said how are we doing he showed evidence of getting it, and it seems like it's gone. This was our unit on fractions. He certainly did very, very well on it, yet the thing that frustrates me as his teacher is that we could do this maybe two weeks later, and it's possible that he will do as well, or it's possible that I could get a score that's very much in the beginning category (TI-1, lines 337-347).

Rae felt very challenged by this fluctuation in performance and wondered if it could be attributed to Martin's reluctance to admit to not understanding the concepts as they were being taught. While there was no way of confirming Rae's suggestion, his clinical interview did provide insight into Martin's mathematical ability. A description of his work on the clinical interview follows.

## Martin

## Clinical interview.

Martin completed the first two items on the protocol accurately with no need for probes. He used mental math despite Rae's assertion that he did not use that cognitive
strategy. Martin was not as successful on the third item on elapsed time. He had no idea about how to go about solving the question.

Clinical Interviewer: So it lasts about 45 minutes. It begins at 3:30.
Martin: (Has one hand at the side of his head as he contemplates the question. He doesn't use the clock that is beside his left hand). I don’t know the answer to this one.
Clinical Interviewer: Ok. Do you want to try and think about it using this clock? Martin: Ok. (Looks at clock but doesn't move hands).
Clinical Interviewer: You can touch the clock, you can move the hands if it makes it easier.
Martin: Umm...
Clinical Interviewer: Or you could just talk about what you know and what you don't know about it.
Martin: Umm...it ends at umm...3:45?
Clinical Interviewer: You think it ends at 3:45? Ok. Thank you (CI, lines 52-73).
This example illustrates again the importance of students’ prior knowledge as it pertains to the efficacy of probes. Martin did not know enough about elapsed time in order to be able to use the probe as a tool to discover the correct answer. Rae had postulated in the first teacher interview that Martin was hesitant to reveal his lack of understanding about concepts. In this case, Martin very early on stated that he did not know how to solve the problem. Maybe the individualized nature of the clinical interview allowed him more freedom to admit his inability to provide an accurate response.

Martin easily solved the pattern item, and he volunteered his strategy when asked by the clinical interviewer. "Um, in my head I saw that it was minusing two" (CI, lines 101). Similarly the coin and probability items presented no challenges and he competently found the answers to both parts of these items. With the bar graph item,
finding the correct answers was more elusive for Martin. He did not understand the language of the first sub-question and that caused him to get the question incorrect.

Clinical Interviewer: Who read twice as many books as Tyesha?
Martin: (He is looking closely at the graph). Jamal.
Clinical Interviewer: Jamal. And how did you decide that?
Martin: Because, um, it says, um, who read twice as many books as Tyesha, and, um, Jamal does because, um, he read on more book than Tyesha.
Clinical Interviewer: He read one more. What does twice as many mean?
Martin: Um ... it means (stops and thinks) I think that it ... one more than.
Clinical Interviewer: Ok. So you think twice as many means one more.
Martin: Hmm, hmm (yes) (CI, lines 233-250).
This misconception was another example of where the language of the directions proved problematic to the students. When he was asked how many books were read in all, his initial response was not due to a misreading of the directions. Rather, his answer affirmed Rae's contention that Martin frequently put down any answer regardless of whether it made sense or not.

Clinical Interviewer: Ok. What does the next one say?
Martin: How many books were read in all?
Clinical Interviewer: Yes.
Martin: Um, how many books ...(Stops and looks at graph). Andrew (CI, lines 254-260)?

A probe from the clinical interviewer allowed him to get the correct answer, but had that not been the case, he would have gotten the sub-question wrong. In fact, for the remainder of sub-questions for this item, the clinical interviewer needed to use probes to get Martin to arrive at the accurate answers. Fortunately, the adult intervention was effective as he could use the information on the graph to help him. This was not the case with the coordinate grid item.

Martin did not realize that two numbers were needed for the coordinate pairs.
Despite a plethora of probes, Martin could not realize his error.
Clinical Interviewer: Ok. And is there anything else you usually do with grids? Martin: Umm ... we usually like, when there is nothing on there, we usually, um, color them in.
Clinical Interviewer: You leave, um, you color in the ones that are empty?
Martin: Yeah, my teacher tells us which ones to color in.
Clinical Interviewer: I see. Ok. So ... you don't put any other identifications there, any other directions for people to find the places on the grid?
Martin: Hmm, she tells us what number, which, and um, which, what, which number on the um ... grid to color in.
Clinical Interviewer: Ok. How does she tell you which grid, which place to color? What numbers does she give you to tell you?
Martin: Like, um, like it's kind of like the same one but it doesn't have letters. Clinical Interviewer: Ok (CI, lines 417-437).

He did, however, know that sometimes the coordinate pairs sometimes had letters and not only numbers. As with the coordinate grid item, probes did not help Martin get the money word problem correct. After he had chosen items that exceeded the maximum amount he could spend, the clinical interviewer tried to get him to see his error.

Clinical Interviewer: I have a question - has he used up all of his money yet?
Martin: Um ... (Puts his hand on his head and looks very unsure).
Clinical Interviewer: How will he know?
Martin: Um, yes.
Clinical Interviewer: You think he has used up all his money? How will he know?
Martin: Um, because he has one dollar and he spent it on a brownie, a cookie, and cupcakes, and on a raisin cookie.
Clinical Interviewer: Ok. Did he have enough to buy all of those?
Martin: Yes (CI, lines 468-483).
Probes were unnecessary on the multi-digit word problem as Martin followed the pattern of his peers in the study and found rides that totaled nine. Adult intervention was also not needed for the first part of the place value item. That was not the case for the second part of the question. When Martin was asked to create a number to match the one
he had just analyzed, he at first did not understand what was being requested. Through a series of probes, the clinical interviewer led him to a realization of what he needed to do.

Clinical Interviewer: Ok. Now can you write another number like 857, but not exactly the same?
Martin: Um ... (There is a long pause while he thinks) ...No.
Clinical Interviewer: What do you think they mean by that?
Martin: Um ... like, make it different and not the same.
Clinical Interviewer: Make it different, but with hundreds, tens, and ones. Could you do that?
Martin: No.
Clinical Interviewer: Could you write a number with hundreds, tens, and ones?
Martin: Oh! Yes.
Clinical Interviewer: Just any hundreds, tens, and ones? Ok. Let me see you write something. Ok. So you could do it. Very nice. Thank you (CI, lines 550572).

The final item on fractions also had two parts. Martin answered the first part correctly, but he confused the numerator with the denominator.

Clinical Interviewer: Ok. And the denominator stands for what?
Martin: The number up top
Clinical Interviewer: The number..
Martin: The first number in the fraction (CI, lines 601-609).
Due to his mistake Martin did not calculate the denominator accurately, and thus got the item incorrect.

To conclude, Martin's performance on the clinical interview illustrated many of Rae's comments. For the items in the areas in which she had evaluated Martin as being at the secure level, he was successful in solving them accurately. Where he was rated as developing or beginning, he experienced more difficulty. This was true except for his subtraction basic facts that he got correct. Probes were mixed for Martin. Sometimes they worked such as when he had enough prior knowledge, or when he could use the information from the item. At other times they were ineffective due to insufficient
knowledge. Martin worked very hard during the clinical interview, but his results varied. Rae's response to his performance is described below.

## Martin

## Post teacher interview.

At the very beginning of the second teacher interview, Rae's comments indicated that she believed that Martin had done well on the clinical interview. "Probably did real good with that (looking at the protocol)" (TI2, lines 39-40). She made no comment about his inability to do the elapsed time problem as it met the expectations she had stated in the first teacher interview. However, as she watched Martin work on the coin problem, she reiterated her concern about being able to accurately determine the level of his understanding.

Rae: You see how he's behaving?
Principal Investigator: Hmm, hmm (yes).
Rae: That's how he behaves in class. He's very, very reluctant to show what he knows and he often, umm, ... confuses me to the degree that I don't know if he really knows it or he's just insecure, unsure, and remember I also felt he's immature (TI2, lines 69-75).

Despite this concern, as she watched the interview, she noted that Martin was using mental math. "I did notice here that he was using mental ...math. Surprised me" (TI2, lines 173,177 ). Her comment corroborates her statements in the first teacher interview where she said he didn't use mental math in his algorithmic computations.

When the interview got to the bar graph item, Rae indicated that the children had trouble with the fact the bars were horizontal.

Rae: Yeah, it's very strange that when the bar graphs go this way (indicates with her hand horizontally).
Principal Investigator: This way

Rae: The kids have trouble. When it goes that way (vertically) they go right to the number (TI2, lines 313-319).

However, the bar's vertical alignment did not affect Martin's answers. Using mental math in the item did influence Rae's opinion of his use of this cognitive strategy. "...I didn't realize he had that strength as much as he seems to" (TI2, lines 341-342). In the same way that she had been concerned about the difference in presentation of the bar graph item, Rae also experienced some disquiet about the coordinate grid item.

She pointed out that the children were accustomed to letters on the axes. The items on both assessments used only numbers. Again, Rae was afraid that the difference in presentation would negatively affect the students’ ability to find the correct answer. She felt that Martin's struggle with the item was due to this dissimilarity between the items.

Rae: ‘Cause, let's see. He’s not getting this because it looks different than what we're accustomed to, 'cause we're doing chess this year.
Principal Investigator: Ok.
Rae: ...and they're really ... teaching them the chess has really helped them, but they're not accustomed to seeing the...
Principal Investigator: Numbers ... right
Rae: the same thing up and down. In other words if this was A, B, C, D, E, F, and G ...
Principal Investigator: ...Hmm, hmm (yes).
Rae: I wanna see if he's ... if he says number/letter ... See, doesn't have the letters. (The student mentioned (at that point) that the teacher used letters to tell him what to color in).
Son of a gun ... (TI2, lines 407-433).
However, there is no way of knowing whether it was the absence of letters that prevented Martin from finding the coordinate pairs. Rae again mentioned the differences in how the items were presented in the students' math program as opposed to how they were represented on the clinical interview protocol. For the money word item Rae felt
that Martin's issue with the question was that the item did not require exact change.
Nonetheless, the other students in the study did not have that challenge with the item and were even able to say how much was left over.

These comments notwithstanding, Rae's main revelation after viewing Martin's clinical interview was the fact that he had used mental math.

Rae: Basically he used mental math ... and that surprised me.
Principal Investigator: Ok. Umm ... and so the next question, preferred method was mental math.
Rae: Mental math 'cause you could see his little head going and his little mouth going and umm ... I was surprised by that (TI2, lines 568; 576-581).

This concluded Rae's comments about Martin’s clinical interview performance. An examination of his NJPASS results revealed that Martin's use of mental math was not restricted to the clinical interview. He had extended them to the NJPASS assessment as well. Details of Martin's NJPASS scores are described in the next section.

## Martin

NJPASS results.
Martin earned 18 out of 30 items correct on the corresponding NJPASS items and 22 out of 40 on the entire assessment. As a result, he was evaluated as being in the Basic level. Moving from these more general scores, the item response record highlights Martin's progress on the individual items in the various categories.

Number sense, operations, and properties.
In this category Martin got some corresponding items correct, and others
incorrect. He got two of the three mental math items correct which underscores Martin’s hitherto unrecognized use of this cognitive strategy. Martin was not successful on the multiple-choice coin item, the money word problem, the multi-digit word problem, and
the place value items on the NJPASS that matched the items on the clinical interview. However, he earned two of the three points on the open-ended item that was one of two items that corresponded with the coin item. He had better fortune with all the items in the next section.

## Measurement.

Martin got both of the corresponding items correct in this section. This included the open-ended item for which he earned the full three points. This was noteworthy as this item was an elapsed time problem. He had been rated as being at the beginning level and had done poorly on the elapsed time item in the clinical interview. However, on the NJPASS he was successful. This answer exemplifies the inconsistent nature of Martin's responses. It also indicates the different levels of challenge between the open-ended clinical interview items and the multiple-choice NJPASS items. In the next section, his performance fluctuated once again.

Spatial sense and geometry.
In this category Martin got the other item that evaluated fractions correct. He was not able to accurately complete the coordinate grid item and thus confirmed his difficulty in this area. Antithetically, he had no challenges with the symmetry item and as a result he earned the full three points for the open-ended item. His uneven performance continued in the next section.

Data analysis, probability, and discrete mathematics.
Martin experienced no challenges with the probability items. However, the bar graph items were as perplexing to him as those on the clinical interview, and on the

NJPASS there were no probes to help him. Additionally he only got one point for the open-ended item that matched the chart item. The same inconsistent pattern of performance was seen on the patterns and algebra section.

## Patterns and algebra.

He got the two corresponding items incorrect. These were the basic facts and measurement items. This was surprising, as he had done well on these concepts on the clinical interview. Although he got these items incorrect, he excelled on the symmetry open-ended item earning the entire three points.

In sum, Martin's work on both the clinical interview and the NJPASS showed a student whose skills were for the most part in the developing stage. He surprised his teacher by revealing that he had more strength in using the cognitive strategy of mental math. For certain items, his prior knowledge was not able to benefit from probes. He definitely fit the classification of inconsistent. Not only on the clinical interview but also on the NJPASS sometimes he understood a concept, while at other times he didn't. Or he did well on a concept on one assessment, but not on the other. Philip, the final student participant was also considered to be an inconsistent worker by his teacher. Further information about him from his teacher and his performances on the clinical interview and the NJPASS are described in the next sections.

## Philip

First teacher interview.
Philip attended E School. He was seven years and ten months old at the time of the study. His teacher, Mary, had taught for a total of 15.5 years, three of them having been spent in her current position in the second grade. She had a Masters degree in
education plus 30 additional credits. During her first teacher interview, Mary rated Philip as being at the developing level in all the conceptual categories. This was with the exception of place value and basic addition facts in which she rated him as working at the secure level. Mary stated that he used mental math in completing algorithms. He used his fingers, paper and pencil, and manipulatives. She felt that while he used written explanations, diagrams and pictures, he was best able to explain his thinking orally. Mary used the same sources as the other teachers to support her evaluations of Philip's mathematical skills and knowledge.

After stating her ratings, Mary added that the First Move chess program was very helpful to her pedagogy.

Mary: I think the chess is really helping me - let's hope, uh, because there's an opportunity for me to naturally give them vocabulary like ... such as ... you know ... horizontal, vertical, diagonal, quadrant. You know... doing the coordinates and things like that, so that has been really helpful for me as a teacher. I hope it carries over to them (TI1, lines 238-242).

This is again important as it shows that Mary also did not teach to the test. She used the word "naturally" to describe how the chess program facilitated her ability not only to expose the children to certain mathematical terminology, but also to allow them the means of becoming increasingly familiar with the words.

Mary's deep understanding and awareness of Philip's work and work habits were revealed through her subsequent comments. When she was asked to summarize Philip's mathematical abilities, Mary was concerned about his inconsistent performance and the length of time it took Philip to complete tasks.

Mary: I think ...He is very emotional. So I ... when I look through his work, there are days when he didn't do as well as other days, and I will pull him aside
or put him in a small group or something, and I know that he knows it. It also takes him a long time to settle in in the morning. He, uh, I think at this age they're trying to kinda navigate their friendships and all that, and I see that he is kinda frequently preoccupied with all of that. So it takes him a while to settle in and a lot of times I do do some math boxes in the morning with them and as a result it takes him a while to get that finished. He can do it. It just takes him a long time.
Principal Investigator: Hmm, hmm (yes).
Mary: He is a bright student. He needs to slow down, and I have many samples of that as well. And he responds well to praise, and working one on one, and in small groups (TI1, lines 252-264).

This brief excerpt is important because it reveals three of Philip's characteristics that affected his performance on the clinical interview and possibly on the NJPASS. The first attribute, his inconsistency, is reminiscent of Martin, and was evident during the clinical interview. The next attribute, the length of time Philip took to complete tasks was very noticeable on the clinical interview. Of all the students, including students like Kevin who were considered slower workers by their teachers, Philip took by far the longest to complete tasks on the clinical interview. The final attribute, working well one on one, could have contributed to the fact that he performed the best on the clinical interview of all the students.

Mary also indicated that Philip did not like to check his work. Again this is similar to Rae's comments about Martin, and Philip revealed this tendency during the clinical interview.

Mary: And he doesn't go back and check. I encourage them all to go back and check, but he really needs to be, you know, enco... very much encouraged to go back and check. I think he is just so happy to get it done that he just, you know (moves her hands to indicate just wanting to be finished with something) (TI1, 290-293).
Clinical Interviewer: Now I do have another question. I asked if you anticipated that he would do well on the NJPASS. If you had asked him those same
questions in a one to one setting, do you think he would have done better? About the same?
Mary: (Thinks for a while). My feeling is he would probably do better. That's just my gut feeling. I mean, I was looking through all of his work, and in a nutshell, I think he just makes careless errors. He is rushing through things or he is distracted or preoccupied about other things, but if I were to sit him down and rein him back in, he’s perfectly capable. Even more so... (TI1, lines 305-314). This may have been true in the classroom, but as will be seen during his clinical interview, Philip sometimes did not check his work despite probes designed to get him to do that. Prior to this section of the interview, the clinical interviewer had asked Mary if she had anticipated that Philip would do well on the NJPASS. This was after she had indicated that he was a bright student. Mary had responded in the affirmative. "I think he will do well. I think he will kinda like be in the middle. I don't think he will be at the high portion. I think he will do well" (TI1, lines 269-270). Given his inconsistency, his careless errors, and his reluctance to check his work, this placement of his scores as being good but not of the highest caliber is warranted. Like the other teachers, Mary clearly had positive but realistic expectations of her student. Her predictions were verified by Philip's performance during the clinical interview. Details about his work on this assessment are presented in the next section.

## Philip

## Clinical interview.

Philip was the only student who performed better on the clinical interview than on the corresponding NJPASS items. Another difference that was immediately apparent between Philip and his peers was that he did not use mental math for the first item. He counted on his fingers. His teacher, Mary, had indicated in the first interview that did use fingers to compute addition and subtraction algorithms. However, given the fact that all
of the other students had used mental math for this item, his usage of his fingers in this instance was striking. Nonetheless, Philip got the first part of the item correct. When asked to write a number sentence to show his answer, he started writing an actual sentence using words. That too was very different from the other students. All of them had understood the term "number sentence". A probe from the clinical interviewer in which she used the term "number model" clarified things for Philip, and he was able to complete the second part of the item correctly.

Clinical Interviewer: You don't have to write the words. No. Just, just the number model.
Philip: Oh (CI, lines 19-21)!
While the first item had presented Philip with a minor complication, the second item on the protocol provided no challenges for Philip. Unfortunately, this was not the case for the third item on elapsed time. After volunteering that he used to play tee-ball, Philip was unable to arrive at the correct solution despite probes from the clinical interviewer.

Clinical Interviewer: It lasts 45 minutes and it begins at 3:30. What time did your practice end? You can use the clock. What do you want to do?
Philip: (He is looking at the question, looking at the clock, but doesn't really manipulate it initially. Then he starts doing that). Um ... I'll ... say ... this time.
Clinical Interviewer: This time ... and what made you decide on that time?
Philip: Um ... I don't know.
Clinical Interviewer: Ok. How long is 45 minutes?
Philip: Three 15 minutes (I am not sure if this is what he said. He had his hand partially over his mouth so some of his words were inaudible).
Clinical Interviewer: Is that more than an hour or less than an hour?
Philip: More?
Clinical Interviewer: Ok. So what time would you write here then for what time practice ended? If it started at ... if practice started at 3:30, what time would you say it ended?
Philip: I'd say he end, um, at 10 ?

Clinical Interviewer: Ok. Would you write that (CI, lines 50-78)?
Despite his having received a rating of developing in this conceptual category, Philip could not answer the item correctly. His answer provided another example of the ineffectiveness of probes if the student's prior knowledge was inadequate. While Philip knew that forty-five minutes could be divided into three fifteen-minute segments, he didn't know that the time frame was less than an hour. This gap in his knowledge precluded his ability to answer the item correctly.

In contrast with his challenges on the elapsed time item, Philip successfully answered the pattern, coin, and probability items. He needed no probes for any of them and in this way surpassed the rating of developing for these categories. For the pattern and coin items he used mental math and confirmed Mary's statement that he did use that computational strategy. All of the students had taken a long time to complete the coin item. However, Philip took by far the longest time. The clinical interviewer and I were concerned as to whether he would be able to finish the remainder of the items in the allotted time. Additionally, compounding the issue, when he colored the first spinner, he also took an extremely long time. As a result, when he did the second spinner, we had to make some changes as to how he would respond.

Clinical Interviewer: Ok. You did a very nice job with that. Now I have a question about this one. You need to make this spinner also have the equal number of red, yellow, blue, and green, but it can't look like this one. Um, so you could show what pattern you'd like to use on this spinner by just putting an R for red with the red crayon, $y$ for yellow with the yellow crayon, $G$ for green with the green crayon, B for blue with the blue crayon. Show me what pattern you would use on that, on that one (CI, lines 158-163).

Even with this modification, Philip still took some time, but eventually completed it accurately. The length of time he took to complete these items substantiated Mary's assertion that Philip took a long time to complete tasks. In the classroom he had distractions, but there were none in the individualized setting of the clinical interview. This leads one to infer that Philip was an inherently slow worker who benefited from extra time when completing assignments.

Time was not an impediment with the bar graph item that came next. Philip's response to the first sub-question illustrated how probes can be effective in enabling students to answer correctly.

Clinical Interviewer: Using the graph can you tell who read twice as many books as Tyesha? ... Read twice as many books as Tyesha.
Philip: Well both are equal, so these two (pointing to Andrew's and Trent's).
Clinical Interviewer: Can you see where Tyesha is?
Philip: (He nods and puts finger on Tyesha's name).
Clinical Interviewer: Ok. Who read twice as many?
Philip: ...twice as many ... Andre?
Clinical Interviewer: Ok. And how did you decide that?
Philip: Um ... because Tyesha's is up here, is right here, and Andre's is right here (pointing to their bars) more than Tyesha's.
Clinical Interviewer: Ok. So can you write that name right here - that's Andrew. Philip: Oh.
Clinical Interviewer: That's ok.
Philip: (He writes name) (CI, lines 179-204).
In this instance, a probe allowed Philip to make the correct choice. He had no difficulty with the next two sub-questions, but he again needed a probe from the clinical interviewer to arrive at the correct solution.

Clinical Interviewer: And the last question there is how many less books did Jamal read than Trent?
Philip: He read four books.
Clinical Interviewer: Four ... four less books.
Philip: Yup.

Clinical Interviewer: Is that what it says?
Philip: Hmm, right here.
Clinical Interviewer: Jamal read four books. How many less books than Trent did he read?
Philip: (He calculates using number grid). Less than Trent? Umm, umm, he read 3.

Clinical Interviewer: He read three fewer?
Philip: Yeah (CI, lines 226-246).
Philip was able to use the probe to rectify the error he had made in reading the graph.
The mistake he made while completing the symmetry was more indicative of his inconsistency in performance rather than of misinterpretation of the language of the question.

For the symmetry item, Philip correctly identified the three figures that were symmetrical. Unfortunately, he also identified one that wasn't. This would match his rating of developing in this category. For the next item on coordinate grids, Philip was the only student who got the pairs correct. His only lapse was writing them as fractions as he did despite a probe by the clinical interviewer.

Clinical Interviewer: Ok. So what would you write to show me that location? Philip: Umm ... one fourth?
Clinical Interviewer: Ok. One, four. Write one, four. Ok. Would you write one, four like that or would you write it side by side?
Philip: Ah, side by side. No, like this.
Clinical Interviewer: You think that's the way you'd write it? Ok (CI, lines 316-327).

Regardless of the probes of the clinical interviewer, Philip stuck with his original representation of the coordinates as fractions. By doing this, he again validated his teacher's rating of developing. Although he knew what to do and found the correct answer, he got confused about how to write them.

Probes again proved ineffectual in altering his response to the money word problem. In fact his initial response was very similar to some of Oliver's more unexpected replies.

Clinical Interviewer: At the bake sale, brownies cost 50 cents, the chocolate chip cookies cost 25 cents, oatmeal raisin cookies cost 20 cents, and cupcakes cost 35 cents. Sam has one dollar. What can he buy?
Philip: One dollar ... he can ... he has one dollar. He can't buy anything because he only has one dollar.
Clinical Interviewer: So there's nothing that costs less than a dollar? Let's read it again. At the bake sale the brownies cost 50 cents ...
Philip: Oh, 20 ?
Clinical Interviewer: The brownies cost 50 cents, the chocolate chip cookies cost 25 cents, the oatmeal raisin cookie costs 20 cents, and the cupcake costs 35 cents. They're all sold separately. What can he buy?
Philip: Umm...
Clinical Interviewer: Can he buy anything?
Philip: The raisin cookie.
Clinical Interviewer: Ok. He could buy a raisin cookie. Could you circle that? Does he have any money left? Can he buy something else?
Philip: Umm ... no (CI, lines 363-392).
Since he had been evaluated as developing in the conceptual category of money, Philip's inability to answer this question correctly despite the probes is surprising. One possible reason could be Philip's unwillingness to check for correctness. The last two probes were an invitation to look at the question again and re-examine the numbers. Philip did not avail himself of this opportunity. He also did not approach the multi-digit item in the same way as his peers had before him.

After the item had been read to him, Philip looked at the item and reread it.
However, unlike the other students, he made no attempt to start trying to solve it. Finally, the clinical interviewer offered him some options.

Clinical Interviewer: Would you like to skip that one and come back to it later if we have time?

Philip: Ok.
Clinical Interviewer: It's up to you. You can keep working on it or we can come back later.
Philip: Umm ... come back later (CI, lines 403-409).
However, the fact that he made no attempt to solve the problem could be an indication that he too had misunderstood the directions. Writing a number sentence would be in keeping with his rating of developing in this area. The clinical interviewer had helped him to remember what constituted a number sentence. Thus, it could be inferred that if he had realized that was what was needed for this item, he would have been able to provide one.

For the chart item, Philip also missed the key terms of the directions like the peers who had gone before him. Like them, he focused on getting nine. In response to the clinical interviewer asking him to explain how he had chosen the rides he had circled, he answered, " $2+4=6+3=9$ " (CI, line 439). Instructions were not an issue for Philip on the place value item and he answered both parts without difficulty or the need for probes. For the final question, things were not quite as simple for Philip. He erred from the very first by shading in eight instead of nine rectangles. Therefore, his answer was understandably incorrect.

Clinical Interviewer: Ok. Now, how many are left un-shaded?
Philip: Umm ... ok ... 7.
Clinical Interviewer: Ok, so you shaded nine and seven are left?
Philip: Yeah.
Clinical Interviewer: Ok. Want to check?
Philip: Umm ... (He looks like he’s counting, but he doesn't change his incorrect answer).
Clinical Interviewer: So could you write down how many are left and could you write it as a fraction?
Philip: Fraction??? Like, like, like ...this??

Clinical Interviewer: No. A fraction is rather like this. (She writes an example on a piece of paper).
Philip: Oh! A fraction ... oh ... (CI, lines 495-517).
The excerpt above is illustrative of Philip's reluctance to checking his work, and completely supports Mary's evaluation of this aspect of his work ethic. Notwithstanding the clinical interviewer's direct invitation to review his working, Philip does not do so. Therefore, unlike the other students, Philip was the only one who got the first part of the fraction item incorrect. This result was directly attributable to his reluctance to look over his work for mistakes. In addition to exemplifying Philip's unwillingness to ascertain that his work was error free, this quotation also serves another function. It explains Philip's confusion on the coordinate grid item, when although he could identify all the pairs and put the numbers in the correct order, he wrote them as fractions. In answering the fraction item, when the clinical interviewer asked him to write his answer as a fraction, his response indicates that he had forgotten what that entailed. This would explain his confusion about representing the coordinate pairs appropriately despite the efforts of the clinical interviewer. His misconceptions or lack of sufficient prior knowledge led him to pick the incorrect visual representation of the coordinate pairs.

In sum, Philip's performance on the NJPASS essentially verified Mary's comments in the first interview. He had good skills that were overall developing, although in some areas like patterns, money, and probability, he worked more at a secure level. Her comments about Philip's inconsistency and his reluctance to check his work were also supported in his work on the clinical interview. While Philip used mental math, he used them to a much lesser degree than the other students. Additionally,
although many of the students were slow workers, Philip was unquestionably the slowest. If the clinical interviewer had not made modifications for the probability item and if he had not deferred completion of the multi-digit subtraction item, Philip would not have been able to complete the protocol in a time that would have been comparable to his peers. Mary had indicated during the first interview that Philip had used every bit of time that had been allotted on the NJPASS so in this area his behavior is consistent. While Philip experienced the same directional miscues for the chart and multi-digit items, he was the exception on the coordinate grid item as he was the only one who clearly knew that two numbers were needed, and who knew the correct order for all the pairs. Philip's performance on the clinical interview matched Mary's evaluations about his knowledge and work habits. His scores on the NJPASS also corroborated her belief that he would score well but not in the highest levels.

## Philip

NJPASS results.
Philip got 19 of the 30 items that corresponded with the clinical interview items correct. He got 24 out of the 40 items on the actual assessment correct. The latter score earned him the level of Proficient on the complete NJPASS assessment. Philip's performance on the individual items on the assessment was documented in the item response record.

## NJPASS Item Response Record

As I have previously mentioned, the Record has five conceptual categories. These are number sense, operations, and properties; measurement; spatial sense and geometry; data analysis, probability, and discrete mathematics; and patterns and algebra.

The 40 items are separated into these classifications. Each section on the NJPASS ends with an open-ended question. Once again, the focus of the analysis will be on items that correspond to those on the clinical interview.

Number sense, operations, and properties.
Philip had mixed success in this category. Of the six corresponding multiplechoice items on the NJPASS he got three correct and three incorrect. The three correct corresponding items matched the money word problem, the coin item, and the place value item. Except for the money word problem, Philip had been successful on these items as well in the clinical interview. The incorrect responses matched the fraction, money word problem, and multi-digit subtraction items on the clinical interview. He had gotten these incorrect on that assessment. For the open-ended item that assessed his knowledge of money, Philip earned only one of three possible points. He did much better on the next category.

Measurement.
On the multiple-choice item that matched the measurement item, Philip was successful. This was similar to his answer on the clinical interview. For the open-ended item that evaluated elapsed time, Philip earned two of the three points. This was considerably better than his performance on the corresponding item in the clinical interview. Philip's answers on the spatial sense and geometry category had an unexpected twist given his performance on the clinical interview.

## Spatial sense and geometry.

He got the second corresponding fraction item correct. Since he had gotten the first item incorrect, this split would be consistent with someone working at a developing level in the concept. The response that was unanticipated was that on the multiple-choice coordinate grid item, Philip was not accurate. He got the more challenging open-ended clinical interview item correct, yet missed the simpler multiple-choice item on the NJPASS. This is clearly a reflection of Philip's inconsistent work habits. For the openended item on symmetry, he was able to earn the full three points for the first time on the NJPASS. That achievement did not carry over to the data analysis category.

Data analysis, probability, and discrete mathematics.
Philip got the two multiple-choice items on probability correct. His response for this item on the clinical interview was similarly accurate. He got both of the bar graph items incorrect. This was not entirely surprising. Without the benefit of probes on the clinical interview he would have gotten some of the bar graph sub-questions incorrect. In contrast with the clinical interview, Philip got two of the three points for the open-ended item that corresponded to the chart item. He was more successful on the open-ended item in the final section of the NJPASS.

## Patterns and algebra.

Philip got the basic fact multiple-choice item correct, but, in concert with his performance on the clinical interview, did not get the corresponding elapsed time item correct. He cemented his mastery of patterns by getting the total three points for the item that evaluated this concept.

To conclude, Philip's performance on the NJPASS was mixed as befitting a student who was still developing his mathematical skills. He was able to get the full points on the open-ended items in only two categories - spatial sense and geometry and patterns and algebra. Philip used mental math less than any of the other students, and he was not successful on these items in the NJPASS as well. Probes proved to be instrumental in helping him get the bar graph items correct on the clinical interview while he missed them on the NJPASS. He was unable to solve the elapsed time items on both assessments. While he got the multi-digit item incorrect, he almost got the chart item correct. So on the NJPASS, the directions for that item were less problematic for him than it had been on the clinical interview. Philip's work on the NJPASS earned him the middle level of proficient and his answers as outlined by the item response record explained this score. His teacher, Mary, had predicted Philip’s achievement. Her comments as she watched his clinical interview are shared in the following section.

## Philip

## Post teacher interview.

Mary clarified that Philip had been ready to write an actual number story for the first item. She found the wording interesting and reiterated that the question should have said number sentence. When the clinical interview reached the coin item, the issue of time, in particular wait time arose. I pointed out that particularly during this item, and especially with this student, the clinical interviewer and I had to wait patiently for the students to respond.

Principal Investigator: Every time we would think he would be ready to say something ...
Mary: Ok. Yeah... is he going to say something?

Principal Investigator: ... he would start again. It was a good thing we learned to just remain silent and just wait and watch them. This was very good for us in terms of wait time.
Mary: That's important in teaching too ... in the classroom (TI2, lines 143-150).
The issue of time became a theme in the study due in part to the fact that all the students took longer than had been anticipated on this item. Watching as he worked on the item, Mary confirmed that Philip was a slow worker.

Principal Investigator: Does he normally take this much time?
Mary: He does.
Principal Investigator: Good. Ok. Good. This is something important for me to know.
Mary: He does. He gets it right, but it takes him a long time (TI2, lines 201-207). The clinical interviews were scheduled to last approximately 45 minutes. This was with the belief that they would last no longer than 30 minutes. The fact that Philip took as long as he needed to complete the items on the clinical interview contributed to his success on this assessment. Despite the inordinate amount of time that he took on the coin item for example, if he had not finished within the allotted time, it would have been possible to extend or reschedule the clinical interview. This scenario would not have been possible with the NJPASS.

When Mary heard that most of the students took longer than expected on the coin and probability items, her response reflected her interest in sharing information and discussing ideas with colleagues on her grade level. "You kept seeing that? That's good for us to talk about as a grade level" (TI2, line 287). This concern with pedagogy foreshadows her response to Philip's performance on the clinical interview. Mary used this opportunity to share some positive opinions about Philip. "He will listen to
instructions and he'll do what you ask...you know...what the teacher is asking. He wants to learn" (TI2, lines 352-353).

The clinical interview tape had progressed to the bar graph items. When Mary saw that Philip needed probes on two of the sub-questions in order to arrive at the correct answers, she wondered how the students would perform on a test without this type of intervention. "But see, I wonder what they'll ... when they have to do it on a test? Without the prompting" (TI2, lines 421-422)? In Philip’s case his NJPASS scores showed that he did not do as well on that assessment, and the lack of probes there could have contributed to those outcomes.

For the coordinate grid item, Mary was distressed that Philip had stuck to writing the coordinates like fractions despite the probes to get him to see his error. She felt better, however, upon noting Philip's strategy in solving this problem. It was possibly the reason why he got the item completely correct in contrast with the other students.

Mary: What did he do?
Principal Investigator: He went like ding ... He went like so and so.
(Demonstrates on the paper). He did ...
Mary: He drew a line joining them together.
Principal Investigator: Yes.
Mary: Ah, so he wrote it like a fraction?
Principal Investigator: Yeah.
Mary: So he stuck to writing it as/like a fraction?
Principal Investigator: Hmm, hmm (yes).
Mary: That's so ... (looks very distressed).
Principal Investigator: What I found though, for each one, he drew the two pieces...
Mary: ...lines like meeting
Principal Investigator: Yes. Sorry, I can't explain. But that's what he did for each one. So he wrote 2, 9
Mary: 2, 9 looking ...two ninths?
Principal Investigator: I don't know if he ... I guess it looked like that ... I
couldn't see from where I was sitting.

Mary: Ok
Principal Investigator: But he did get that it was 2, 9.
Mary: He knew it was the bottom and ... (points to the side) (TI2, lines 471-504).
Philip's strategy of drawing the line connecting the numbers proved that unlike his peers in the study, he recognized that the number on the bottom was the first number in the pair. Mary noted this technique and stated, "He is a smart boy" (TI2, line 514), once again indicating her positive expectations for this student. The fact that she noted the strategy is another example of her interest in the pedagogical ideas that Philip's clinical interview could provide for her.

Also in her recap of the clinical interview, Mary noted his use of the number grid. Although Philip used less mental math than the other students during the clinical interview, he apparently did not use the number grid much in the classroom.

Mary: Well, he used his fingers, and he used the grids. And I am glad to see that he was using the grid. I don't know how you solve this problem. In the classroom, I don't think he's as open to use the grids.
Principal Investigator: Ok.
Mary: Because maybe he feels his peers aren’t using the grids, so I don’t know how to solve that problem (TI2, lines 693-700).

She felt that despite his use of the grid, his main computational strategy had been paper and pencil. The clinical interviews were conducted after the NJPASS, but Mary was still focusing on exposing her students to the different ways in which familiar concepts could be presented.

In conclusion, Mary confirmed that Philip was a very slow worker, and was a student who wanted to learn. Insufficient prior knowledge sometimes prevented Philip from utilizing probes. However, when he was able to use them, they contributed to his success on the clinical interview. His record of mixed success on both the clinical
interview and the NJPASS are indicative of a student who was on the cusp of acquiring solid mathematical skills and knowledge.

Just as Philip was unique among the students in terms of being the only student who performed better on the clinical interview than on the NJPASS, his teacher, Mary, also viewed the information she gathered from the clinical interview through a different lens. She was interested not only in Philip's performance on the clinical interview for what it revealed about his mathematical understanding. Mary also wanted to use whatever she learned as an opportunity to enhance the mathematical knowledge of the students in her class and on the grade level as a whole. While her focus may have been unique, the other teachers also had positive reactions to their students' performances.

In sum, the seven student participants displayed disparate ability levels and earned different scores on the clinical interviews and the NJPASS. Six of the seven students performed better on the NJPASS standardized test than on the matching clinical interview items. Thus, the answer to the first research question is that the clinical interview as a complementary assessment did not enable the students to show that they knew more mathematic concepts than they were able to reveal on the NJPASS. On the entire mathematics section of the NJPASS, one student achieved the advanced level, four students attained the level of proficient, and two achieved the basic level. In addition, the teachers generally had accurate evaluations of their students' academic abilities in the various concepts. The impact on the teachers of the information that they had gathered from the clinical interviews is presented in the next section.

## Results

## Research question \# 2.

The second research question was: What might teachers learn about their students' mathematical ability, and how might this knowledge affect their views of those students? The table below presents the results for this part of the study's inquiry.
Table10

Summary Of Changes In Teachers' Attitudes

TEACHER

LIANE Kevin and Sam had both performed
DANIELLE Oliver and David had both performed

RAE

MARY
as she had expected
DAHLIA Jake had performed as expected BECCA Jake had performed as expected

STUDENT'SPERFORMANCE AS ANTICIPATED

Philip had performed as she had expected

STUDENT'S PERFORMANCE UNANTICIPATED

No change
No change
No change
No change
Martin used more Her opinion of mental math than expected Martin's math ability improved

Used information to make pedagogical changes in her classroom. The teacher also planned to present her ideas for pedagogical changes to her peers on the grade level.

Excerpts from the post teacher interviews provide supporting details for these findings. They are organized to match the order in which the teachers were presented previously.

## Liane

Kevin.
Liane felt that Kevin’s performance during the clinical interview did not provide any surprises for her with reference to his mathematical skills. His responses were what she had anticipated. Liane noted that his preferred cognitive strategy during the clinical interview was mental math. She thought that this elevated the level of his computation as using mental math was more challenging. "Which is good. Great that he could do that. That's a higher level" (TI2, line 678). Liane believed that his mental computing verified her evaluation in the first teacher interview. She felt that Kevin had done really well during the clinical interview. In her opinion, the one to one format of the assessment enhanced his performance. She ended the post interview by reiterating that she had not learned anything new about Kevin's mathematical abilities as a result of watching the clinical interview. "Uh, I didn’t have any surprises from Kevin" (TI2, line 688). Her opinion about Sam did not change either.

## Liane

Sam.
In the second teacher interview, after watching Sam's performance during the clinical interview, Liane said that he had done as she had thought he would. She felt that he didn't have a preferred cognitive strategy during the interview and had used a combination of mental math and paper and pencil. There were a few items that she had thought he would not have answered accurately to which he had responded correctly. "I
knew he would have a little difficulty with ... actually one or two that I thought he might not get that he did, but pretty much, I'll just put yes" (TI2, lines 1335-1337). Liane felt that overall his responses were as she had anticipated and that her prior evaluations had been supported. "He performed pretty much the way I thought he would, so..." (TI2, lines 1288-1289). Therefore, her views of his mathematical ability had remained the same. Liane was very proud of both Kevin's and Sam's performances during their clinical interviews. She found the clinical interviews informative. "It was very interesting" (TI2, line 1341). Danielle didn’t verbalize her curiosity, but it was evident from the intentness with which she watched Oliver's clinical interview.

## Danielle

Oliver.
As she watched Oliver solve the items, Danielle recognized that he definitely used mental math and some paper and pencil. She felt that his body language supported his performance. "I would guess because he did so much mental math he was really thinking so he would have his hand on his head, and he would, you know ...' (TI2, lines 684-685). When asked if there had been any surprises for her, Danielle responded, " Just that he wasn't more enthusiastic and he seemed a little draggy. Is that the right terminology" (TI2, lines 692-693)? However, in the following excerpt, there seems to have been two elements of Oliver's performance on the clinical interview that Danielle had not expected to see.

Principal Investigator: Were the student's responses as you anticipated? Danielle: For the most part, yes. I, I ... he's pretty confident with his math.

Principal Investigator: Umm ... so how would you summarize what you learned from the clinical interview about your student's mathematical knowledge? What would you say?
Danielle: That he uses a lot more mental math than I thought he did.
Principal Investigator: Ok.
Danielle: I knew he could explain a lot orally. He likes to do that. So that wasn't a surprise, but the, the, as much mental math as he used was a surprise.
Principal Investigator: Ok. So he used a lot more mental math than you would have...
Danielle: Right. Than I guess I realized that he does. And that he had to have so many directions reworded for him. That caught me a little bit by surprise for him. 'Cause you know...I think what he did was to ... he actually wanted to be sure of himself too. He asked a few times what was it, I... what was the number. He asked to be sure a few times. He asked for it to be repeated not that he didn't understand also, but that he just wanted to make sure before he did it. I felt like he did that a few times too (TI2, lines 700-724).

After sharing these insights into Oliver's mathematical ability, namely his extensive use of mental math as well as needing many directions reworded for him, Danielle indicated that his responses were as she had anticipated and reflected his confidence in math. Based on the clinical interview, Danielle stated that her view of Oliver's mathematical knowledge had remained the same. "I would say pretty much the same. I knew he was a confident math student, and he seemed to show that today" (TI2, lines 729-733). At the end of Oliver’s DVD, Danielle was just as keen to view David’s clinical interview.

## Danielle

## David.

During the second teacher interview, when asked whether David's responses had been as she had expected, Danielle answered in the affirmative. She stated that she had found nothing surprising in his performance. In an unconscious contradiction, she
admitted that she found his ability to verbalize some of his responses unanticipated. "I was surprised at some of them he did do ... that he was able to verbalize" (TI2, line 280). A little later in the interview, there was some disparity between her interpretation of his body language and mine. Based on my professional experience, I thought that when David looked up and away from his work, he was actually computing mentally. This belief was supported by the fact that after these pauses, David would usually jot something down or arrive at a solution. Danielle, however, viewed these incidents during the clinical interview as a demonstration of David's lack of concentration.

Principal Investigator: Body language - anything you noticed about his body language?
Danielle: Umm...sometimes he looked a bit distracted...he would look up...away...
Principal Investigator: Ok, ok. Could he have been visualizing at that point?
Danielle: Well, uh...
Principal Investigator: I don't know. Because sometimes, you know, when some kids look off, they're kind of seeing it in their mind's eye.
Danielle: Right (TI2, lines 284-295).

This was the end of the discussion on David's use of mental math until the very end of the post interview when Danielle agreed that he had used some mental math as well as paper and pencil in his computations. Her response to my statement that David had used some mental math and some paper and pencil was, "Yeah, he did some of each. And then I think when he was explaining he had to kind of go step by step to continue to explain I noticed. You know, but I think he had both" (TI2, lines 350-351). Danielle believed that her prior evaluations were supported.

Principal Investigator: So, were your prior evaluations supported or negated?
Danielle: I think supported (TI2, lines 302-304).

In summarizing what she had learned about David's mathematical knowledge, she said that he had benefited from repetition and rewording, and that he was inconsistent.

Principal Investigator: Ok. Uh, have your views about your student's mathematical knowledge changed or remained the same after viewing his performance in the clinical interview?
Danielle: Umm ... remained the same. Yeah (TI2, lines 331-334).
Thus, at the conclusion of her post teacher interview, Danielle maintained that her opinions of David's mathematical capabilities remained unchanged after viewing his clinical interview. Dahlia, Jake’s homeroom teacher also felt that her evaluation of his mathematical ability had been confirmed.

## Dahlia

Jake.

After viewing Jake's clinical interview, Dahlia noted that Jake had used a combination of paper and pencil as well as mental math as his preferred computational strategies. Also, she noted that he explained himself orally.

Principal Investigator: In terms of his knowledge of the concepts were there any surprises? Or his performance, or how he, you know, answered the questions?
Dahlia: No, I don't think that for him ... No.
Principal Investigator: So I guess the student's responses were as you anticipated then.
Dahlia: Yes. I figured he would stay focused and you know ...
Principal Investigator: And his knowledge was kind of where you figured it would be?
Dahlia: Yes, yes. So it's not like you know, all of a sudden, say for instance, he's not strong in counting money and all of a sudden I saw him count ... you know, I didn't ... I didn't say well, wait, he doesn't do that for me. So I didn't see anything (TI2, lines 785-798).

As this excerpt shows Dahlia believed that her prior evaluations had been supported. Dahlia finished the post interview by noting that she could not exclusively
work one on one with Jake in the classroom. She was therefore very pleased at being given the opportunity to see the strategies and processes he used during his problem solving.

Dahlia: It's just nice to see like ... in a class of $19 \ldots$ not being able to totally work one on one to see how he goes about trying to solve ... how he approaches the problem ... the process... we have opportunities, but limited ... when I pull for small group, but this was nice to see ... how he's going around, moving the coins, you know... charting on the, you know, clock (TI2, lines 812-813; 825826; 830-831).

Becca was also happy with Jake's performance on the clinical interview.

## Весса

Jake.
During the post teacher interview, in sharing what had been surprising for her as she watched Jake's clinical interview, Becca focused on the speed at which he answered some of the items on the protocol.

Becca: At the speed at which he attempted that subtraction example, which makes me think that the peace and quiet, not looking around and seeing everybody picking up their pencil first made a huge difference. Umm...that class especially has a lot of kids that whisper or say answers out loud, and you see his shoulders go down because he didn't even get to try it (TI2, lines 591-595).

In summarizing what she had learned from Jake's clinical interview, Becca focused on the areas in which he had experienced challenges. Her detailed analysis supported her beliefs that Jake had performed as she had anticipated on the clinical interviews.

Principal Investigator: How would you summarize what you learned from the clinical interview about your student's mathematical knowledge?
Becca: Well, we know it's in there. I know he needs time ...
Principal Investigator: It's there, needs time ...
Becca: And definitely did better in quieter surrounding ...

Principal Investigator: Hmm, hmm ... (yes).
Becca: Umm ... did show me that the more complicated the diagrams, he has trouble visually, so reading that graph, he was asked to do two different things... that was a problem. Measurement is weak. That is a visual.
Principal Investigator: Problem (repeating words as writes them down). Measurement weak, and that's also visual.
Becca: And the fractions, and fractions, that was developing. That was difficult, that was difficult for him in class.
Principal Investigator: Ok. All-right. Umm ... have your views about your student's mathematical knowledge changed or remained the same after viewing his performance in the clinical interview?
Becca: Uh, remained the same. He is a determined student. But it showed him even better in a quiet setting that he did even better.
Principal Investigator: Ok. All-right. Thank you very much!
Becca: And he didn't appear nervous ...
Principal Investigator: No, no.
Becca: He didn't do the erasing ... some days I see ... with the fractions there was a lot more erasing.
Principal Investigator: Good.
Becca: He looked confident and willing. He is a great student. I am glad this portrayed him as a good student (TI2, 621-661).

Consequently, her views of Jake's mathematical knowledge remained the same.
Also unchanged were her positive views of this student and his abilities. Unlike Becca, Rae's views about Martin's mathematical abilities were altered due to the information she learned from watching his clinical interview.

Rae

## Martin.

During the post teacher interview Rae was asked to reflect on Martin's computational strategies. She said that he had primarily used mental math. This was very surprising to her. Rae stated that he had more strength or confidence in that area than she had thought. She wondered if this was because he didn't have a number grid.
(All students had access to the number grid during the clinical interview).

Rae: Only that he was using mental math more than I thought he was, and I don't really know if that's because he didn't have the number grid, because the number grid is right on his desk. I'll have to observe today when we do math to see, and I'm very curious to do that now (TI2, lines 597-600).

She said that he had used primarily oral explanations. Although she was surprised
at how much mental math he had used, she wasn't surprised he had done as well as he had because when he took his time, checked his work, and used the strategies that she had taught him, he did do well. She felt that her prior evaluations were supported. "My prior evaluations were supported. Yes" (TI2, line 627).

In describing what she had gained from Martin's clinical interview, Rae shared that she felt that her class in general would have struggled with the same items that Martin had found challenging.

Principal Investigator: How would you summarize what you have learned from the clinical interview about your student's mathematical knowledge?
Rae: Well, again it's gonna go back to, umm ... I feel he did well, umm...the problems, uh, the problems that he had difficulty with (writing as she speaks) ... I write slowly ...
Principal Investigator: No, that’s fine.
Rae: The problems that he had difficulty with as I do feel a good part of my class
Principal Investigator: Would have had the same problems
Rae: Problems
Principal Investigator: Ah.
Rae: I do feel a good part of my class, umm, would also have difficulty because they were in a different format
Principal Investigator: Thank you
Rae: Especially the graph one. Honestly, probably only about three of the kids in my class would have gotten it (TI2, lines 642-664).

With reference to her views of his mathematical knowledge, she felt that they had changed due to his extensive use of mental math. They had also changed because she noticed that with some additional prompting he seemed to do much better.

Rae: And also from the point of the mental math, however, also, umm... with some additional prompting, and strictly prompting, not given answers, he seems to do much better (TI2, lines 682-684).

In the earlier segment of her post-teacher interview, Mary had wondered about the effect on the students if they had to do without the prompting that had helped them to get items correct. Nonetheless, her primary focus remained on identifying pedagogical practices that she could employ in her classroom and share with colleagues.

## Mary

## Philip.

Commenting on Philip's performance in the clinical interview, Mary felt that her evaluations had been supported. Philip was doing as she had anticipated. Her views of Philip's mathematical knowledge had not changed. She still believed he was very capable, but some of the strategies she would use or do in the classroom would change. Mary's interest in pedagogical ideas was evident when she asked for a copy of the questions on the protocol so that she could review them with her class. She planned to review the items with her class in conjunction with the notes that she had taken during the interview. These included teaching the students about money, paying more attention to the language used in problems, and providing enough time and tools.

Principal Investigator: How would you summarize what you learned from the clinical interview about your student's mathematical knowledge? He ... you know ... I think you said it. He's doing as you anticipated, but there are certain areas that you could work on.
Mary: Yes. Exactly. Money, just wording things too, like with the number sentence and number story. You really have to be care... and providing them with that time, just like you said ... ok, as we know ... I have to get ready for this, I have to do this and that ... Just providing them the time.
Principal Investigator: Ok. So wording, providing time ...
Mary: Yup ... Providing him the tools ... you know ... encouraging them ... everyone learns differently and it's ok to use those tools (TI2, lines 770-782).

When Mary was asked about what surprised her about Philip's mathematical knowledge, she identified his challenges with some of the items dealing with money. "The coins. I know that they have a tough time with coins, but that was eye opening because there were two questions about the coins he didn't ... the second one he said he couldn't buy anything with the 20 cents. Right" (TI2, lines 729-731). Mary continued her focus on pedagogical issues in her discussion of the bar graph items.

Mary: The bar graph. We wanted to look at that too, because we want to see how many... I don't really think they show them a horizontal bar graph a lot in our math program and that is what they showed them on the test.
Principal Investigator: That’s why I did it horizontally.
Mary: Yes. So I was showing them some samples that were horizontal, but usually in the math program that we use, it's usually (gestures up and down). Principal Investigator: Vertical
Mary: Yes (TI2, lines 740-749).
Mary ended the post teacher interview by reiterating that Philip was a capable student, but that she would change the strategies she would use or the things she would do in her classroom. This decision was in response to the difficulties that her student encountered with some of the items on the clinical interview. While individual students experienced challenges with different items on the protocol, as a group, certain items proved to be particularly difficult.

To conclude, most of the teachers felt that they had not learned anything additional about their students from the clinical interviews that substantively changed their original opinions of their students' abilities in the concepts that were examined. One of the teachers changed her opinion of her student's mental math capabilities and
another made pedagogical changes based on the information garnered from her student's clinical interview.

The themes presented below provide further details about the boys' performances on both assessments. They also reflect cross analyses of segments of the data from the clinical interview and the NJPASS and provide another perspective of the study's findings. The students' varied efforts on these questions were one of the many themes that emerged from analysis of the data. In describing the topics, I will first address the ones that arose from the clinical interviews followed by those that were gathered from the teacher interviews.

Common Trends or Patterns Among the Students

## Themes from the Clinical Interview Findings

Challenging items.
A factor that all the boys had in common was the nature of their responses to certain items on the clinical interview. As was evident from their clinical interviews, four items in particular were particularly challenging for most of the students. The items were number nine - the coordinate grid problem; number eleven - the multi-digit word problem; number twelve - the chart; and number fourteen - the fraction problem. While a couple of the boys got numbers nine and fourteen correct, none of the students got items eleven or twelve correct on the interview. Since the items on the clinical interview were primarily open-ended, while those on the NJPASS were predominantly, but not exclusively, multiple-choice, I wanted to examine whether this difference affected the boys' performances on both assessments. Further analysis ascertained how the students
performed on the items on the NJPASS that corresponded with the items on the clinical interview that all or almost all of the students found challenging. The table below outlines the students' scores on these challenging items on both the clinical interview and the NJPASS.

Table 11

Comparison of Challenging Items

| Student | Grid/ <br> OE | Grid/ <br> MC | Multi/ <br> OE | Multi/ <br> MC | Chart/ <br> OE | Ch/NJ/ <br> OE | Fract/ <br> OE | Fract/ <br> MC |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| KEVIN | NO | YES | NO | YES | NO | $0 / 3$ | $2 / 2$ | $1 / 2$ |
| SAM | $4 / 6$ | YES | NO | YES | NO | $1 / 3$ | $1 / 2$ | $1 / 2$ |
| OLIVER | NO | YES | NO | YES | NO | $3 / 3$ | $2 / 2$ | $2 / 2$ |
| DAVID | NO | YES | NO | YES | NO | $3 / 3$ | $1 / 2$ | $2 / 2$ |
| JAKE | NO | YES | NO | NO | NO | $1 / 3$ | $1 / 2$ | $2 / 2$ |
| MARTIN | NO | NO | NO | NO | NO | $1 / 3$ | $1 / 2$ | $2 / 2$ |
| PHILIP | YES - <br> $6 / 6 ~$ | NO | NO | NO | NO | $2 / 3$ | $0 / 2$ | $1 / 2$ |

As the table shows, most of the students found the coordinate grid item easier as a single multiple-choice item rather than as a six point open-ended item. The multiplechoice format appeared to have helped four of the students to get past the directional miscue that occurred with the clinical interview item that had prevented any of the seven student participants from solving the multi-digit item correctly. Despite the fact that the chart item on the NJPASS was also open-ended, the students, except for Kevin, all did better on the NJPASS item. The fraction item was fairly equal in terms of performance between the clinical interview and the NJPASS. In general, judging from the students' scores, the open-ended questions on the NJPASS appeared to be less challenging for the boys. The table below summarizes information about the students' performances on the
challenging items on the clinical interview and the corresponding open-ended items on the NJPASS.

## Table 12

## Summary of Correct NJPASS Open-ended Items

| NAME | THREE POINTS | TWO POINTS | ONE POINT | ZEROPOINTS |
| :--- | :---: | :---: | :---: | :---: |
| KEVIN | 4 |  |  | 1 |
| SAM | 2 | 2 | 1 |  |
| OLIVER | 4 | 1 |  | 1 |
| DAVID | 2 | 1 | 1 | 1 |
| JAKE | 2 | 2 |  |  |
| MARTIN | 3 | 1 | 1 |  |
| PHILIP | 2 | 2 | 1 |  |

As the table shows, all of the students were able to get the full points for at least two open-ended items. Most were also able to score at least one point on the items. In sum, although they had found these four items challenging on the clinical interview, the students found the corresponding multiple-choice and open-ended NJPASS items easier to solve successfully. In addition to the issue of challenging items other themes emerged from analysis of the data from both the clinical and teacher interviews. For example, as seen in the analysis of the students' performances on the challenging items, misunderstanding or incorrectly processing the language of the questions proved to be problematic for the students. Thus, the issue of language emerged during analysis as one of the themes of the study.

## Language.

The issue of language emerged as a common element in all the students’ responses to varying degrees. Language was particularly relevant in the area of
directions. All the students missed the key words "as many different rides as possible" (Appendix F, pp. 323-328) in the chart item, and so got it incorrect. On the multi-digit item, they all did not register that only a number model was required. In addition, directions sometimes had to be repeated or clarified for some of the students in order for them to comprehend what was required of them. For example, as I have described, probes from the clinical interviewer enabled Oliver to understand what was required of him for the place value item. It was his lack of understanding of the directions that created his difficulty with completing the second part of the problem.

Other students were reluctant to explain their thinking or working as they solved the items when asked by the clinical interviewer despite having arrived at the correct answers. Although he had arrived at the correct total on the first part of the coin item, Sam did not share his working initially, and only did so reluctantly. For the spinner item, he did not respond at all to clinical interviewer's attempt to understand the strategies he used to solve the problem.

Some of the students did not understand the meaning of phrases and sentences such as "twice as many" or "what was the greatest amount read by one person" (Appendix F, pp. 323-328). This confusion sometimes interfered with their ability to correctly answer some problems. For example, as seen previously in Martin’s clinical interview the probes from the clinical interviewer clearly showed that he did not know he meaning of the phrase "twice as many." Consequently, the probes were not effective here because Martin did not have the requisite prior knowledge that would have enabled
him to utilize them successfully. David had a similar experience during his clinical interview when he tried to answer a sub-question from the bar graph item.

Clinical Interviewer: The next question says what was the greatest amount of books read by one person.
David: Trent (CI, lines 236-238).
Although probes later helped David to give the amount, he was confused by the fact that the problem ended with the phrase "by one person" (Appendix F, pp. 323-328). Since those words were the last he heard, he focused on them and lost sight of the real question that the item was asking. Some of the students' inability to navigate between certain vocabulary terms or their misunderstanding of the wording of certain directions sometimes resulted from misconceptions that prevented them from answering certain items successfully.

The challenges the students had with language could be attributed to the fact that the students were still learning "academic language, or the language through which school subjects are taught and assessed (Schleppegrell \& O’Hallaron, 2004). While academic language is generally used in the context of second language learners and in middle and secondary schools, it is relevant to this study because the students were learning the language of subjects to which they were being introduced in school, specifically mathematics. I suggest that the recommendations that Schleppegrell and O’Hallaron (2004) make with regards to academic language instruction for English language learners in secondary school situations are applicable to elementary school students who are being exposed to the language of various content areas for the first time. Schleppegrell and O’Hallaron (2004) state:

Academic language instruction occurs in the moment-to-moment work of teaching, and students need to support engagement with content in multiple ways. Teachers can support this by connecting with the language students bring to school, whether from first language contexts or from contexts of informal interaction that draw on everyday registers, and by creating classroom contexts where all students participate in cognitively challenging work. In addition, academic language development requires flexible pacing that responds to students' needs and growth in both language and content, and students need regular feedback that focuses on meaning as well as form.

In my opinion, the above suggestions for instruction in academic language can be applied to elementary school students as well. Like the second language learners, the students would benefit from instruction that recognizes the complexity of the language of content areas. In textbooks, it is not only the vocabulary that differs from student's daily language interactions, but also the grammatical structures of these texts Schleppegrell (as cited in Schleppegrell, Achugar, \& Oteiza, (2004). Further, for students with limited exposure to text and language, including poorer African American students, this kind of instruction could be invaluable in promoting their academic success in school.

## Misconceptions.

The discussion of misconceptions follows the examination of the issue of language because many of the students' misconceptions had some connection to language. For example, some of the boys, particularly Jake and Martin, confused the numerator with the denominator while attempting to solve the fraction problem. Martin explained the meaning of the term denominator by saying that it was the number "up top" (CI, line 603). This fundamental error inevitably precluded him from arriving at a correct answer for this item.

Misconceptions not only stemmed from language issues, but also from gaps in the students’ prior knowledge. For example, Jake knew that a centimeter and an inch were very different. However, partially due to his visual challenges with differentiating the measurements on the ruler, and incorrectly processing information about these two measures, he attributed to each the characteristics of the other.

Clinical Interviewer: The next question says here are three pieces of yarn. They're different lengths. Which string would you think would measure about 1 cm ?
Jake: (He points to the second longest string).
Clinical Interviewer: Ok. And how did you decide on that one?
Jake: Because this one is bigger...
Clinical Interviewer: Ok.
Jake: And this one is like small, and this one is like a little bit uh...
Clinical Interviewer: A little bit ...
Jake: ...inch
Clinical Interviewer: An inch. Ok. And is a cm the same as an inch?
Jake: Um ... no.
Clinical Interviewer: No? Ok. So...which one do you think is the cm ?
Jake: This one. (Still points to the one that looks like an inch) (CI, lines 25-48).

Additionally, Jake knew that to find the time on a clock, it was necessary to count by fives. Unfortunately, as we saw in the excerpt for the elapsed time item, he got the incorrect starting point, and this caused him to miss the correct answer by five minutes.

While some misconceptions such as those described above had the unfortunate result of causing an incorrect answer, others were less harmful. For example, Philip knew that coordinates needed to have two numbers. Since fractions also needed two numbers, he confused the two concepts. He correctly identified the coordinates and put them in the proper order. However, he wrote them as fractions. A probe designed to allow him to see his error proved to be unsuccessful. In general, misconceptions based
on gaps in prior knowledge negatively impacted the students' abilities to correctly solve items on the clinical interview protocol.

This was because these misconceptions were usually impervious to the positive influences of probes.

## Probes.

During the clinical interview probes were used with varying levels of success to help the students perform to the best of their abilities. Their impact was as nuanced as the different types of probes that were used and was as varied as their effects on the individual students. Patton (2002) describes probes as questions that are designed to learn more about an interviewee's response. They are also intended to give clues to the latter about the kind of response that is needed (p. 372).

Some probes are "detail-oriented", and are designed to get "a complete and detailed picture of some activity or experience" (Patton, 2002, p. 373). During the clinical interviews, the clinical interviewer sometimes used detail-oriented probes to get a clearer picture of what the student participant was thinking as they solved the problems. This example is from Kevin's clinical interview as he completed the multi-digit word item.

Clinical Interviewer: How did you do ... decide on that answer?
Kevin: I counted 10, 20, 30, 31, 32, 33, 34, 35, 36, 37, $30 \ldots 10,20,30,31,32,33$, $34,35,36,37,38,8,39$, 39 (He then stops and erases and changes his answer) (CI, lines 419-424).

It showed how Kevin attempted to solve the multi-digit item mentally, albeit unsuccessfully.
"Elaboration probes" (Patton, 2002, p. 373) are designed to keep the participant talking. The best non-verbal cue for achieving this goal is nodding encouragingly. The verbal equivalent of a nod is, "uh, huh" (p. 373). More precise verbal elaboration probes can ask the participant to elaborate, say more, or give more detail (Patton, 2002, p. 373). During the clinical interviews the clinical interviewer and principal investigator utilized elaboration probes such as the following example with Jake.

Clinical Interviewer: A dollar and ninety-five cents. And did you count? How do you know how much that is?
Jake: Because, umm....
Principal Investigator: That's how much? The quarters make....
Jake: That's how much the quarters make...
Principal Investigator: No, I am asking you how much money do the quarters make. That. The four quarters is how much?
Jake: The four quarters is a dollar
Principal Investigator: Ok.
Jake: These are six qu...six dimes. These are four nickels, four pennies... (CI, lines 150-165).

The clinical interviewer used "clarification probes" (Patton, 2002, p. 374) when she needed more information from the participants. This was especially the case when she wanted them to explain a strategy or a skill, or to have them rework a problem when they had a computational error. An example from David's clinical interview as he tries to solve a sub-question on the bar graph item illustrates this type of probe.

David: 27
Clinical Interviewer: 27? Ok. Can you write that? And how did you figure that out?
David: 'Cause I just, um, counted all the books.
Clinical Interviewer: You counted all the books? Ok. So what did you use to help you count?
David: Um, I counted the boxes.
Clinical Interviewer: You counted the boxes. Ok. Let me see. Could you just show me how you counted? Go ahead and do it again for me. Let me just see that.

David: 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 (He counted the bars on the graph).
Clinical Interviewer: Ok. That's an interesting way to do it. Do you have any thoughts on that?
David: Um ... (He shakes head meaning no) (CI, lines 213-232).
In some of the clinical interviews, the interviewer used a "contrast probe" (Patton, 2002, p. 374) to try to get the participants to compare one thing with another. This was usually an attempt to get the students to realize they had made a mistake or that they were laboring under a misconception. An example of a contrast probe occurred during Oliver's clinical interview as he answered the symmetry item. Although this example was already shared earlier, it is the best example of this type of probe from all the clinical interviews. It shows how the clinical interviewer kept presenting Oliver with figures against which he could contrast the crayon and thus see how he had erred in drawing the line of symmetry.

Clinical Interviewer: Um... So if the crayon would be symmetrical, why wouldn't this be symmetrical? (She points to the picture of an apple that is the first picture in the group).
Oliver: There is no leaf there (correct answer.)
Clinical Interviewer: I see. Ok. And how about this one?
(Points to mitt.)
Oliver: This one doesn't have an end
Clinical Interviewer: ... and how about this one (points to pig)?
Oliver: There is no ear on the back
Clinical Interviewer: And how about this one (points to ring which was symmetrical)?
Oliver: (His answer indicated that he didn't think it was symmetrical, but I couldn't hear his exact words).
Clinical Interviewer: Ok. And the crayon is symmetrical?
(Places crayon in front of him, next to the other unsymmetrical objects)
Oliver: Yes (CI, lines 288-312).
The clinical interviewer had to decide at any moment during the clinical interviews whether to probe, and if so, what kind of probe would be beneficial in a
particular concept. Unlike interview questions, probes are not usually written out in an interview so it is necessary to listen carefully to what is said and not said. In short the interviewer has to be adept at understanding what is needed by the person being interviewed at any given moment during the interview (Patton, 2002, p. 375).

A positive outcome from the probes was not guaranteed, and the success was dependent on some factors. One factor was the depth of the student's prior knowledge and his ability to tap into that knowledge and make the connections necessary to solve the question. This relates to the Zone of Proximal Development (Vygotsky, 1978) as scaffolding and giving children cues do not work if they don't have the background knowledge on which the scaffolding can be built.

It is one reason why some probes were effective and some weren't. Some of the students were on the cusp of understanding (Vygotsky, 1978) about some of the concepts, and so were unable to connect all the pieces to answer some of the questions correctly. Others had sufficient prior knowledge so that the probes were utilized to arrive at the correct answer. Probes provided the students with the time to reflect and this is crucial to the clinical interview method. Having the time to think about their answers and to then retrieve necessary information from their prior knowledge repertoire was helpful to most of the students as they attempted to solve the items.

Time.
Despite the divergence in the responses to probes, the topic of time was one area in which the boys displayed similar behaviors. It had been anticipated that the clinical interviews would last no longer than thirty minutes based on empirical data from similar
studies, the recommendations of Nowak and Gowin (1984), and on the pilot I had done with one student. However, all of the participants took over forty minutes to complete the protocol. Here are a few possible reasons.

First, some of the students worked slowly even in a classroom setting. Four of the seven - Kevin, David, Jake, and Philip - were described by their teachers as being slow workers in comparison to their peers. Second, the item with the coins took them the longest time. This was due to the open-ended nature of the item and to the fact that it was more challenging to distinguish the differences among the plastic coins than it would have been to differentiate among real coins. Additionally, the students had to use higher levels of Bloom's Taxonomy (1956) as they had to decide which coins would total the requisite amount and then explain their choices. This was considerably different than filling in a circle next to an amount from multiple choices on the NJPASS. The probability item diverged from their customary assignments with this concept. First, they had to figure out how to arrange the same colors differently on each spinner. Then they had to color in the segments on the spinners and this took some time.

Generally, all the students took as long as they felt they needed to answer the items. Being able to take as much time as they needed contributed to the boys' feeling relaxed and comfortable during the clinical interview. Further, wait time (Rowe, 1974) the time given after asking a question in which a student thinks about how to answer proved to be important for all students. Some benefited more, in particular, Jake, David, and Philip. Although the clinical interviews lasted longer than had been anticipated, none exceeded 45 minutes. The time allotted for the first session of the NJPASS assessment
was also 45 minutes, while the time for the second session was 35 minutes. In addition to time, another unifying factor that was exhibited during the boys’ clinical interviews was their persistent use of mental math. It was overwhelmingly the primary cognitive strategy used during the interviews.

Mental Math.
All the boys used mental math as a computational strategy. Some of the boys Kevin, Sam, and Oliver - were particularly proficient. David, Jake, and Martin revealed during the clinical interviews that they were more adept in this area than their teachers had previously thought. To ascertain the degree to which the students had used mental math on the NJPASS, and how that compared with their usage on the clinical interview as well as with the teachers' evaluations, I examined the Item Response Record. In addition to providing the detailed score for each item on the NJPASS, this document also presented data about the students' proficiency in mental computation on the NJPASS.

There were three items on the NJPASS that were specifically identified as evaluating mental computational strategies. They were listed in the category of number sense, operations, and properties. They were numbers fifteen, eighteen, and twenty nine. Of the three items, Kevin got all three correct which dovetails with his teacher’s evaluation and his performance on the clinical interview. Sam got two of the three items correct. This also matches his teacher's evaluation as she said he used mental math along with other strategies. On the clinical interview, he mostly computed using mental math. Oliver got three out of three items correct, which matched both his teacher's evaluation and his performance on the clinical interview.

David got only one of the mental math items on the NJPASS correct. His teacher had indicated on the first teacher interview that he didn't use mental math. During the second teacher interview she had interpreted what appeared to be the outward manifestations of his mental computation as indications of his distraction. During the clinical interview he used mental computation quite frequently, but he was not consistently successful. He tended to get the single digit computations correct while the double digits problems proved to be more problematic.

Jake got one of the three items correct, but this score tallied with his teachers’ evaluations as Dahlia said he didn't use mental math and Becca felt that he didn't use it much. On the clinical interview he did more mental computation than they anticipated, but with mixed success, so this score also matches his performance there. Martin got two out of three items correct. His teacher had been surprised at the amount of and the quality of his mental computation during the clinical interview. This score shows that he used mental math computation successfully on the NJPASS as well.

Philip got none of the mental math items on the NJPASS correct. His teacher had indicated that he used mental math computation in class. However, unlike the other participants, he did not use much mental math computation on the clinical interview. This was the case even for the basic fact items. For example, he used a number grid to compute nine minus three. Therefore, his score on the mental math items on the NJPASS tallies with his performance on the clinical interview. His score does not support his teacher's evaluation. Given the teachers' ratings for their students in this area, the topic of mental math was important to the teachers' findings as well.

## Themes from Teacher Interview Findings

## Mental Math.

The boys' use of mental math was a recurring motif throughout the teacher interviews. It proved to be an area in which some of the teacher evaluations from the first teacher interviews were shown to be in need of revision. After viewing the clinical interviews in the second teacher interviews, some teachers were surprised that their students had used mental math at all. Others were surprised at the extent to which their student used mental math and some were unsure about whether or not their students had even used the cognitive strategy. Given that proficiency in mental math was and will continue to be evaluated on the NJPASS in the foreseeable future, the teachers need to accurately determine the boys' efforts in this area. In addition to mental math, another theme that was applicable to more than one student was the structure of the clinical interview itself.

## Individualized Setting of the Clinical Interview.

Some teachers expressed the viewpoint that some of the boys had performed better in the individualized format of the clinical interview. They noted that their particular student had worked better in the one on one setting of the clinical interview than in their classroom. This was particularly relevant for Kevin and Jake. This was because Kevin was younger than his homeroom peers, and because Jake had learning disabilities including difficulty with processing information. During the clinical interview, these students felt increasingly comfortable away from the classroom distractions and pressures. Their body language as viewed on the DVDs as well the
following teachers' comments are supportive of this observation. Dahlia is referring to Jake and Liane is referring to Kevin.

Dahlia: Oh, look at that. I see he's feeling more comfortable... starting now.
Principal Investigator: From his body language
Dahlia: Yeah, and then when he's willing to ask, can I do this, can I do ... you know... (TI2, lines 406-410).

Liane: Isn't that interesting, though? Really. Now he's like chh, chh, chh (making sounds to indicate how fast he is working now in comparison to before). Ok, I'll do another one if that's what you want. He's working much faster (TI2, lines 466-468).

Additionally, the individualized setting allowed for the personalized use of probes. The effective use of probes by the clinical interviewer was often instrumental in allowing some of the students to rectify errors. This was generally the case if the error was as the result of careless computation rather than as a result of missing prior knowledge. This ability to rethink and rework answers worked better in the clinical interview setting than on the NJPASS for some of the students. This was particularly true for the boys who had been described as not liking to check their work or whose teachers had considered them to be inconsistent workers.

Inconsistency.
Another idea mentioned by the teachers was the inconsistent mathematical performance of some of the students as they completed their classwork. Some of the boys, David, Martin, and Philip in particular, would seem to know concepts and then would do poorly when presented with them on subsequent occasions. Their teachers found this pattern very troubling and revealed their concerns during the teacher
interviews. In addition to inconsistency, some of the boys were also described as wanting to finish problems and paying scant attention to whether or not the problems were correct.

## Completion over Accuracy.

A couple of these boys' tendency to favor completion over accuracy was another issue that surfaced. They felt that finishing assignments was more important than accuracy in their working. Their teachers had reported in the first teacher interviews that these students were inclined to focus more on completing assignments than on verifying the accuracy of their answers. In the clinical interview there was adult intervention that sometimes helped these students recognize and remediate their errors. Some errors were the result of computational lapses. To the teachers' surprise and relief, other mistakes were not triggered by changes in the presentation of the concepts.

Differences in the Presentation of Concepts.
The differences between how concepts were presented in the district's math program as contrasted with the manner in which they were posed on the NJPASS and the clinical interview protocol was a recurring theme among the teachers. Teachers were concerned that the students' unfamiliarity with how the concepts in the items were presented could prevent them from successfully solving the problems. The item that was used most frequently to illustrate these differences was the bar graph item since the bars were shown horizontally instead of vertically on both assessments. The teachers were fearful that due to the changed representation of this item, their students would be unable to solve it accurately.

## Bar Graph Item.

All the teachers commented that for the bar graph item on both the clinical interview and the NJPASS, the students were not accustomed to seeing the bars going horizontally. The math program presented bar graphs with vertical bars, and there was concern that this difference would cause the students to be unsuccessful in solving the items. Using the Item Response Record I analyzed how the students performed on the comparable items, numbers ten and eleven, on the NJPASS. Two of the students, Sam and Oliver got both items correct. Two others, Kevin and Jake got one of the two items correct. David, Martin, and Philip were unsuccessful on both items. These performances on the NJPASS differed from the clinical interview assessment where the students were generally more successful. However, some of that success, particularly with Martin and Philip was due to the fact that probes helped them get the correct answers for two of the sub-questions that they had originally answered incorrectly. When they were left to solve the problems independently on the NJPASS they were unsuccessful.

Incorrect answers for this item on the clinical interviews were therefore not as a result of the bars being horizontal, but were primarily due to difficulty interpreting the language of the sub-questions. Despite their teachers' concern about the lack of familiarity with the orientation of the item's bars, five of the seven students got at least three of the four sub-questions correct on the clinical interview, while four got at least half of the item correct on the NJPASS correct. For this item, unfamiliarity with the directionality of the bars was not a significant barrier to success.

Another area of discrepancy was the coordinate grid item. The students were more accustomed to seeing letters on one axis and numbers on the other. Both the NJPASS and clinical interview items had numbers on both axes. Further evidence of disparity was seen with the coin item on the clinical interview that required that the students had to select the actual amount from a group of coin manipulatives instead of completing the correct bubble on the NJPASS. Additionally, for the probability item on the clinical interview, the boys had to create their own spinners rather than choosing the likelihood of a spinner landing on a particular spot.

The teachers were relieved and pleased whenever their particular student successfully answered one of these items. This was especially true if the student did so without adult intervention. They realized that the errors for these items were not as a result of the students' unfamiliarity with their format. It was evident that most of the students were not adversely affected by the changes in the format of these items. This was particularly true of the students who received extra support from parents as well as teachers.

## Individual Variations Among Students

## Teacher and Parental Support

Extra teacher support or working in small groups was stated by many of the teachers to be beneficial for all of the students and not just the student who received special services. Furthermore, the additional parental support for two of the students contributed to their increased mathematical confidence and competence. On more than
one occasion during the clinical interview, Sam attributed his actions during problem solving to the work done with him by his father at home.

Clinical Interviewer: I like the way you labeled it too. How did you learn to label things so well?
Sam: My dad teaches me at home (CI, lines 379-382).
Clinical Interviewer: Thank you very much. You did a
very nice job. Do you do these kinds of problems some place else?
Sam: (Nods his head in affirmation.)
Clinical Interviewer: Where do you do them?
Sam: With my dad... (CI, lines 582-589).
Sam's experience as a recipient of parental support as well as the fact that his teacher noted that he benefited from small group support no doubt contributed to his math confidence and competence. His performance on the clinical interview displayed evidence of his versatility and adaptability (Sfard \& Linchevski, 1994).

## Students' Versatility and Adaptability

The students showed varying levels of versatility and adaptability (Sfard \& Linchevski, 1994). These levels varied depending on the requirements of the particular tasks and according to the amount of prior knowledge the students had about the concept. Some such as Kevin, Oliver, and Sam seemed to have a more developed set of tools especially in the area of mental math. They consistently selected the cognitive strategies that worked best for them as they tried to solve the different items. For example, Kevin's strength in rote memorization and in mental computation led him to use mental math to find the correct answers for items one, three, four, five, seven, ten, and fourteen. Oliver solved the elapsed time problem almost instantaneously using mental math.

Similarly, Sam computed the answer for the total number of books read on the bar graph item mentally. This involved counting six numbers and arriving at the correct answer of 29. Sam explained that his strategy was to start by counting the highest number first, and then going to the smaller numbers.

David, Jake, Martin, and Philip were more inconsistent overall in applying their problem solving strategies effectively. Sometimes they seemed to have the requisite pieces of the puzzle. However, they sometimes put the pieces together incorrectly. For the item that dealt with finding the missing numbers in the pattern, David calculated the missing numbers. He knew the pattern moved by twos, but did not say whether the pattern increased or decreased. Then after a probe to try and ascertain this, he answered incorrectly. David did not say as he had previously that he had counted by twos to get the correct answer. He knew that the pattern was skipping numbers but did not repeat his previous answer and instead said it was skipping only one number.

Martin was challenged by the vocabulary or language used in some of the items. This was seen in the previously cited excerpt from his clinical interview that illustrated that he didn’t know what the phrase "twice as many" (Appendix F, pp. 323-328) meant. For the measurement item, Jake was unable to correctly identify the piece of yarn that would measure a centimeter. Despite the probe, Jake did not pick the string that had the correct length. He knew that an inch and a centimeter were different, but could not isolate their differences sufficiently in order to choose the right piece of yarn. Philip knew that there needed to be two numbers for coordinates. He also knew their correct order. His error was that he wrote them like fractions. These examples show that some
of these students were on the cusp of understanding (Vygotsky, 1978) and so did not as yet have in their mathematical repertoires the necessary prior knowledge or cognitive strategies to correctly solve some items.

In sharing their views about their students' mathematical abilities, the teachers took this factor into consideration. As they evaluated the students' abilities they indicated areas in mathematical knowledge and skills in which the boys experienced challenges or successes. Their evaluations were largely corroborated by the analyses of the boys' performances on the assessments.

## Analysis and Interpretation of Findings

Qualitative interpretation begins with clarifying meanings (Patton, 2002, p. 477). Meaning making also comes from comparing stories and cases and can take the form of inquiring into and interpreting causes, consequences, and relationships (Patton, 2002, p. 479). The analysis of the findings for the first research question revealed that six of the students performed better on the standardized test items while one performed better on the clinical interview items. Oliver and David performed considerably better on the standardized test items than on the clinical interview items, while Philip had greater success on the clinical interview items. The other students had less dramatic differences in performance between the two assessments. Certain factors could have contributed to these results.

Three of the boys who performed better on the standardized test items - Kevin, Sam, and Oliver - had mathematic skills and abilities that were ranked in the higher levels, that is, in the secure to developing range, by their teachers. They had no rating in
the beginning levels. These students were considered by their teachers to be confident and competent math students. Their mathematical literacy was higher than that of their peers as evidenced by the fact that they got higher scores on both assessments. Their teachers said that all three students (Kevin, Sam, and Oliver) enjoyed math and this positive attitude no doubt contributed to their success in that subject. This mindset, aligned with their possession of the requisite prior knowledge and good math skills allowed them to be successful on both the regular and complementary assessments. Two of the three -Kevin and Sam - were described as having trouble following directions but this did not detract from Liane's belief in their overall ability in the subject. These two students - Kevin and Sam - also had the advantage of having fathers who worked consistently with them at home, honing their math skills.

A fourth student, David, who also did noticeably better on the standardized test, did not fall in the same category as these three students. On the rating scale, he had conceptual categories in which he was working at a beginning level. Additionally, his teacher thought that his mathematical performance was inconsistent. She further stated that he needed much help with his mathematics in order to be successful, and that he benefited from repetition and rewording. This support was available on the clinical interview and would not have been on the NJPASS. Nonetheless, David performed much better on the NJPASS items than on the clinical interview items.

A possible answer to David's success on the NJPASS is that maybe he had more skills than had been noticed in the classroom, and was able to use them to successfully navigate the standardized test. Additionally, the primarily multiple-choice format of the

NJPASS could have contributed to his success, as he did not have to explain his thinking on those items. He performed better on those items than on the clinical interview items. In contrast, he did poorly on three of the five corresponding open-ended items on the NJPASS while performing well on the other two. This pattern of inconsistent performance where sometimes he performed well, and at other times did not, could have contributed towards the very different performances on the clinical interview and the NJPASS assessment.

Of the two other students who performed better on the standardized test items, Jake had less variation between his performances on both assessments. He had the lowest scores on both assessments of all the participants. Jake's disabilities, specifically his processing, visual, and fine motor challenges were contributing factors to these results. Additionally, gaps in his prior knowledge plus slower developing math skills inhibited Jake's ability to benefit from adult intervention to the same degree as some of his peers.

With reference to the latter, Jake was very similar to Martin who was the other student who performed better on the NJPASS. He too had gaps in his prior knowledge that created misconceptions that sometimes prevented him from answering items correctly. Consequently, like Jake, probes were sometimes not effective with Martin. This could be an explanation for the small percentage difference between his scores on both assessments and for the fact that his scores were some of the lowest on both assessments.

Martin's teacher had described him as an inconsistent and careless worker. Rae had noted that he preferred to be finished with a problem rather than to be correct, and he was not careful about checking his working. These last two traits were revealed on the clinical interview when he ignored a probe designed to get him to reflect on an answer to an item. Despite the repeated probes of the clinical interviewer as to whether he had used up all his money or whether he had enough money to buy the various products, Martin did not use them to reflect on the accuracy of his answer.

Philip was the only student who got better scores on the clinical interview. Like Martin, he was described as not always being concerned with accuracy when solving problems. Similarly to Martin and David, Philip was described as being inconsistent. However, he benefited from probes on the clinical interviews to a greater degree than Jake and Martin for example. He got problems correct that, if he had been left to his own devices, he would have gotten incorrect. The adult intervention in the clinical interview helped to enhance this student's performance on that assessment considerably. This improvement was also due to the fact that his prior knowledge was adequate enough to allow him to take advantage of many of the probes that were provided to him. His teacher had given him levels of developing and one secure on the evaluation scale. His teacher's "gut feeling" (TI-1, lines 309-310) that he would perform better on the clinical interview than on the standardized test were realized.

## Conclusion

The findings show that while six of the seven students performed better on the standardized test items than on the clinical interview items, there were variations within
these results. The students who were more mathematically competent, those whose mathematical literacy was more honed, were able to perform successfully on the standardized test. They also did well, albeit to a lesser degree, on the complementary assessment. The format of the NJPASS, with multiple-choice items that contained the correct answers could have contributed to the differences in scores. On the standardized test, the students did not have to generate as many answers as they did on the clinical interview. This suggests that African American students who have sufficient knowledge of mathematical concepts that are supplemented by good skills and cognitive strategies can perform well on different types of assessment.

It can also be postulated from the findings that African American students with a sufficiently robust understanding of mathematical concepts, who are not always careful or consistent workers, can potentially show more knowledge of mathematical concepts on an assessment such as the clinical interview than they can on a standardized test. This possibility is due to the twin advantages of adult intervention and ample time that are possible with the clinical interview. The student who performed better on the clinical interview benefited from the extended time and from probes from the interviewer. Probes provided him with the opportunity to see most of his errors and to make corrections. He also had the adequate prior knowledge that enabled him to take advantage of this adult intervention. These findings would therefore suggest that the clinical interview as a complementary assessment could be beneficial to students with reasonably robust academic ability who are not careful workers and who are inconsistent in their application of their mathematical knowledge and skills.

The clinical interview was useful as a "window" into the cognitive processes of students who had well-honed mathematical skills and knowledge. For example, this assessment allowed us to discover Sam's strategy for adding a series of numbers mentally and showcased Oliver's remarkable mental computation on the clock item. However, it was also helpful for those who had gaps in knowledge that created misconceptions. The clinical interview also revealed Jake's misconceptions about counting time on clocks and Martin's inability to distinguish between the numerator and the denominator. Such knowledge provided teachers with information that complemented the data that they obtained from standardized test scores. From the perspective of a cost-benefit analysis, the clinical interview would be particularly helpful to inexperienced teachers as opposed to very experienced educators. Their inexperience would cause them to be less adept at accurately evaluating their students’ abilities. Insights from the clinical interview would therefore potentially be more useful to them. The clinical interview is an assessment that gives teachers accurate data about their students that they can use to meet the individual needs of these students. It can be another resource to help them accurately assess their students' mathematical knowledge and thus have realistic but affirming expectations of their students.

The findings for the second research question showed that the teachers' views of their students' mathematical competency were generally positive. Their initial evaluations from the first teacher interviews were substantiated by examples of their students' academic work. Since they had had the opportunity to review their students’ mathematical achievements prior to the first teacher interviews, the teachers’ evaluations
were usually consistent with their students' performances on the clinical interviews. Therefore, it was not surprising that the majority of the teachers said that their views of the students' abilities had not changed after viewing the clinical interview.

These findings show that the teachers had a comprehensive understanding of their students' academic abilities based on more than standardized test scores. They possessed realistic yet positive views about their students' capabilities. Additionally, these teachers were amenable to engaging in reflection on their practice. Two were willing to go even further and make pedagogical changes to address issues that they observed during the clinical interviews. Their behavior is typical of what Irvine (1990) outlines as part of the characteristics of effective teachers of minority students. "They restructure the learning activities, assuming that the child has not yet mastered the materials, not that the child is incapable or unwilling to learn" (p. 94).

Without exception, all the teachers appreciated being able to watch their students' clinical interviews. As previously mentioned excerpts indicate, the teachers found the interviews fascinating and helpful. While their views about their students' abilities may have remained the same, they universally expressed enjoyment and appreciation about being afforded the opportunity to watch their students' problem solving efforts. They valued the opportunity the clinical interviews afforded to view the mental processes of their students. The interviews confirmed or informed the teachers' opinions of their students' abilities. The teachers also appreciated the opportunity to evaluate their pedagogical practice in light of the differences between their math curriculum and the NJPASS and clinical interview items.

Therefore, while the clinical interview did not provide all of these African American students with the opportunity to show that they knew more than they would be able to do on a standardized test, it revealed much information about their knowledge of mathematical concepts as well as the cognitive processes they had used in problem solving. This knowledge can be helpful to teachers in their goal to improve the academic performance of their African American students. This quest was the reason for undertaking this study. Chapter Five discusses and summarizes the study and its findings.

## CHAPTER FIVE

Discussion and Summary

The performance of African American students on standardized assessments has been and continues to be an intransigently puzzling issue for educators and policy makers nationwide (Evans, 2005). Some African American students perform as well as or better than their White peers on standardized tests (Noguera, 2003; Slavin \& Madden, 2006; Stinson, 2006). However, as a group, many African American students continue to underperform their White peers on standardized tests (Haney, 1993; Gardner \& Miranda, 2001; Ladson-Billings, 1994).

This underachievement is evident across social classes as some middle-class African Americans also earn lower scores than their White peers (Ferguson et. al, 2002; Viadero, 2000). For decades, the focus of much educational reform has been to use accountability testing to create improvements in the classroom with the goal of improving students' academic performance (Linn, 2006). These reform initiatives have involved increasingly more standardized testing (Linn and Hambleton, 1991), especially in the wake of the latest reform initiative, the NCLB Act of 2001 (Solley, 2007).

Since the results of these tests are used as indicators of how the nation's educational system is faring (Miller, Linn, and Gronlund, 2009) standardized testing is likely to remain an important factor in the educational system for the foreseeable future. This is problematic, as many - though not all - African American students find it challenging to perform well on standardized tests (Flowers, 2007; Slavin \& Madden,
2007). Consequently, exploring other methods of assessment for these African American students in addition to standardized tests would be beneficial for them and the educational system (Gordon, 1995). While the research highlighted the students' performance on standardized assessments, more investigation into complementary testing methods was necessary.

## Gap in Literature

The gap in the literature that this exploratory study addressed was the examination and identification of potential complementary assessments for African American students. To facilitate the inquiry, a very different kind of assessment, the clinical interview, was used to open up a "window" about what these students knew about mathematics. Learning more about their students' mathematical ability would be an added bonus for teachers. My desire to understand more about the thinking that supported African American students’ problem solving efforts developed over more than two decades of teaching.

## Origin of Study

In my capacity as an African American elementary school teacher in a predominantly White suburban school district, I witnessed the chronic underachievement of many of my African American students. These students, frequently, though not uniformly, underperformed on tests in comparison with White peers. This interest increased concomitantly with the current importance of testing in our educational system. As an African American teacher and mother, my quest had both professional and personal aspects.

National test data (NAEP, 2009) and (NCES, 2009) corroborated my professional observations. Desirous of contributing to efforts aimed at ameliorating the situation, I shared my concerns with other educators. As a result of this collaboration, I was invited to participate in two studies conducted in my school district. Through these studies I was introduced to the clinical interview as it was being used in educational research. While viewing tapes of the clinical interviews of students during these two studies, I saw how much information children of different races and ages revealed about their knowledge and understanding of mathematics. My colleagues and I learned a great deal about the students' skills and knowledge as a result of watching the interviews. The experiences led me to wonder if the clinical interview would provide an in-depth perspective of their mathematical understanding. Specifically, I wanted to explore whether this different kind of assessment would allow African Americans to reveal the real extent of their mathematical understanding better than they would be able to do on standardized tests.

## Research Questions

The aim of this study was to see if the clinical interview would enhance the ability of African American boys to show what they knew about mathematical concepts. The inquiry also explored what their teachers could learn from watching their clinical interviews and if this would alter the views these teachers had about their students.

The research questions were: 1) Will a complementary assessment show that the interviewed students actually know more about mathematics than they are able to reveal on the standardized test? 2) What might teachers learn about their students’ mathematical ability and how might this knowledge affect their views of those students?

## Overview of Design

This was a qualitative inquiry that utilized clinical interview methodology. In addition to the clinical interview protocols, the primary data sources were the clinical interview transcripts, teacher interview transcripts, and the NJPASS scores. The vignettes in Chapter four were compiled from this data and provided the supporting details for the findings. The constructivist ideas of Piaget $(1969,1978)$ and Vygotsky (1978) provided the theoretical framework of the study. The influence of this schema on the study was evident in the following examples. The vignettes outlined the boys' metacognitive processes as they solved the items on the clinical interview protocol. They showed how the clinical interviewer used probes to learn more about the students' mental cogitations during the clinical interviews. In calculating how effective these probes were, I found the Zone of Proximal Development (Vygotsky, 1978) for the students’ responses that were assisted by adult intervention. Some of the themes that emerged from data analysis such as misconceptions were the results of the students' attempts to make sense of what they were being asked to do. Finally, as the teachers watched the clinical interviews they had to construct an understanding of what they were learning about their students. This reflection was crucial as it provided the rationale for the teachers' decisions about whether or not their opinions about their students had evolved from their earlier evaluations.

Using clinical interview methodology, I was able to focus on each participant's individual response and perspective in the vignettes. The methodology also enabled me to reveal the similarities and differences among the participants. While I identified,
described, and analyzed the varied viewpoints, I recognized that my personal background and experiences influenced what I had observed and understood about them (Patton, 2002, p. 569). To counteract these researcher effects, I asked the adult participants to verify that the transcripts of their interviews accurately reflected their views and statements. Additionally, when I had ascertained the results of the study, I shared them with the teachers and asked for their comments. The responses will be shared later in this chapter. Involving the participants' feedback or "review by inquiry participants" (Patton, 2002, p. 560) is one of the methodological choices that I implemented to ensure the credibility of the study and its findings.

The setting and sample were purposefully and conveniently chosen (Patton, 2002, pp. 235, 241). They affected the findings due to the location of the school and to the characteristics of the teachers. The school was located in a suburban area and the students were exposed to a comprehensive district mandated curriculum that met the requirements of the state core content standards. The students were not taught to the test. While the teachers prepared the students for the NJPASS, they did not do so by narrowing the curriculum. On the contrary, whenever they felt that the curriculum had provided inadequate coverage of a particular topic, they supplemented the district material with outside resources. The teachers also welcomed district introduced enrichment activities such as the chess program, First Move. The high quality of the curriculum enabled three of the boys to perform well and two others to perform creditably on both assessments. Only two of the seven students did not earn scores of proficient or higher on the NJPASS.

The curriculum was not the only factor that affected the boys' achievement. The other important element was the students' teachers. All the teachers in the study had attained the highly qualified status required by the state. Many had been teaching for decades, and most had spent many years in their assignment at the time of the study. The great degree of qualifications and experience of the teachers contributed to the high caliber of the boys' mathematics instruction. Their professionalism and years of experience also enabled the teachers to form realistic but unbiased opinions of their students' mathematical abilities. The opinions essentially rendered my second hypothesis about the results of the study moot.

## Findings

There were two main hypotheses of the inquiry. The first hypothesis was that the students would show the depth of their mathematical understandings better on the clinical interview than on the standardized test. The second was that the teachers would learn new things about their students that would improve their views of these students’ mathematics ability. The findings showed that six of the participants performed better on the standardized test items, while one performed better on the clinical interview items. Despite the empirical support for the predominant and historic academic underperformance of African American students on standardized tests, the majority of the boys in this study performed better on the corresponding NJPASS items. By this I do not mean that they all did well on the entire NJPASS assessment, but rather that they got a higher percentage of the corresponding items correct. Consequently, although two of the students earned higher scores on the corresponding standardized test items in the study
they did not do well on the complete math segment of the NJPASS. Jake got $47 \%$ of the NJPASS corresponding items correct, but his total NJPASS score was 45\%. Martin got $60 \%$ of the NJPASS corresponding items correct, but his total NJPASS score was $55 \%$. Therefore, Jake and Martin earned scores of basic on the complete mathematics section indicating that their math skills were less than proficient. Their results added to the plethora of data that document the underperformance of African American students on standardized tests (NAEP, 2009; NCES, 2009).

Keeping the study's results in mind, if viewed as a "window" into African American students' mathematical abilities, the clinical interview was useful. The clinical interviews provided detailed pictures of the cognitive strategies, prior knowledge, mathematical aptitude and ability for each of the seven students. The information that the interviews revealed supported empirical claims that the clinical interview exposes the thinking that accompanies students' problem solving efforts (Ginsburg, 1997).

As a complementary assessment that provided African American students with the opportunity to show their knowledge and skills better than they could on a standardized test, the findings were equivocal. While six of the seven boys earned better scores for the corresponding standardized test items, there was a 50 percent difference between the highest and lowest score for these NJPASS items. There was a 35 percent difference in the highest and lowest scores for the clinical interview items. These score differences highlight the differences in ability among the students, but are not intended to conclude that one assessment is more capable of discerning academic differences than the other.

Although all these boys attended schools in a middle class school district, and were not subjected to a narrowed curriculum like many of their urban peers, these scores and data from the clinical interviews revealed that they had distinctly varying levels of prior knowledge and mathematical abilities. Of the students who performed better on the standardized test items, two (David and Martin) were identified as being inconsistent workers. One of these two students (Martin) was also described as being more motivated by a quest for completion rather than accuracy when solving problems. A third student (Jake) had learning disabilities that impacted his performance on both assessments. These three students revealed more gaps in their prior knowledge and less math skills than their peers. These differences help to explain their poorer performance in comparison with the other participants.

For the two boys who did well on both assessments (Kevin and Sam), their robust skills and prior knowledge contributed to their success on both the clinical interview and NJPASS assessment items. Further, these students’ levels of mathematical literacy were enhanced by parental support at home. For the two boys (Oliver and David) who did considerably better on the standardized test items than on the clinical interview items one (Oliver) belonged to the group who generally exhibited more robust skills and knowledge. The other (David) belonged to the group that displayed less mathematical competency in his clinical interview. Possible reasons for David's increase in scores on the NJPASS despite his weaker mathematical skills can only be surmised. There is not enough information from the NJPASS results to make a determination.

The findings of the study were not only limited to the scores on the assessments. They also included ideas that emerged from data analysis. One important element that arose from the clinical interviews was the efficacy of probes with the students. As Vygotsky (1986) noted, "the development of concepts, or word meanings, presupposes the development of many intellectual functions" (Vygotsky, 1986, pp. 149-150). During the clinical interview when the required intellectual functions were present, adult intervention by way of probes made it possible for these students to be successful on some of the items on the clinical interview. Probes and their mixed effectiveness in helping the students to correctly answer items were integral to the findings of the study. The themes that will be discussed next were identified during analysis of the clinical and teacher interviews. With specific reference to this study, a theme is an issue or idea that appeared repeatedly in the student and teacher transcripts. These themes are: probes, language, misconceptions, time, mental math, differences in concept presentation, and parental support.

## Theme

## Probes.

Students experienced mixed success with probes during the clinical interviews. Sometimes they worked and sometimes they did not. The efficacy of the probes depended to a considerable degree on whether or not the student's prior knowledge enabled them to apply the hints contained in the probes appropriately. This ability, or lack thereof, determined whether or not they arrived at the correct answers to the items. As Farkas (2003) states, the skills and experiences that students bring to the task are
important for student learning to occur. "If prior learning has not created sufficient skill and knowledge for the student to be ready to cope with the assigned instructional tasks, little achievement of new learning will occur" (Farkas, 2003, p. 2). Similarly, Bainbridge and Lasley II (2002) assert, " Prior learning affects the future achievement of all students, regardless of race" (p. 427).

The students' level of success depended on the depth of their prior knowledge in the particular concept being assessed. If they were on the cusp of understanding with reference to a particular concept (Vygotsky, 1978), probes in that area generally proved to be ineffective. If they were adequately competent in the concept and had good prior knowledge the probes were usually effective. Due to the varying levels of prior knowledge, some students needed much scaffolding in order to answer the items correctly, while others needed less. Sometimes this difference was apparent when children tried to solve the same item. Nonetheless, overall, probes were crucial in allowing many of the students to arrive at correct answers.

Although the NJPASS and the clinical interviews were similar in that the questions were read to the students, a significant difference between the two assessments was that during the NJPASS probes were not permitted, while during the clinical interviews, probes were not only permitted, but were an integral part of the procedure. On several occasions, students would have answered items incorrectly had they not been the recipients of probes from the clinical interviewer.

Many of the probes were in response to confusion about the language in some of the items. Boaler (2003) states that language can have a significant effect on students'
success on standardized tests. It also had a noticeable impact on the students' success on the clinical interviews. As I explain below, language had an important impact on the findings as it affected whether or not the boys got some items correct or incorrect. Theme

## Language.

Language - processing and understanding what the directions meant, recognizing key phrases - was an important element in the clinical interviews. More than one participant struggled with the term "twice as many" (Appendix F, pp. 323-328) and did not know what it meant. A shift in emphasis away from identifying shaded portions of figures to determining the un-shaded parts was also problematic for some of the students on the clinical interview protocol and the NJPASS assessment. These challenges were due in part to the fact that the students' math program sometimes used different wording than the NJPASS and the clinical interview protocol. A few of the teachers made this point during the teacher interviews. However, other issues with language were not as a result of these differences.

For example, some students gave the name of the person who had read the largest number of books rather than the number itself as asked. Others struggled to understand what seemed like simple directions - 'Can you write another number that is like this one?’ All the students missed the crucial direction in the chart problem directing them to pick "as many different rides" (Appendix F, pp. 323-328) as possible even when the word was emphasized. It was enlightening to see that even with the problems being read aloud, and reread as many times as necessary, and with the availability of probes,
language comprehension provided challenges for all of the participants at some point during the clinical interviews. Matteson (2006) has written about the need to recognize the importance of reading comprehension on multiple-choice standardized tests, but its importance to open-ended assessments as well was revealed during the clinical interviews.

An examination of the boys' success rate with the open-ended questions on the NJPASS (Table 12, p. 243) revealed that the students did not appear to find them as challenging as some of the open-ended items on the clinical interview. Consequently, an explanation for the boys' getting more of the standardized test items correct could be this varying level of difficulty. While all the boys got some of the open-ended items correct on the clinical interview, they did better overall on the corresponding multiple-choice items on the NJPASS. The challenge with the open-ended items was primarily seen in the student's misunderstanding of the language used in the directions.

The implications here are that language comprehension could be very important for the success of African American students on assessments in general, as well as for learning. It suggests that they should be exposed to an adequate amount of mathematical vocabulary so as to improve their ability to understand word problems (Matteson, 2006). The teachers of some of the participants had indicated that math was a favorite subject or the student's strength. Teachers could use these positive attitudes to hone their students' comprehension of the language of mathematics. In order to explain their thinking the children need to learn to use language effectively. The results for the first research
question depended a great deal on the students' comprehension of and facility with written and oral language during the clinical interviews.

Consequently, constructing meaning through language was a critical factor of the entire study. Our perspectives of the world are "shaped by cultural and linguistic constructs" (Patton, 2002, p. 96). During the clinical interview, language was crucial in understanding what was being asked, and in conveying responses to the interviewer. Language was also important with reference to probes and their effectiveness in helping students solve the items successfully. For example, the probes that the clinical interviewer used with the students were all verbal. In this way, her attempts to get the students to reveal what they knew or to offer clarification were effected through the use of language. The language in the clinical interview's open-ended items not only affected the students' ability to understand the directions. It also helped to reveal some of their misconceptions. Many of the misconceptions that were revealed by the students during the clinical interviews would not be as easily detected on a standardized test.

## Theme

## Misconceptions.

Misconceptions due to language confusion and lack of prior knowledge led to incorrect answers to items during the clinical interviews. The misconceptions that the boys had as a result of gaps in their prior knowledge substantiates the findings of researchers such as Farkas (2003) who described the importance of this type of academic base for students' academic success in school. Almost all the students had at least one misconception, and unsurprisingly, the students with weaker math skills revealed more
misunderstandings. With respect to the relationship between misconceptions and probes, the implications are that the clinical interview is better equipped than a standardized test in providing teachers with information that can allow them to detect misconceptions in their students' knowledge. The exploration of students' thinking through the use of probes is not possible during a standardized test, and it illustrates the usefulness of the clinical interview in opening up a "window" into students’ cognitive processes. Gathering information based on students' errors is evocative of Piaget's interest in analyzing student's mistakes (Evans, 1973). In addition to shedding light on students’ misconceptions, the individualized setting of the clinical interview meant that students did not have to vie for time with peers. They had uninterrupted time in which to utilize their problem solving strategies.

## Theme

Time.
The students may have had different misconceptions. Nonetheless, they had similar experiences when it came to the length of time they took to complete the clinical interview protocol. All of the participants took over forty minutes to complete the protocol. Possible reasons stem from characteristics of both the students and of some items on the protocol. For example, some of the students were identified as being slow workers in comparison to their peers. Further, a couple of the items - the coin and probability items - required the students to spend extra time on the manipulatives and to work at higher cognitive levels.

All the boys in the study reacted positively to being able to work noncompetitively at their own pace during the clinical interviews. This suggests that ample time is an important factor for some African American students being successful on assessments. The student who worked the slowest on the clinical interview was also the one whose performance was definitively better on that assessment. The student with the learning disability also benefited from the wait time (Rowe, 1974) that the clinical interview permitted in contrast with the standardized test. Additionally, the students who were more concerned with completion rather than accuracy were able to correct errors due to the fact that the structure of the clinical interview permitted the time for extensive probing. Finally, six of the seven students used mental math in their computations. When they were engaged in their mental computations, they often did not respond immediately after the items had been read. Having enough time to do mental and other computations was invaluable for these students.

## Theme

## Mental math.

Mental math was the cognitive strategy of choice for the majority of the students during the clinical interview. All but one student used this strategy while solving the items on the clinical interview. Mental math was also the area in which the teachers expressed the most surprise while watching the clinical interviews. A few teachers expressed surprise that their student had used mental math at all, others were surprised at the amount of mental math that their student had used, and one did not think that her student had used mental math.

The teachers acknowledged differences with their initial evaluations about their students' use of mental math. However, when asked if they had learned any new information from the clinical interviews that had changed their views of their students' mathematical abilities, only one admitted that her student's use of mental math constituted something new.

One reason for this response may have been that the teachers' focus was more on their students' performance on the actual concepts that were assessed. Thus, they may have been less concerned with their students’ cognitive strategies. This possibility may have been caused by the fact that most assessments including standardized tests examine students' knowledge of concepts.

However, the NJPASS does evaluate mental computation so the implications for the teachers are that they should be more attentive to their students' abilities in this area. Some of the teachers may not have been sufficiently familiar with the indications that their students were using mental math during daily lessons. With the importance of mental math as a strategy on the NJPASS, it is necessary for these teachers to become more familiar with their students' use of mental math.

In addition, mental math is a higher order cognitive activity (Bloom, 1956). Students proficient in mental math would be exhibiting signs of increasing mathematical literacy (Moses, 2001). Honing students in this skill would be instrumental in improving their mathematical competency and commensurately their performance on standardized tests and other assessments.

In contrast to their reactions to their students' use of mental math, the teachers were all interested in the differences in presentation of the concepts among the school curriculum, the clinical interview protocol, and the NJPASS items. They were worried that their students' unfamiliarity with the way the concepts were presented on some of the items in the two assessments could compromise the boys’ ability to solve the problems successfully.

## Theme

## Differences in concept presentation.

With regard to concept presentation, the teachers were specifically concerned about the following: the bar graph item that displayed horizontal bars instead of the vertical bars to which the children were accustomed; the fact that the students were used to seeing letters and numbers on the coordinate grids rather than numbers and numbers as was the case with both the clinical interview and NJPASS items; the probability item on the clinical interview which had the children create their own spinners instead of identifying the chances of the spinner landing on a particular color as presented in both the district curriculum and the NJPASS; and the fact that the students were more accustomed to being asked to find the shaded and not the un-shaded part of a figure as was the case with the clinical interview fraction item. The teachers believed that these differences in presentation added to the challenges that some of the students faced during the clinical interviews. However, the student's responses allayed the concerns. For each of these items, some of the students solved them independently. Others completed them correctly with the help of probes. The results showed that in solving the problems
outlined above, all of the students were able to do what was required rather than only what was familiar, albeit with varying levels of competency.

The implications for this accomplishment are that the students had been exposed to a sufficiently varied curriculum. They were not subjected to the narrowed curriculum that so many African American students in poorer neighborhoods experience (Barton \& Coley, 2008; Koretz, 2008). Wider exposure allowed them to take changes in presentation in stride. The instruction that is provided for students is one of the main sources of their conceptual knowledge. This teaching also has significant impact on children. It decides the quality of their mental development (Vygotsky, 1986, p. 157). The teachers' concern about differences in the way that concepts were presented on the two assessments indicates that they are very aware of the effect of their instruction on the students' conceptual knowledge. Their concern also indicates that they were not teaching the students to the specific test.

This result raises the topic of teaching a concept so specifically that students are unable to transfer that knowledge to anything else that does not meet the same criteria. It also confirms the need for African American students to be taught curricula that are not narrowed to meet testing goals (Orfield \& Wald, 2000). Instead, they should be taught to develop a fluidity of thinking that allows them to apply what they already know about one concept to another about which they may know less (Koretz, 2008). Instead of being instructed in ways that limit them to performing at the lower levels of Bloom's Taxonomy (1956) their curricula should optimally engage them at all its levels. This kind of instruction will have positive consequences for their abilities to perform well on
standardized tests and other assessments, as they will have a sufficiently diverse and broad academic foundation.

## Themes

## Conclusion.

The themes provided detailed information about the mathematical background and competency of the students. They created pedagogical questions or issues for the teachers' reflections. This analysis of the themes suggests that in providing teachers with insights about their students' mental process, the clinical interview is an assessment that gave them the kinds of information about their African American students not readily accessible from the typical standardized test. Although the majority of the teachers in the study did not learn anything substantively new about their students' mathematical abilities, the clinical interview served the purpose of confirming their initial evaluations.

The clinical interview provided the teachers with the opportunity to reflect on their practice with particular reference to their dual responsibilities towards their students. They are required to teach the students the district curriculum. However, these students also take a standardized test for which they need to be adequately prepared. The teachers examined the clinical interview items for similarities and differences with both the NJPASS items and their math curriculum. They then determined how best to use this information in their pedagogical practices. After doing this analysis, two of the teachers in the study, decided to adjust aspects of their pedagogy to reflect this information. Implications of Findings

The study's results indicate that even for students who performed well on standardized tests, the clinical interview provided the teachers with insights into their students' mathematical abilities that were not readily available on the mandated assessments. However, while the clinical interview provided two teachers with a deeper insight into the mental processing of their African American students, the teachers' responses to this assessment were varied. What the teachers learned was contextualized and individualized depending on the students involved.

The findings also indicate that African American students who were inconsistent workers, who are more concerned about completion than accuracy, or who need ample time could benefit from a complementary kind of assessment such as the clinical interview. These students could be helped by the adult intervention (Vygotsky, 1978) that is available during the clinical interviews by way of judicious probing (Patton, 2002, p. 365).

An assessment that allowed students to work not only at the zone of their actual development, but also at the zone of their proximal development, (Vygotsky, 1978), could be very useful to African American students, regardless of their levels of mathematical understanding and skills. This assessment could be a way of improving or honing their mathematical literacy by giving their teachers information that they needed to make changes in their instructional practice. This could be a topic for future inquiry.

The study's findings not only have implications for future research, they also contribute to current empirical knowledge. The good quality of the mathematical competency displayed by most of the students in the study reinforces the evidence from
empirical research that positive teacher expectations have a beneficial effect on students’ academic performance (Montague \& Rinaldi, 2001). During the second teacher interviews, the teachers' pride in their students' accomplishments was clearly evident. The positive attitudes they held about their students were reflected in their thoughtful responses to the study's findings.

## Teachers' Responses to Study’s Findings

The following comments by the teachers in the study in response to the findings continue the habit of reflection that they exhibited throughout the study. They exemplify the type of educators that Rousseau and Tate (2003) believe are essential in today's educational environment. They are professionals who think about various elements of their praxis as well as the effects those factors might have on the academic success of their students. Jake's teacher, Becca, was surprised that he had done better on the standardized test items than on the clinical interview items. She thought that the feedback that he got during the interview had helped him, and so she was surprised that he did better without it. He had used strategies that he hadn't normally used in class, and he had appeared to be very comfortable during the clinical interview.

Becca also provided insight into Jake's performance on the standardized test. As a special education student, his Individual Educational Plan (IEP) allowed him to take the standardized test in a small group setting. There he used a lot of mental math and manipulatives, but his score was low even with the extra time that he was allowed. He made basic mistakes and didn't complete all the steps to accurately solve some problems. She felt that this performance was reflective of his work in the classroom.

She further reflected that his performance on the NJPASS showed that accommodations alone did not provide sufficient help for him. The models he had in the classroom could not be provided during the NJPASS assessment, and he was unable to recreate them adequately on his own. Becca's response suggests that although Jake scored five percentage points better on the standardized test items, in her opinion, the quality of his performance on the clinical interview was better than on the NJPASS.

Dahlia, Jake's classroom teacher, indicated in her response to the findings that she had found his clinical interview to be extremely interesting because it provided her with such an in-depth look into how he approached and perceived mathematics problems. She reiterated that Jake struggled and still struggles with math. While he had the basics he did not use his prior knowledge unless prompted.

This statement underscores the importance of prior knowledge in this study due to the substantial differences in this area among the students. Dahlia stated that Jake benefited from wait time (which also corroborates this area of the findings), and she reiterated that he worked very slowly and cautiously. Dahlia ended by saying that she felt his performances on both the NJPASS and the clinical interview were similar. She felt that both were consistent with his abilities.

Liane, Kevin's and Sam's teacher shared that she was pleasantly surprised by the knowledge that her students had displayed and credited the program used by the district as being partially responsible for their success. This comment supports my hypothesis that the teachers in the study did not teach a narrowed curriculum that focused primarily on the test.

At the end of her comment, Liane wondered if the conditions, that is, one on one versus testing in a familiar classroom, contributed to the difference in scores. This is very interesting as it counters the findings indicating the benefits of the individualized setting for the participants. It also contradicts her response in the second teacher interview when she observed how well Kevin worked in that more private setting.

In her response, Rae indicated that she had found the findings surprising as she too had expected that the clinical interview would provide a higher score or give the students the opportunity to demonstrate their knowledge better on a standardized test. She acknowledged that the scores for her students were very close, but she had anticipated a greater disparity between them. Since her student had difficulty with math, she had felt he would score higher on the clinical interview items. This was due to the fact that the questions would be read to him providing him with the opportunity to hear the questions orally. He sometimes read quickly and missed key words.

Rae affirmed that she had learned more about Martin due to the clinical interview than she had from her previous observations of him in class. She reiterated that she had not known that he often used mental math to solve problems. She wished she had recognized his capability in this area earlier. She felt that if she had recognized his use of this cognitive strategy, maybe she could have done more to improve his skills in that area. Also she might have realized that his use of mental math could have indicated that he had forgotten a math diagram or strategy.

Mary was pleased to learn that her prediction that Philip would perform better on the clinical interview items had been substantiated. She was the teacher of the sole
student whose scores on the clinical interview items were better than his scores on the NJPASS items. In her response to the findings, she reiterated that Philip worked better without time constraints. This supports the study's results with reference to the issue of students having ample time during the clinical interviews. She often had to break tasks up for him, and he seemed to do well when given the opportunity to verbalize his responses. Mary indicated that he responded well to authentic praise and encouragement. As a result of these factors she felt that the clinical interview was more conducive to his academic style and needs than a standardized assessment such as the NJPASS.

Danielle shared that she found the findings interesting and surprising. She indicated that she was surprised by the fact that the NJPASS scores were higher. Then she stated that both students had scored as she had expected on the clinical interview protocol. She thought this was intriguing as they had taken the clinical interviews separately. She posited that working independently in that setting could have made them nervous, and perhaps they performed better with a whole group. So, like Liane, Danielle suggested that the difference in the nature of the two assessments - individualized versus group - could have contributed to the variations in the scores. Ladson-Billings (1997) referred to the cultural preferences that children brought to school. Could this difference have been reflection of the students' cultural inclinations? The teacher responses that were summarized above contributed to the study's results. The responses were not the only part of the study that affected the findings. They were also affected by the study's design that in turn contributed to its limitations.

## Limitations and Future Research

This was a qualitative study that utilized clinical interview methodology. The research questions that were explored using this methodology contributed to the limitations as they dictated that the student participants had to be African American males enrolled in elementary school. They also required that their teachers be included in the study. It was a small sample comprised of a total of thirteen participants - seven students and their six teachers.

They were chosen because I had access to the students, teachers, and a standardized test with which to compare the clinical interview items. The sample was therefore selected because of its convenience. In educational research this type of sampling is the least preferred choice (Patton, 2002, p. 242).

The time line of the study was driven by the need to administer the clinical interviews after the NJPASS in order to limit carryover effects. Thus the teacher and clinical interviews had to be scheduled around the standardized assessment. As a result, the time actually spent with the participants was only two weeks. Thus, due to the study's limited scope, the findings of the study may not be generalized to a different population or to different settings.

Another limitation was my inexperience as a researcher. An example of this lack of expertise was the insufficient information about coding the sample protocols that I shared with the first rater. This communication deficit was prompted by my desire not to bias her scoring in any way. However, my perspective did not take into account that if I had wanted to ensure that we graded consistently, I should have provided a more detailed scoring guide than I provided. This guide would have clarified the questions on the
protocol, provided desired answers, and possible student solution methods. This example underscores the importance of the researcher in qualitative study where "the researcher is [italics added] the instrument" (Patton, 2002, p. 14). My decision, though ultimately resolved, did negatively affect the inter-rater reliability process and the results themselves.

By the very nature of the clinical interview, the clinical interviewer had a tremendous impact on the methodology of the study as its results were greatly influenced by her actions. By this I mean that any omissions or errors on her part even where not clearly identifiable in the outcome of the clinical interviews may have influenced the outcome of the study. Consequently, while using one clinical interviewer contributed to the consistency of the study's methodology, the effect of her actions on the results was a limitation of the study.

Limitations such as the ones described above are inherent in any research study. The seeds for future inquiry are also present in these flaws. For example, the study's purpose necessitated using African American male participants. If the sample was changed to focus on girls, gender differences in conceptual knowledge could be investigated. Also, since most of the boys in the study used mental math, gender differences in the use of cognitive strategies could also be explored. Gender differences were not the only avenues for future inquiry that emerged from the data in this study. The topic of modifying teachers' instructional practices to benefit both male and female students also became apparent.

In her response to the findings, Mary indicated that she wanted to better prepare her students for the standardized test. She wanted her students to become more cognizant of the connection between the curricula they were learning in the classroom and the concepts on which they would be assessed on the standardized test. Future inquiries could investigate whether the clinical interview could be incorporated into teachers' pedagogical practices in a way that helped them hone their students' mathematic literacy skills to best enhance performance on standardized tests.

Additionally, future inquiry could not only address regular education students of both genders, but also students who had Individual Educational Plans. Jake was the study's only student who was classified as having learning disabilities. In her reflections on the findings, his special education teacher, Becca, indicated that while working independently during the standardized test, Jake had been unable to recreate the helpful strategies that had been modeled for him in the classroom. Dahlia also indicated that Jake needed to be reminded to use his prior knowledge. Investigations could explore the possibilities of using the individualized setting of the clinical interview to help students like Jake practice previously modeled strategies. During a clinical interview, problemsolving opportunities that would allow them to practice these modeled strategies could be provided. Also, it could be investigated whether individual educational plans could be modified to allow for certain types of adult interventions such as probes for special needs students.

As well as exploring interventions for special education students, the potential impact of training teachers in the art of academic language instruction could be evaluated.

Would instruction that explains the complexity of academic language positively affect the academic performance not only of English language learners, but also of African American students with limited exposure to text? In many ways, these students' experience prior to entering school mimics those of second language learners in that their everyday oral communication does not adequately prepare them for the rigors of content area language that they encounter in schools. Instead of beginning academic language instruction in middle and secondary schools, inquiry could be undertaken that would explore the efficacy of implementing such instruction in elementary schools.

Based on Liane's and Danielle's responses, there could be inquiry into whether the clinical interview with its individualized format could be counter to the cultural preferences of some African American students. The current research supports the notion that interviews are a culturally acceptable educational tool for these students (LadsonBillings, 1997).

Finally, indicators suggest that parental support influenced the degree of success of two students in the study. Kevin and Sam did well on both assessments and their utilization of a variety of strategies assumed a level of automaticity that was not as evident in the other children. Future explorations could investigate the impact of consistent parental coaching on African American students’ ability to perform well on standardized tests and assessments in general.

These potentially transformative avenues of inquiry involving the clinical interview could add to the body of knowledge on accountability testing. Although some African American students perform well on these assessments, many others do not.

Educational research will continue seeking for ways to remedy this problem.
Investigating ways of using complementary assessments like the clinical interview could be part of this inquiry.

## Significance of Study

In designing and conducting this study, I provided another perspective to the problem of African American academic achievement that I experienced as a teacher and mother, and that is evident in our country today. This issue has been the topic of many studies, and has been written about in many practitioner articles. It has been the focus of countless reform initiatives and continues to dominate our national discourse about education.

In this study I explored a gap in the current literature concerning accountability testing and its effects on African American students. Specifically I explored using a complementary assessment, the clinical interview, as a means of learning more about African American students’ mathematical capabilities.

## Conclusion

Prior to conducting this study, I had experienced firsthand the ability of the clinical interview to reveal previously unrecognized information about students’ mathematical ability. I had seen how surprised teachers had been by these revelations and how they had resulted in some teachers viewing their students' mathematical capabilities through more positive lenses. These teachers welcomed the ability to see what their students really knew about mathematics. The clinical interview provided them with that opportunity.

As a result of these previous experiences with the clinical interview, I had anticipated that all the students in the study, especially those with less mathematical proficiency, would perform better on the clinical interview items. This expectation was not fulfilled. Different factors: the good quality of the instruction that the students received; positive teacher expectations; and the fact that there were more open-ended items on the clinical interview, contributed to these results.

Nonetheless, the clinical interviews did replicate my previous experiences with them in terms of what they revealed about the students' mathematical abilities. Students' conceptual knowledge as well their preferred cognitive strategies emerged with microscopic clarity. Thus, the clinical interview, used as a "window" into African American students' mathematical abilities was successful in accurately portraying the extent of these boys' mathematical literacy as applied to the evaluated concepts.

Upon reflection, I realized that my anticipation of the teachers' responses had been influenced by the observations from the past studies in which I had participated. I had expected that the students' performances on the clinical interview would have provided the teachers with significant new information about their students' mathematical ability.

Although some teachers did learn new information about their students’ use of mental math, whether or not this knowledge was recognized, the clinical interviews in this study confirmed and supported the teachers' views of their students. This could be attributed to the fact that they were experienced teachers of high caliber. As such they
were able to make informed evaluations about their students' mathematical abilities that were generally confirmed by the findings.

The participants - teachers and students - were completely cooperative and helpful. Every one of the students worked hard and consistently during the clinical interviews and maintained their focus throughout. As they had just finished three days of standardized testing, this was highly commendable. The teachers were invested and all sought to bring value to the study. They were accustomed to engaging in reflection about their students and their practice. They were comfortable sharing their thoughts with the clinical interviewer and me.

In addition to the participants, the clinical interviewer was also critical to the success of the study. Her experience was invaluable especially in the area of probes. Her ability to decide whether or not to probe, to choose the kind of probe that would elicit the best information from the student impacted the results of the clinical interview. The participants and the clinical interviewer were valuable contributors not only to the study, but also to its findings and their significance. As a result of their cooperation and effort, the study provided another perspective into the academic performance of African American students.

A slogan from the National Association for the Advancement of Colored People (NAACP) reads, "A mind is a terrible thing to waste." This has always been and will remain true, but its veracity has assumed a new urgency in the face of today's global economics (Hunter \& Bartee, 2003). In order for the United States to maintain and
improve the standard of living for its citizens, it needs to have the efforts of everyone who is capable (Barton \& Coley, 2008; Hale, 2004).

As Hale (2004) asserts, the education of low-income and middle-class African Americans affects all of society. All children need to be prepared to assume leadership roles in society (Hale, 2004). Providing requisite strategies and skills, and utilizing complementary assessments like the clinical interview are some factors that could facilitate this process.

## References

American Educational Research Association (2000). Position statement of the American Educational Research Association concerning high-stakes testing in preK-12 education. Educational Researcher, 29(8), 24-25.
Bainbridge, W. L., \& Lasley II, T. J. (2002). Demographics, diversity, and K-12 accountability: the challenge of closing the achievement gap. Education and Urban Society, 34(4), 422-437.
Baker, J. M. (1999). Teacher-student interaction in urban at-risk classrooms: differential behavior, relationship, quality, and student satisfaction with school. The Elementary School Journal, 100(1), 57-70.
Barthe, P., \& Mitchell, R. (2006). Standardized tests and their impact on schooling: Q\&A. www.centerforpubliceducation.org/Main-Menu/Instruction/High-stakes-testing-and-effects-on-instruction-At-a-glance/
Baker, J. M. (2008). Achievement testing in U. S. elementary and secondary schools. New York, NY: Peter Lang Publishing, Inc.
Barton, P. E. (1999). Too much testing of the wrong kind; too little of the right kind in K12 education. Princeton, NJ: Educational Testing Service. Available: www.ets.org/research/pic (for downloading) - March 1999.
Barton, P. E. (2003). Parsing the achievement gap -baselines for tracking progress. Princeton, NJ: Educational Testing Service.
Barton, P. E., \& Coley, R. J. (2008). Windows on achievement and inequality. Princeton, NJ: Educational Testing Services.
Benner, A. D., \& Mistry, R. S. (2007). Congruence of mother and teacher educational expectations and low-income youth's academic competence. Journal of Educational Psychology, 99(1), 140-153.
Berliner, D. (2009). Poverty and potential: Out of school factors and school success, a policy report published jointly by EPRU (at the University of Arizona) and Education and the Public Interest Center (at the University of Colorado, Boulder.
Bloom, B. S. (1956). Taxonomy of Educational Objectives, Handbook 1: The Cognitive Domain. New York, New York: David McKay Co. Inc.
Boaler, J. (2003). When learning no longer matters- standardized testing and the creation of inequality. Phi Delta Kappan, 84(7), 502-506.
Bourdieu, P. (1986). The forms of capital. In J. E. Richardson (Ed.), Handbook of theory of research for the sociology of education. (R. Nice, Trans.) Connecticut: Greenwood Press.
Bourdieu, P., \& Passeron, J. C. (1977). Reproduction in education, society, and culture. (R. Nice, Trans.) Beverly Hills: Sage.
Bourque, M. L. (2005). Leave no standardized test behind. In R. P. Phelps (Ed.), Defending standardized testing (pp. 227-254). Mahwah, NJ: Lawrence Erlbaum Associates.
Boykin, A. W., Tyler, K. M., Watkins-Lewis, K., \& Kizzie, K. (2006). Culture in the sanctioned classroom practices of elementary school teachers serving low-income

African American students. Journal of Education for Students Placed at Risk, 11(2), 161-173.
Brown v. Board of Educ., 347 U.S. 483 (1954).
Causey-Bush, T. (2005). Keep your eye on Texas and California: A look at testing, school reform, No Child Left Behind, and implications for students of color. The Journal of Negro Education, 74(4), 332-343.
Center for Teaching Excellence, "Improving your test questions." www.cte.illinois.edu/testing/exam/test_que.html
Charles, K., \& Nason, R. (2000). Young children's partitioning strategies. Educational Studies in Mathematics 43(2), 191-221.
Cizek, G. J. (2001). Unintended consequences of high stakes testing - P-12. Educational Measurement: Issues and Practice, 20(4), 19-27.
Cizek, G. J. (2005). High-stakes testing: Contexts, characteristics, critiques and consequences. In R. P. Phelps (Ed.) Defending Standardized Testing, (pp. 2354). Mahwah, NJ: Lawrence Erlbaum Associates.

Committee on Incentives and Test-Based Accountability in Public Education at the National Research Council, 2011.
Cooper, B. (1998). Assessing national curriculum mathematics in England: Exploring children's interpretation of key Stage 2 tests in clinical interviews. Educational Studies in Mathematics, 35, 19-49.
Crozier, G. (2005). ‘There’s a war against our children- black educational underachievement revisited. British Journal of Sociology of Education, 26(5), 585-598.
Darling-Hammond, L. (2007). Third Annual Brown Lecture in Education Research -The flat earth and education: how America's commitment to equity will determine our future. Educational Researcher, 36(6), 318-334.
Denzin, N. (1978). The research act: A theoretical introduction to sociological methods. New York, NY: McGraw-Hill.
Denzin, N. (2006). Sociological methods: A sourcebook. Aldine Transaction. ISBN 9780-202308401. (5 ${ }^{\text {th }}$ Edition)
Dillon, S. (2009, April 28). 'No Child’ law is not closing a racial gap. The New York Times. Retrieved November 19, 2011 from http://www.nytimes.com/2009/04/29/education/29scores.html
diSessa, A. A. (2007). An interactional analysis of clinical interviewing. Cognition and Instruction, 25(4), 523-565.
Draper, R. J., \& Siebert, D. (2004). Different goals, similar practices: Making sense of the mathematics and literacy instruction in a standards-based mathematics classroom. American Educational Research Journal, 41(4), 927-962.
Dunphy, E. (2006). The development of young children's number sense through participation in sociocultural activity: Profiles of two children. European Early Childhood Education Research Journal, 14(1), 57-76.
Earthman, G. I. (2004). Prioritization of 31 criteria for school building adequacy. Retrieved November 19, 2011 from www.schoolfunding.info/policy/facilities/ ACCUfacilities-report1-04.pdf

Elliott, J. (2003). Dynamic assessment in educational settings: realizing potential. Educational Review, 55(1), 15-32.
Empson, S. B., \& Turner, E. (2006). The emergence of multiplicative thinking in children's solutions to paper folding tasks. The Journal of Mathematical Behavior, 25(1), 46-56.
English, F. W. (2002). On the intractability of the achievement gap in urban schools and the discursive practice of continuing racial discrimination. Education and Urban Society, 34(3), 298-311.
ETS Policy Notes, 2009.
Evans, R. (2005). Reframing the achievement gap. Phi Delta Kappan, April 2005, 582589.

Evans, R. I. (1973). Jean Piaget - The man and his ideas. New York, NY: E. P. Dutton \& Company, Inc.
Farkas, G. (2003). Racial disparities and discrimination in education: what do we know, how do we know it, and what do we need to know? Teachers College Record, 105(6), 1119-1146.
Ferguson, A. A. (2001). Bad boys: Public schools in the making of black masculinity. Ann Arbor, MI: The University of Michigan Press.
Ferguson, R. F. (2003). Teachers' perceptions and expectations and the black-white test score gap. Urban Education, 38, 460-507.
Ferguson, R. F., Clark, R., \& Stewart, J. (2002). Closing the achievement gap in suburban and urban school communities, a policy issue from North Central Regional Lab, Naperville, Illinois.
Fleming, J. (2000). Affirmative action and standardized test scores. The Journal of Negro Education, 69(1/2), 27-37.
Flowers, L. A. (2007). Recommendations for research to improve reading achievement for African American students. Reading Research Quarterly, 42(3), 424-428.
Fraivillig, J. L., Murphy, L. A., \& Fuson, K. C. (1999). Advancing children's mathematical thinking in everyday mathematics classrooms. Journal for Research in Mathematics Education, 30(2), 148-170.
Frankenberg, E., Lee, C., \& Orfield, G. (2003). A Multi-racial society with segregated schools: Are we losing the dream? The Civil Rights Project at Harvard University. Available: www.civilrightsproject.harvard.edu
Friedman, T. L. (2005). The world is flat: A brief history of the twenty-first century. Union Square, New York: Farrar, Straus, and Giroux.
Freud, S. (1938). The history of the psychoanalytic movement. In A. A. Brill (Ed.), Basic Writings of Sigmund Freud (pp. 933-977).
Gardner III, R., \& Miranda, A. H. (2001). Improving outcomes for urban African American students. The Journal of Negro Education, 70(4), 255-263.
Gardner, III, R., \& Talbert-Johnson, C. (2001). School reform and desegregation: The real deal or more of the same? Education and Urban Society, 33(1), 74-87.
Garibaldi, A. (2007). The educational status of African American males in the $21^{\text {st }}$ century. The Journal of Negro Education, 76(3), 324-333.
Gerstner, Jr., L. V. (2002, March 14). The tests we know we need. The New York Times.
www.nytimes.com/2002/03/14/opinion/the-tests-we-know-we-need.html Ginsburg, H. P. (1997). Entering the child's mind: The clinical interview in pyschological research and practice. New York, NY: Cambridge University Press.
Ginsburg, H. P., \& Pappas, S. (2004). SES, ethnic, and gender differences in young children's informal addition and subtraction: A clinical interview investigation. Journal of Applied Developmental Psychology, 25(2), 171-193.
Ginsburg, H. P., Mast, J. V., \& Snow, M. (2010). Child study/lesson study: A catalyst for teacher curiosity. In C. J. Craig \& L. F. Deretchin (Eds.), Cultivating curious and creative minds: The role of teachers and teacher educators, Part I. The Teacher Education Yearbook XVIII. Lanham, MD: Rowman \& Littlefield Education.
Goldin, G. A. (2000). A scientific perspective on structured, task-based interviews in mathematics education research. In A. E. Kelly \& R. A. Lesh (Eds.), Handbook of research design in mathematics and science education (pp. 517-545). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Gordon, E. W. (1995). Toward an equitable system of educational assessment. The Journal of Negro Education, 64(3), 360-372.
Greene, J. P., \& Winters, M. A. (2006). Getting farther ahead by staying behind: A second-year evaluation of Florida's policy to end social promotion. (Manhattan Institute for Policy Research Civic Report, No. 49). www.manhattaninstitute.org/html/cr_49.htm
Grigorenko, E., \& Sternberg, R. J. (1998a). Dynamic testing. Psychological Bulletin, 124(1), 75-111.
Grigorenko, E., \& Sternberg, R. J. (1998b). Dynamic testing: Erratum. Psychological Bulletin, 124(3), 443.
Grodsky, E., Warren, J. R., \& Telfts, E. ( 2008). Testing and social stratification in American education. Annual Review of Sociology, 34, 385-404.
Groth, R. E. (2005). An investigation of statistical thinking in two different contexts: Detecting a signal in a noisy process and determining a typical value. The Journal of Mathematical Behavior, 24(2), 109-124.
Hale, J. E. (2001). Learning while Black: Creating educational excellence for African American children. Baltimore, MD: The Johns Hopkins University Press.
Hale, J. E. (2004). How schools shortchange African American children. Educational Leadership, November 2004, 34-39.
Haney, W. (1993). Testing and minorities. In L. Heis and M. Fine (Eds.), Beyond voices: Class, race, and gender in U. S. schools (pp.45-59). New York: State University of New York Press.
Harris, D. M. (2012). Varying teacher expectations and standards: Curriculum differentiation in the age of standards-based reform. Education and Urban Society 44(2), 128-150.
Haskins, J. (1998). The dream and the struggle - separate but not equal. New York: Scholastic Inc.
Haywood, H. C., \& Lidz, C. S. (2007). Dynamic assessment in practice - clinical and
educational applications. New York, NY: Cambridge University Press. Hebert, T. P., \& Reis, S. M. (1999). Culturally diverse high-achieving students in an urban high school. Urban Education, 34(4), 428-457.
Hill, P. T., Guin, K., \& Celio, M. B. (2003). The chasm remains. Education Next, Spring 2003.

Hout, M., \& Elliott, S. W. (Eds). (2011). Incentives and test-based accountability in education. Washington, D.C.: The National Academies Press. www.nap.edu
Hunter, R. C., \& Bartee, R. (2003). The achievement gap: Issues of competition, class, and race. Education and Urban Society, 35(2), 151-160.
Huntley, M. A., Marcus, R., Kahan, J., \& Miller, J. L. (2007). Investigating high-school students’ reasoning strategies when they solve linear equations. Journal of Mathematical Behavior, 26, 115-139.
Inoue, N. (2005). The realistic reasons behind unrealistic solutions: the role of activity in word problem solving. Learning and Instruction, 15(1), 69-83.
Irvine, J. J. (1990). Black Students and School Failure. Policies, Practices, and Prescriptions. Westport, CT: Greenwood Press, Inc.
Jacobs, B. (2007). No Child Left Behind’s emphasis on 'teaching to the test’ undermines quality teaching. Endeavors, 10(17).
www.education.umd.edu/infofor/alumni=friends/endeavors/endeavors0712/endea vors0712.pdf
Jencks, C., \& Phillips, M. (1998). The Black-White test score gap: An introduction. In C. Jencks and M. Phillips (Eds.), The Black-White Test Score Gap (pp. 1-51). Washington, DC: The Brookings Institution.
Johnston, R. C., \& Viadero, D. (2000). Unmet promise: Raising minority achievement. Education Week, 19(27), 18-19.
Kantor, H., \& Brenzel, B. (1992). Urban education and the "truly disadvantaged:" The historical roots of the contemporary crisis. Teacher College Record, 94 (2), 278314.

Koretz, D. M. (2002). Limitations in the use of achievement tests as measures of educator's productivity. Journal of Human Resources, 37(4), 752-777.
Koretz, D. (2008). Measuring up: What educational testing really tells us. Cambridge, MA: The President and Fellows of Harvard College.
Kozol, J. (2005). The shame of the nation: The restoration of apartheid schooling in America. New York: Three Rivers Press.
Kuklinski, M. R., \& Weinstein, R. S. (2001). Classroom and developmental differences in a path model of teacher expectancy effects. Child Development, 72(5), 1554-1578.
Kunjufu, J. (2002). Black students: Middle class teachers. Chicago, Illinois: African American Images.
Ladson-Billings, G. (1994). The dreamkeepers: Successful teachers of African American children. San Francisco, CA: Jossey-Bass Inc., Publishers.
Ladson-Billings, G. (1997). It doesn't add up: African American students' mathematics achievement. Journal for Research in Mathematics Education, 28(6), 697-708.

Landsman, J. (2004). Confronting the racism of low expectations. Educational Leadership, 62(3) 28-32.
Lan, W., \& Lanthier, R. (2003). Changes in students’ academic performance and perceptions of school and self before dropping out of schools. Journal of Education for Students Placed at Risk, 8(3), 309-332.
Lareau, A. (2003). Unequal childhoods: Class, race, and family life. Berkeley, CA: University of California Press.
Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress towards equity? Educational Researcher, 31(1), 3-12.
Linn, R. L. (1995). Assessment-based reform: Challenges to educational measurement. Princeton, NJ: Educational Testing Service.
Linn, R. L. (2001). A century of standardized testing: Controversies and pendulum swings. Educational Assessment, 7(1), 29-38.
Linn, R. L. (2006). Educational accountability systems. (CSE Technical Report 687). National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/University of Colorado at Boulder.
Linn, R. L. (2007). Validity of Inferences from test-based educational accountability systems. Journal of Personnel Evaluation in Education, 19, 5-15.
Linn, R. L., \& Hambleton, R. K. (1991). Customized tests and customized norms. Applied Measurement in Education, 4(3), 185-207.
Lubienski, S.T., \& Bowen, A. (2000). Who's counting? A survey of mathematics education research 1982-1998. Journal for Research in Mathematics Education, 31(5), 626-633.
McKnight, K. (2011). Opting out of NCLB testing. www.katherinemcknight.com/opting-out-of--nclb-testing
Madon, S. J., Jussim, L., \& Eccles, J. (1997). In search of the powerful self-fulfilling prophecy. Journal of Personality and Social Psychology, 72, 791-809.
Malone, J., Douglas, G., Kissane, B., \& Mortlock, R. (1980). Measuring problem solving ability. In S. Krulik \& R. Reys (Eds.), Problem solving in school mathematics (1980 Yearbook of the National Council of Teachers of Mathematics) (pp. 204-215). Reston, VA: National Council of Teachers of Mathematics.
Martinez, M. (2011). TAKS test taking a bite out of budget? www.valleycentral.com/news/storyaspx?id=611612
Mathis, W. J. (2005). Bridging the achievement gap: A bridge too far? Phi Delta Kappan, April 2005, 590-593.
Matteson, S. M. (2006). Mathematical literacy and standardized mathematical assessments. Reading Psychology, 27, 205-233.
Miller, D. M., Linn, R.L., \& Gronlund, N. E. (2009). Measurement and assessment in teaching. Upper Saddle River, NJ: Pearson Education Inc.
Miller, E., \& Almon, J. (2009). Crisis in the Kindergarten: Why children need to play in school. www.allianceforchildhood.org/restoring_play
Milliken v. Bradley, 418 U. S. 717 (1974).
Mitchell, R. (2006). A guide to standardized testing: The nature of assessment.
www.centerforpubliceducation.org/Main-Menu/Evaluating-performance/A-guide-to-standardizedtesting-The-nature-of-assessment
Montague, M., \& Rinaldi, C. (2001). Classroom dynamics and children at risk: a follow-up. Learning Disability Quarterly, 24(2), 75-83.
Moses, R. P. (2001). Radical equations: Math literacy and civil rights. Boston, MA: Beacon Press.
Moseley, B., \& Brenner, M. E. (2009). A comparison of curricular effects on the integration of arithmetic and algebraic schemata in pre-algebra students. Instructional Science, 37, 1-20.
Mulvenon, S. W., Connors, J. V., \& Lenares, D. (2011). Impact of accountability and school testing on students: Is there evidence of anxiety? www.eric.ed.gov, ED 460155, November, 2011.
National Assessment of Educational Progress, 2009. www.nces.ed.gov/nationsreportcard
National Center for Education Statistics, 2009. The Nation's Report Card. US Department of Education - Institute of Education Sciences. [On-line]. www.nces.ed.gov/nationsreportcard/mathematics.
National Council of Churches Committee on Public Education and Literacy. Ten moral concerns in the implementation of the No Child Left Behind Act. www.ncccusa.org/pdfs/LeftBehind.pdf
National Council of Teachers of Mathematics. (2000). Principles and Standards for School Mathematics. Available: www.nctm.org/standards/content.aspx?id=16909
New Jersey Department of Education, 2010-2011; 2009-2010
Nichols, S. L., \& Berliner, D. C. (2008). Why has high-stakes testing so easily slipped into contemporary American life? Phi Delta Kappan, May 2008, 672-676.
Noguera, P. A. (2003). The trouble with black boys: the role and influence of environmental and cultural factors on the academic performance of African American males. Urban Education, 38(4), 431-459.
North Central Regional Educational Laboratory. (2002, December). Closing the achievement gap in suburban and urban school communities (Issue 13). Naperville, IL: Ferguson, R. F., Clark, R., \& Stewart, J.
Nowak, J. D., \& Gowin, D. B. (1984). Learning how to learn. New York, NY: Cambridge University Press.
O’Connor, C. (1997). Dispositions towards (collective) struggle and educational resilience in the inner city: A case analysis of six African-American high school students. American Educational Research Journal, 34(4), 593-629.
Ogbu, J. U. (1993). Frameworks - Variability in minority school performance: A problem in search of an explanation. In E. Jacob and C. Jordan (Eds.), Minority education: Anthropological Perspectives (pp. 83-111). New Jersey: Ablex.
Ogbu, J., \& Simons, H. D. (1998). Voluntary and involuntary minorities: a culturalecological theory of school performance with some implications for education. Anthropology \& Education Quarterly, 29(2), 155-188.
Ohanian, S. (2002). Collateral vomitage. www.susanohanian.org, March 14, 2002.
Olson, L. (2001). Study questions reliability of single-year test-score gains. Education Week, 20(37), 9.

Orfield, G., \& Wald, J. (2000). Testing, testing. The Nation, June $5^{\text {th }}$, 38-40.
Patton, M. Q. (2002). Qualitative research \& evaluation methods, $3^{\text {rd }}$ Edition. Thousand Oaks, CA: Sage Publications.
Phelps, R. P. (2002). Estimating the costs and benefits of educational testing programs. www.education-consumers.org/research/phelps2.htm
Phelps, R. P. (2012). The effect of testing on achievement: Meta-analyses and research Summary, 1910-2010. International Journal of Testing, 12(1), 21-43.
Piaget, J. (1952). Jean Piaget. In E. G. Boring, H. Werner, H. S. Langfeld, \& R. M. Yerkes (Eds.), A history of psychology in autobiography (pp. 237-256). Vol. IV. Worcester, MA: Clark University Press.
Piaget, J. (1969). Genetic epistemology. The Columbia Forum, XII (3).
Piaget, J. (1978). Success and understanding. Cambridge, MA: Harvard University Press. Plessy v. Ferguson, 163 U.S. 537 (1896).
Public Agenda (2002). Where's the backlash? Students say they don't fret standardized tests. www.publicagenda.com/press-releases/wheres-backlash-students-say-they-don't-fret-standardized-tests
Ramful, A., \& Olive, J. (2008). Reversibility of thought: An instance in multiplicative tasks. The Journal of Mathematical Behavior, 27, 138-151.
Ravitch, D. (2007, October 3). Get Congress out of the classroom. The New York Times. www.nytimes.com/2007/10/03/opinion/03/ravitch.html
Ravitch, D. (2010). The death and life of the great American school system. New York, NY: Basic Books.
Rhee, M. (2011, May 24). Accommodate don't discriminate. The Huffington Post. www.huffingtonpost.com/michelle-rhee/accommodate-don't-discriminate
Rist, R. C. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. Harvard Educational Review, 40(3), 411-451.
Rizga, K. (2011, May 19). What standardized tests miss. Mother Jones. www.motherjones.com/politics/2011/05/students-teachers-obama-NCLB
Roscigno, V. (1998). Race and the reproduction of educational disadvantages. Social Forces, 76(3), 1033-1061.
Rosenthal, R., \& Jacobson, L. (1968, 1992). Pygmalion in the classroom -teacher expectations and pupils' intellectual development. Wales, UK: Crown House Publishing Limited.
Rosenthal, D. (2011, April 4). Don't make educational policy based on cheaters. www.hlpronline.com
Rothstein, R. (2004). Class and schools: Using social, economic, and educational reform to close the black-white achievement gap. Washington, DC: Economic Policy Institute.
Rothstein, R. (2004). The achievement gap. Educational Leadership, November, 2004, 40-43.
Rothstein, R. (2005). The limits of testing. American School Board Journal, February, 2005, 34-37.
Rousseau, C., \& Tate, W. F. (2003). No time like the present: Reflection on equity in school mathematics. Theory into Practice, 42(3), 210-216.

Rowe, M. (1974). Wait-time and rewards as instructional variables. Journal of Research in Science Teaching, 11(2), 81-94.
Rubie-Davis, C. M. (2010). Teacher expectations and perceptions of student attributes: Is there a relationship? British Journal of Educational Psychology, 80, 121-135.
Sacks, P. (1999). Standardized minds: The high price of America's testing culture and what we can do to change it. New York, NY: Perseus Publishing.
Schleppegrell, M. J., Achugar, M., \& Oteiza, T. (2004). The grammar of history: Enhancing content-based instruction through a functional focus. TESOL Quarterly, 38(1), 67-93.
Schleppegrell, M. J., \& O’Hallaron, C. L. (2011). Teaching academic language in L2 secondary settings. Annual Review of Applied Linguistics, 31, 3-18.
Schoenfeld, A. H. (2002). Making mathematics work for all children: Issues of standards, testing, and equity. Educational Researcher, 31(1), 13-25.
Schorr, R. Y., Firestone, W. A., \& Monfils, L. (2003). State testing and mathematics teaching in New Jersey: The effects of a test without other supports. Journal for Research in Mathematics Education 34(5), 373-405.
Sfard, A., \& Linchevski, L. (1994). The gains and the pitfalls of reification: The case of algebra. Educational Studies in Mathematics, 26, 191-228.
Shepard, L. A. (2000). The role of assessment in a learning culture. Educational Researcher, 29(7), 4-14.
Singh, P. (2000). Understanding the concepts of proportion and ratio constructed by two grade six students. Educational Studies in Mathematics, 43, 271-292.
Slavin, R. E., \& Madden, N. A. (2006). Reducing the gap: Success for All and the achievement of African American students. The Journal of Negro Education, 75(3), 389-400.
Solley, B. A. (2007). On standardized testing: An ACEI position paper. Childhood Education, 84(1), 31.
Stinson, D. W. (2006). African American male adolescents, schooling, and mathematics: Deficiency, rejection, and achievement. Review of Educational Research, 76(4), 477-506.
Strauss, V. (2010, June 16). Ravitch: Are we in an era of national stupidity? The Washington Post. www.washingtonpost.com
Strauss, V. (2011, April 26). Unanswered questions about standardized tests. The Washington Post. www.washingtonpost.com
Stringfield, S. (2007). Improvements in academic achievement among African American students over time: national data and an urban case study. Journal of Negro Education, 76(3), 306-315.
Strutchens, M., Lubienski, S. T., McGraw, R., \& Westbrook, S. K. (2004). NAEP findings regarding race/ethnicity: Students’ performance, school experiences, attitudes/beliefs, and family influences. In P. Kloosterman., \& F. T. Lester, Jr. (Eds.), Results and interpretations of the 1990 through 2000 mathematics assessments of the National Assessment of Educational Progress (pp. 269-304). Reston, Virginia: NCTM.

Toppo, G. (2011, March 17). When test scores seem too good to believe. USA Today. www.usatoday.com/news/education/2011-03-06-school-testing-N.htm
Tyler, K. M., \& Boelter, C. M. (2008). Linking black middle school students’ perceptions of teachers' expectations to academic engagement and efficacy. Negro Educational Review, 59(1/2), 27-44
U.S. Department of Education (1999). Taking responsibility for ending social promotion. www.ed.gov, May 1999.
U.S. Department of Education (2004). Testing: Frequently asked questions. www.ed.gov, November 17, 2004.
Valli, L., \& Croninger, R. (2011). High quality teaching of foundational skills in mathematics and reading. www.drdc.uchicago.edu/community/project.phtml?projectID=79
Viadero, D. (2000). Even in well-off suburbs, minority achievement lags. Education Week, 19(27), 22-23.
Villegas, A., \& Lucas, T. (2002). Educating culturally responsive teachers: A coherent approach. Albany, NY: State University of New York Press.
Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.
Vygotsky, L. S. (1986). Thought and language. Cambridge, MA: MIT Press.
Walberg, H. J. (2011). Stop the war against standardized tests. Defining Ideas, May 2011. www.hoover.org/publications/defining-ideas/article179431
Walker, E. N. (2006). Urban high school students’ academic communities and their effects on mathematics success. American Educational Research Journal, 43(1), 43-73.
Wang, L., Beckett, G. H., \& Brown, L. (2006). Controversies of standardized assessment in school accountability reform: A critical synthesis of multidisciplinary research evidence. Applied Measurement in Education, 19(4), 305-328.
Weisglass, J. (2002). Inequity in mathematics education: Questions for educators. The Mathematics Educator, 12(27), 34-39.
William, D. (2010). Standardized testing and school accountability. Educational Psychologist, 45(2), 107-122.
Xueqin, I. (2010, December 8). The test Chinese schools still fail. The Wall Street Journal. www.wsj.com/search/term.html?Keywords=The\ 5test\ \%Chinese\ scho ols\%20 still\%20fail
Yeh, S. S. (2005). Limiting the unintended consequences of high-stakes testing. Educational Policy Analysis Archives, 13(43), 1-21.
Zagursky, E. (2011). Smart? Yes. Creative? Not so much. www.wm.edu/research/ideation/professions/smart-yes.-creative-not-somuch.5890.php
Zembat, I. O. (2008). Pre-service teachers' use of different types of mathematical reasoning in paper-and-pencil versus technology-supported environments. International Journal of Mathematical Education in Science and Technology, 39(2), 143-160.

Zhao, Y. (2010). John Richard Schrock: Why doesn't China get off the teach-to-the-test system? www.zhaolearning.com/2010/12/29/john-richard-schrock-why-doesn’t-china-get-off-the-teach-to-the-test-system

Appendices
Appendix A- Student Participants

| Student | School | Age | Teacher |
| :--- | :--- | :--- | :--- |
| Kevin | B School | 7 years 8 months | Liane |
| Sam | B School | 7 years 11 months | Liane |
| Oliver | E School | 7 years 8 months | Danielle |
| David | E School | 7 years 11 months | Danielle |
| Martin | E School | 7 years 9 months | Rae |
| Jake | E School | 7 years 10 months | Dahlia |
| Jake | E School | 7 years 10 months | Becca |
| Philip |  |  | Mary |

Appendix B - Teacher Participants

| Teacher | Position | Total Years of Experience | Years in Current Position | Degrees/Certificates <br> Earned |
| :---: | :---: | :---: | :---: | :---: |
| Liane | $2^{\text {nd }}$ Grade <br> Teacher | 35years | 15 years | BA in Education <br> Certified PreK-3 \& K-8 <br> MAT in Teaching |
| Danielle | $2^{\text {nd }} \text { Grade }$ <br> Teacher | 15 years | 9 years | BA in Education <br> Certified K-8 <br> Masters in Reading <br> 30 credits above that <br>  <br> Supervision |
| Dahlia | $2^{\text {nd }} \text { Grade }$ <br> teacher | 23 years | 8years | BA in Education Nursery and Elementary |


|  |  |  |  | certificates <br> Masters in Education |
| :---: | :---: | :---: | :---: | :---: |
| Becca | Special <br> Education <br> teacher | 16 years | 11years | BA in Education+ 20 <br> credits <br> Preschool, <br> elementary, and K-8 <br> certificates <br> Wilson Reading <br> Training |
| Rae | $2^{\text {nd }} \text { Grade }$ <br> Teacher | 30 years | 12 years | BA in Education w/concentration in <br> Early Childhood <br>  <br> Elementary K-8 <br> certificates <br> Masters in Teaching <br> +30 credits |


| Mary | 2nd Grade <br> Teacher | 15.5 years | 3 years |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | BA in Education <br> Elementary K-8 <br> certificate |
| Masters in Education |  |  |  |  |
| +30 credits |  |  |  |  |

## Appendix C- Mathematical Concepts for Teacher Interview \#1

Here is the list of concepts that we will be discussing during Teacher Interview \#1.

> I will be asking you to evaluate whether your student(s) is (are) performing at a secure, developing, or beginning level in these areas.

I decide to use the scale from the Chicago Math Program as I hoped that this would be easier since it is so familiar to you.

Thank you!!

## MATHEMATICAL CONCEPTS

- Part/Whole
- Measurement
- Money
- Probability
- Elapsed Time
- Telling Time
- Patterns
- Symmetry
- Bar Graphs, Grids, Charts
- Place Value to 100 s
- Basic + \& - facts
- Word Problems


## Appendix D - Teacher Survey \# 1

Student's Name
Evaluating Teacher

Please check the level of proficiency attained by this student in each category below.
Secure Developing Beginning

Part/Whole
Measurement
Money
Probability
Elapsed Time
Telling Time
Patterns
Symmetry
Bar Graphs, Grids,
Charts
Place Value to 100s
Basic + \& - facts
Word Problems

Please check the responses that best describe this student's consistent strategies for solving problems.

In completing basic addition and subtraction algorithms, does this student use:

## YES

NO

Mental math $\qquad$

Fingers

Paper \& pencil $\qquad$

Manipulatives such as cubes or counters? $\qquad$

In explaining his thinking when solving problems, does this student prefer:
YES
NO
Oral Explanations $\qquad$

Written Explanations $\qquad$

Diagrams $\qquad$

A combination of these methods? $\qquad$

Please identify the resources you used to make your evaluations by checking any of the following that apply.

Workbook pages

Open-ended questions

Pre-tests

End-of-unit tests

Math games

Teacher created assessments

Teacher/student conferences

Teacher/student observations

Other (please describe in the space below)

## Appendix E - Post Interview Sheet

# Study Name: Using the Clinical Interview as a Complementary Assessment for Minority Elementary Students to Determine Their In-Depth Understanding of Mathematical Concepts. 

## Principal Investigator: Nicola Crisp

## Teacher Interview \#2

Remind teacher of purpose of this second interview.

This interview is for you to discuss the results of your student's clinical interview which you have now had a chance to review.

Possible probes to help discussion:

What did you anticipate about student's responses?

What surprised you about the student's responses?

What did you notice about the student's verbal responses?

What did you notice about the student's body language?
Did the body language contradict the verbal response?

Were your prior evaluations (Teacher interview \#1/Teacher survey \#1) supported or negated?

How would you summarize what you learned from the clinical interviews about your student's mathematical knowledge?

Have your views about your student's mathematical knowledge changed or remained the same after viewing his performance in the clinical interview?

Appendix F - Clinical Interview Protocol

## STUDY NAME: USING THE CLINICAL INTERVIEW AS A COMPLEMENTARY ASSESSMENT FOR MINORITY STUDENTS TO DETERMINE THEIR IN-DEPTH UNDERSTANDING OF MATHEMATICAL CONCEPTS.

## CLINICAL INTERVIEW PROTOCOL

Jon has 9 Yugioh cards. He gave 3 to his best friend Carl. How many cards does he have left?

Show the problem and the answer as a number model. You may draw pictures before writing the number model.

Here are 3 pieces of yarn of different lengths. Which string would measure about 1cm? Explain how you decided on your choice of string.

Your tee-ball practice lasts 45 minutes. If it begins at $3: 30$ p.m., at what time does your practice end? (A student clock or a visual representation will be provided for the students).

36, 34, 32, 30, __ _ , 24, __. Fill in the missing numbers. Explain your answer by telling about the pattern that you found.

At the book fair, the next book in the series by Jamar's favorite author cost \$1.95. He has 6 quarters, 6 dimes, 4 nickels, and 5 pennies. List 2 different ways he could pay for the book using these coins.

Draw 2 spinners on the circles that are provided. Each spinner needs to have 3 colors - red, yellow, and green. Draw each spinner so that each has equal amounts of red, yellow, and green spaces.


Use the bar graph above to answer the following questions:

Who read twice as many books as Tyesha?

How many books were read in all?

What was the greatest amount of books read by one person?

How many less books did Jamal read than Trent?

Here are some pictures of different objects. Some are symmetrical and some are not symmetrical. Please identify which objects are symmetrical and which are not. Explain the reasons for your answers. Can you name any other object, (not shown here), that you believe to be symmetrical? Give or show reasons for your choice.

On the coordinate grid below, identify the locations of the rectangle, hexagon, triangle, rhombus, trapezoid, and the pentagon. (In the grid provided to students, the abbreviations will be explained as to what they represent. The abbreviations are: $R=$ Rectangle, $H=$ Hexagon, $T=$ Triangle, $R H=$ Rhombus, $T Z D=$ Trapezoid, $P=$ Pentagon).

| 10 |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 9 |  | H |  |  |  |  |
| 8 |  |  |  |  |  | P |
| 7 |  |  | T |  |  |  |
| 6 |  |  |  |  |  |  |
| 5 |  |  |  | RH |  |  |
| 4 | R |  |  |  |  |  |
| 3 |  |  |  |  |  |  |


| 2 |  |  |  |  | TZD |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 |  |  |  |  |  |  |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 |

At the bake sale, the brownie costs 50 cents, the chocolate chip cookie costs 25 cents, the oatmeal raisin cookie is 20 cents, and the cupcakes cost 35 cents. Sam has $\$ 1$. What can he buy?

Mr. Sal needs to set up 137 chairs in the lunch room for the meeting tonight. He has set up 89 chairs so far. How many more chairs does he need to set up in order to be done? Write a number sentence that will show how to solve this problem. You may draw pictures before writing the number sentence.

Anwar has 9 tickets to use at the Halloween Fair. He wants to go on as many different rides as possible.

List the rides he can choose so that he has no tickets left over.
Write a number sentence that explains your answer choices.

## Halloween Rides

## Number of Tickets

Hayride 2
Haunted House 4
Mirror House 3
Ghoul Maze 2
Goblin Ride 2
Look at the number 857. How many hundreds are there? How many tens? How many ones? Can you write another number that is like this one?

Shade 9 of the squares in the figure below.
How many squares are left?

Write your answer as a fraction.

|  |  |  |  |  |
| :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |
|  |  |  |  |  |

# Appendix G - Guide for Viewing Clinical Interviews 

Study Name: Using the Clinical Interview as a Complementary Assessment for Minority Elementary Students to Determine Their In-Depth Understanding of Mathematical Concepts.

Principal Investigator: Nicola Crisp

## Guide for Reviewing Clinical Interviews

As you review the clinical interview for your student, please consider the following factors:

- The levels of mathematical knowledge displayed by the student in each of the conceptual categories.

Please use the provided list of conceptual categories, as well as the clinical interview protocol with the assessed mathematical concepts for each problem highlighted, to help you.

- Methods of computation

Did the student use mental math, paper and pencil, diagram, writing, or a combination of some or all of the above? Throughout the interview was there a preferred method(s) of computation used by the student?

- Methods students chose to explain their problem solving

Oral

Written

- Did body language contradict or support verbal response?
- Were there any surprises?
- Were the student's responses as you anticipated?

Appendix H - Comparison of Scores for Clinical Interview Items \& NJPASS Items
Name: Kevin

| CI Item \# | CI Score | NJPASS Item \# | NJPASS Score |
| :---: | :---: | :---: | :---: |
| 1a | 1 | 2 | 1 |
| 1b | 1 | 2 |  |
| 2a | 1 | 4 | 1 |
| 2b | 1 | 4 |  |
| 3 | 1 | 19, 27 | 4 |
| 4a | 1 | 8 | 3 |
| 4b | 1 | 8 |  |
| 5a | 1 | 3, 28 | 4 |
| 5b | 1 | 3, 28 |  |
| 6 a | 1 | 6, 21 | 2 |
| 6b | 1 | 6, 21 |  |
| 7a | 1 | 10, 11 | 1 |
| 7b | 1 | 10, 11 |  |
| 7c | 1 | 10, 11 |  |
| 7d | 1 | 10, 11 |  |
| 8 a | 1 | 14 | 3 |
| 8b | 1 | 14 |  |
| 9a | 0 | 17 | 1 |
| 9b | 0 | 17 |  |
| 9c | 0 | 17 |  |
| 9d | 0 | 17 |  |
| 9 e | 0 | 17 |  |
| 9f | 0 | 17 |  |
| 10 | 1 | 18, 24 | 1 |
| 11 | 0 | 23 | 1 |
| 12a | 0 | 26 | 0 |
| 12b | 0 | 26 |  |
| 13a | 1 | 30 | 1 |
| 13b | 1 | 30 |  |
| 14a | 1 | 1,13 | 1 |
| 14b | 1 | 1,13 |  |
|  |  |  |  |
| Total Correct | 22/31 |  | 24/30 |
| Percent Correct | 71\% |  | 80\% |

Note: NJPASS items 3, 8, 14, 19 and 26 are open-ended.

Comparison of Scores for Clinical Interview Items \& NJPASS Items
Name: Sam

| CI Item \# | CI Score | NJPASS Item \# | NJPASS Score |
| :---: | :---: | :---: | :---: |
| 1a | 1 | 2 | 0 |
| 1b | 1 | 2 |  |
| 2a | 1 | 4 | 1 |
| 2b | 1 | 4 |  |
| 3 | 1 | 19, 27 | 4 |
| 4a | 1 | 8 | 3 |
| 4b | 1 | 8 |  |
| 5a | 1 | 3, 28 | 3 |
| 5b | 0 | 3, 28 |  |
| 6a | 1 | 6,21 | 2 |
| 6b | 1 | 6,21 |  |
| 7a | 0 | 10, 11 | 2 |
| 7b | 1 | 10, 11 |  |
| 7c | 1 | 10, 11 |  |
| 7d | 1 | 10, 11 |  |
| 8a | 1 | 14 | 2 |
| 8b | 1 | 14 |  |
| 9a | 1 | 17 | 1 |
| 9b | 0 | 17 |  |
| 9c | 1 | 17 |  |
| 9d | 1 | 17 |  |
| 9 e | 0 | 17 |  |
| 9f | 1 | 17 |  |
| 10 | 1 | 18, 24 | 2 |
| 11 | 0 | 23 | 1 |
| 12a | 0 | 26 | 1 |
| 12b | 0 | 26 |  |
| 13a | 1 | 30 | 1 |
| 13b | 1 | 30 |  |
| 14a | 1 | 1,13 | 1 |
| 14b | 0 | 1,13 |  |
|  |  |  |  |
| Total Correct | 23/31 |  | 24/30 |
| Percent Correct | 74\% |  | 80\% |

Note: NJPASS items 3, 8, 14, 19 and 26 are open-ended.

Comparison of Scores for Clinical Interview Items \& NJPASS Items
Name: Oliver

| CI Item \# | CI Score | NJPASS Item \# | NJPASS Score |
| :---: | :---: | :---: | :---: |
| 1a | 1 | 2 | 1 |
| 1b | 1 | 2 |  |
| 2a | 1 | 4 | 1 |
| 2b | 1 | 4 |  |
| 3 | 1 | 19, 27 | 4 |
| 4a | 1 | 8 | 3 |
| 4b | 1 | 8 |  |
| 5a | 1 | 3, 28 | 3 |
| 5b | 1 | 3, 28 |  |
| 6 a | 1 | 6, 21 | 2 |
| 6b | 0 | 6, 21 |  |
| 7a | 1 | 10, 11 | 2 |
| 7b | 1 | 10, 11 |  |
| 7c | 1 | 10, 11 |  |
| 7d | 1 | 10, 11 |  |
| 8a | 0 | 14 | 3 |
| 8b | 0 | 14 |  |
| 9a | 0 | 17 | 1 |
| 9b | 1 | 17 |  |
| 9c | 0 | 17 |  |
| 9d | 0 | 17 |  |
| 9 e | 0 | 17 |  |
| 9f | 0 | 17 |  |
| 10 | 1 | 18, 24 | 2 |
| 11 | 0 | 23 | 1 |
| 12a | 0 | 26 | 3 |
| 12b | 0 | 26 |  |
| 13a | 1 | 30 | 1 |
| 13b | 1 | 30 |  |
| 14a | 1 | 1,13 | 2 |
| 14b | 1 | 1,13 |  |
|  |  |  |  |
| Total Correct | 20/31 |  | 29/30 |
| Percent Correct | 65\% |  | 97\% |

Note: NJPASS items 3, 8, 14, 19 and 26 are open-ended.

Comparison of Scores for Clinical Interview Items \& NJPASS Items
Name: David

| CI Item \# | CI Score | NJPASS Item \# | NJPASS Score |
| :---: | :---: | :---: | :---: |
| 1a | 1 | 2 | 1 |
| 1b | 1 | 2 |  |
| 2a | 1 | 4 | 1 |
| 2b | 1 | 4 |  |
| 3 | 0 | 19, 27 | 1 |
| 4a | 1 | 8 | 3 |
| 4b | 0 | 8 |  |
| 5a | 0 | 3, 28 | 3 |
| 5b | 0 | 3, 28 |  |
| 6a | 1 | 6, 21 | 2 |
| 6b | 1 | 6,21 |  |
| 7a | 1 | 10, 11 | 0 |
| 7b | 0 | 10, 11 |  |
| 7c | 1 | 10, 11 |  |
| 7d | 0 | 10, 11 |  |
| 8 a | 1 | 14 | 3 |
| 8b | 1 | 14 |  |
| 9a | 0 | 17 | 1 |
| 9b | 0 | 17 |  |
| 9c | 0 | 17 |  |
| 9d | 0 | 17 |  |
| 9 e | 0 | 17 |  |
| 9f | 0 | 17 |  |
| 10 | 1 | 18, 24 | 1 |
| 11 | 0 | 23 | 1 |
| 12a | 0 | 26 | 0 |
| 12b | 0 | 26 |  |
| 13a | 1 | 30 | 1 |
| 13b | 1 | 30 |  |
| 14a | 1 | 1,13 | 2 |
| 14b | 0 | 1,13 |  |
|  |  |  |  |
| Total Correct | 15/31 |  | 20/30 |
| Percent Correct | 48\% |  | 67\% |

Note: NJPASS items 3, 8, 14, 19 and 26 are open-ended.

Comparison of Scores for Clinical Interview Items \& NJPASS Items
Name: Jake

| CI Item \# | CI Score | NJPASS Item \# | NJPASS Score |
| :---: | :---: | :---: | :---: |
| 1a | 1 | 2 | 1 |
| 1b | 1 | 2 |  |
| 2a | 0 | 4 | 1 |
| 2b | 0 | 4 |  |
| 3 | 0 | 19, 27 | 1 |
| 4a | 1 | 8 | 2 |
| 4b | 1 | 8 |  |
| 5a | 1 | 3, 28 | 0 |
| 5b | 0 | 3, 28 |  |
| 6 a | 1 | 6, 21 | 1 |
| 6b | 0 | 6, 21 |  |
| 7a | 0 | 10, 11 | 1 |
| 7b | 0 | 10, 11 |  |
| 7c | 1 | 10, 11 |  |
| 7d | 0 | 10, 11 |  |
| 8a | 1 | 14 | 2 |
| 8b | 1 | 14 |  |
| 9a | 0 | 17 | 1 |
| 9b | 0 | 17 |  |
| 9c | 0 | 17 |  |
| 9d | 0 | 17 |  |
| 9 e | 0 | 17 |  |
| 9f | 0 | 17 |  |
| 10 | 1 | 18, 24 | 1 |
| 11 | 0 | 23 | 0 |
| 12a | 0 | 26 | 1 |
| 12b | 0 | 26 |  |
| 13a | 1 | 30 | 0 |
| 13b | 1 | 30 |  |
| 14a | 1 | 1,13 | 2 |
| 14b | 0 | 1,13 |  |
|  |  |  |  |
| Total Correct | 13/31 |  | 14/30 |
| Percent Correct | 42\% |  | 47\% |

Note: NJPASS items 3, 8, 14, 19 and 26 are open-ended.

Comparison of Scores for Clinical Interview Items \& NJPASS Items
Name: Martin

| CI Item \# | CI Score | NJPASS Item \# | NJPASS Score |
| :---: | :---: | :---: | :---: |
| 1a | 1 | 2 | 0 |
| 1b | 1 | 2 |  |
| 2a | 1 | 4 | 1 |
| 2b | 1 | 4 |  |
| 3 | 0 | 19, 27 | 3 |
| 4a | 1 | 8 | 3 |
| 4b | 1 | 8 |  |
| 5a | 1 | 3, 28 | 2 |
| 5b | 1 | 3, 28 |  |
| 6a | 1 | 6, 21 | 2 |
| 6b | 1 | 6,21 |  |
| 7a | 0 | 10, 11 | 0 |
| 7b | 1 | 10, 11 |  |
| 7c | 1 | 10, 11 |  |
| 7d | 1 | 10, 11 |  |
| 8 a | 1 | 14 | 3 |
| 8b | 1 | 14 |  |
| 9a | 0 | 17 | 0 |
| 9b | 0 | 17 |  |
| 9c | 0 | 17 |  |
| 9d | 0 | 17 |  |
| 9 e | 0 | 17 |  |
| 9f | 0 | 17 |  |
| 10 | 0 | 18, 24 | 1 |
| 11 | 0 | 23 | 0 |
| 12a | 0 | 26 | 1 |
| 12b | 0 | 26 |  |
| 13a | 1 | 30 | 0 |
| 13b | 1 | 30 |  |
| 14a | 1 | 1,13 | 2 |
| 14b | 0 | 1,13 |  |
|  |  |  |  |
| Total Correct | 18/31 |  | 18/30 |
| Percent Correct | 58\% |  | 60\% |

Note: NJPASS items 3, 8, 14, 19 and 26 are open-ended.

Comparison of Scores for Clinical Interview Items \& NJPASS Items
Name: Philip

| CI Item \# | CI Score | NJPASS Item \# | NJPASS Score |
| :---: | :---: | :---: | :---: |
| 1a | 1 | 2 | 1 |
| 1b | 1 | 2 |  |
| 2a | 1 | 4 | 1 |
| 2b | 1 | 4 |  |
| 3 | 0 | 19, 27 | 2 |
| 4a | 1 | 8 | 3 |
| 4b | 1 | 8 |  |
| 5a | 1 | 3, 28 | 2 |
| 5b | 1 | 3, 28 |  |
| 6a | 1 | 6,21 | 2 |
| 6b | 1 | 6,21 |  |
| 7 a | 1 | 10, 11 | 0 |
| 7b | 1 | 10, 11 |  |
| 7c | 1 | 10, 11 |  |
| 7d | 1 | 10, 11 |  |
| 8a | 1 | 14 | 3 |
| 8b | 1 | 14 |  |
| 9a | 1 | 17 | 0 |
| 9b | 1 | 17 |  |
| 9c | 1 | 17 |  |
| 9d | 1 | 17 |  |
| 9 e | 1 | 17 |  |
| 9f | 1 | 17 |  |
| 10 | 0 | 18, 24 | 1 |
| 11 | 0 | 23 | 0 |
| 12a | 0 | 26 | 2 |
| 12b | 0 | 26 |  |
| 13a | 1 | 30 | 1 |
| 13b | 1 | 30 |  |
| 14a | 0 | 1,13 | 1 |
| 14b | 0 | 1,13 |  |
|  |  |  |  |
| Total Correct | 24/31 |  | 19/30 |
| Percent Correct | 77\% |  | 63\% |

Note: NJPASS items 3, 8, 14, 19 and 26 are open-ended.

## Kevin

| Categories/Skills | \# of Points Possible | \# of Points Earned | \% of Points Earned |
| :--- | :---: | :---: | :---: |
| Number Sense, <br> Operations, and | 14 | 12 | 86 |
| Properties | 6 | 5 |  |
| Measurement <br> Spatial Sense and <br> Geometry | 6 | 6 | 100 |
| Data Analysis, <br> Probability, and <br> Discrete <br> Mathematics | 7 | 3 | 43 |
| Patterns and <br> Algebra <br> TOTAL | 7 | 7 | 100 |
| Kevin earned a level of Proficient on the NJPASS based on the scores outlined above. |  |  |  |

## Sam

| Categories/Skills | \# of Points Possible | \# of Points Earned | \% of Points Earned |
| :--- | :---: | :---: | :---: |
| Number Sense, | 14 | 10 | 71 |
| Operations, and <br> Properties | 6 |  |  |
| Measurement | 6 | 5 | 83 |
| Spatial Sense and <br> Geometry | 7 | 5 | 67 |
| Data Analysis, <br> Probability, and | 7 |  | 71 |
| Discrete <br> Mathematics | $\mathbf{4}$ |  |  |
| Patterns and | $\mathbf{5 0}$ | $\mathbf{2 9}$ | 71 |
| Algebra |  |  |  |
| TOTAL |  |  |  |

Sam earned a level of Proficient on the NJPASS based on the scores outlined above.

Appendix I - NJPASS Individual Performance Profiles - Continued

## Oliver

| Categories/Skills | \# of Points Possible | \# of Points Earned <br> Number Sense, | 14 |
| :--- | :---: | :---: | :---: | of Points Earned

Oliver earned a level of Advanced on the NJPASS based on the scores outlined above.

## David

| Categories/Skills | \# of Points <br> Possible | \# of Points Earned | \% of Points Earned |
| :--- | :---: | :---: | :---: |
| Number Sense, | 14 | 11 | 79 |
| Operations, and <br> Properties | 6 | 4 | 67 |
| Measurement <br> Spatial Sense and <br> Geometry | 6 | 6 | 100 |
| Data Analysis, <br> Probability, and <br> Discrete <br> Mathematics | 7 | 2 | 29 |
| Patterns and | 7 | 5 | 71 |
| Algebra |  |  |  |
| TOTAL | $\mathbf{7 0}$ | $\mathbf{2 8}$ | $\mathbf{7 0}$ |

David achieved a level of Proficient on the NJPASS based on the scores outlined above.

## Martin

| Categories/Skills | \# of Points <br> Possible | \# of Points Earned | \% of Points Earned |
| :--- | :---: | :---: | :---: |
| Number Sense, <br> Operations, and <br> Properties | 14 | 7 | 50 |
| Measurement | 6 |  |  |
| Spatial Sense and <br> Geometry | 6 | 4 | 67 |
| Data Analysis, <br> Probability, and | 7 | 3 | 67 |
| Discrete <br> Mathematics | 7 | 4 | 43 |
| Patterns and <br> Algebra <br> TOTAL | $\mathbf{4 0}$ | $\mathbf{2 2}$ | 57 |

Martin achieved a level of Basic on the NJPASS based on the scores outlined above.

## Jake

| Categories/Skills | \# of Points <br> Possible | \# of Points Earned | \% of Points Earned |
| :--- | :---: | :---: | :---: |
| Number Sense, | 14 | 4 | 29 |
| Operations, and <br> Properties | 6 | 2 | 33 |
| Measurement <br> Spatial Sense and <br> Geometry | 6 | 4 | 67 |
| Data Analysis, <br> Probability, and <br> Discrete <br> Mathematics | 7 | 3 | 43 |
| Patterns and <br> Algebra <br> TOTAL | 7 | 5 | $\mathbf{7 1}$ |

Jake achieved a level of Basic based on the NJPASS based on the scores outlined above.

Appendix I - NJPASS Individual Performance Profiles - Continued

## Philip

| Categories/Skills | \# of Points <br> Possible | \# of Points Earned | \% of Points Earned |
| :--- | :---: | :---: | :---: |
| Number Sense, | 14 | 5 | 36 |
| Operations, and <br> Properties | 6 | 4 | 67 |
| Measurement <br> Spatial Sense and <br> Geometry | 6 | 5 | 83 |
| Data Analysis, <br> Probability, and <br> Discrete <br> Mathematics | 7 | 4 | 57 |
| Patterns and <br> Algebra <br> TOTAL | 7 | 6 | 86 |
| Prina | $\mathbf{4 0}$ | $\mathbf{2 4}$ | $\mathbf{6 0}$ |

Philip achieved a level of Proficient on the NJPASS based on the scores outlined above.

## Appendix J

## Consent Form for Adults

## CONSENT FORM FOR ADULTS

Please read below with care. You can ask questions at any time, now or later. You can talk to other people before you fill in this form.

Study's Title:Using the clinical interview as a complementary assessment for male African American elementary students in order to determine their in-depth understanding of mathematics concepts.

Why is this study being done?

This study is being done to see if male African American students can show what they really know about mathematics in a different kind of assessment - the clinical interviewthan they are able to show on a standardized test. The study will also look at what teachers can learn from this assessment and how this might affect their views of these students.

What will happen while you are in the study?

I will interview you twice - once before the clinical interviews and once after the interviews. In the first clinical interview I will ask you to tell me how you view your students' math abilities, how they solve math problems, and also to explain how you came to these views.

In between the interviews we will watch the tapes of the clinical interviews together so that you can think about what they showed about your students' math abilities before the second interview.

In the second interview we will talk about what you learned about your students' math performance and whether this knowledge changed your views about your students' math abilities.

The teacher interviews will be taped in order to help me be accurate when I am working with the data. Only I will be able to use the tapes, and I will store all the tapes securely in order to protect your privacy.

Time:This study will take about four hours spread out over a two week period.

The first interview should last between 45 minutes to an hour.
The time for viewing the clinical interviews should be about $30-40$ minutes per interview, and the total time should be no more than two hours as no teacher should have more than 3 students who are being interviewed.

The second and final interview should last between 45 minutes to an hour.

Risks: You may be surprised if your views about your students' mathematical abilities are different from what the clinical interviews show.

However, the risks of participating in this study are no greater than those in ordinary life.

Benefits: You may benefit from this study by gaining information about your students' mathematical abilities that you did not know before.

The students in your class who took part in the study may benefit from this study as they too will have a better idea of what they know and don't know and may also feel better about their mathematical ability.

There is continued and increasing concern about the weaker performance of African American students, especially boys, on standardized tests in comparison with their White and Asian peers. If using another kind of assessment allows these students to show that they know more about mathematical concepts than they are able to show on standardized tests, this would be a benefit to our society as this other assessment could also be used to help us get a more accurate picture of these students' mathematical knowledge.

Who will know that you are in this study? You will not be linked to any presentations. We will keep who you are confidential according to the law. Therefore, in writing up the results of the data in the study, the real names of the participants, including yours, will not be used. The district superintendents and your principals will also know that you participated as I had to get their permission to do the study.

## Do you have to be in the study?

You do not have to be in this study. You are a volunteer! It is okay if you want to stop at any time and not be in the study. You do not have to answer any questions you do not want to answer. Nothing will happen to you.

Do you have any questions about this study? can be contacted at the following address - Nicola Crisp, P.O. Box 11504, Chattanooga, TN 37401 before and after the study. The phone number there is 423-228-5240. During the study I can be reached at 1061 Ironbound Avenue, Plainfield, NJ 07060. The phone number there is 908-754-1817. My cell phone number is 908-612-0302. My email address is ncrisp@comcast.net.

Do you have any questions about your rights? Phone or email the IRB Chair, Debra Zellner (reviewboard@mail.montclair.edu or 973-655-4327).

It is okay to use my data in other studies:
Please initial: $\qquad$ Yes $\qquad$ No

I would like to get a summary of this study:
Please initial: $\qquad$ Yes $\qquad$ No

When the investigator is videotaping individual subjects, add the following two statements:

It is okay to videotape me while I am in this study:
Please initial: ___ Yes No

It is not okay to videotape me while I am in this study:
Please initial: ___ Yes No

The copy of this consent form is for you to keep.

If you choose to be in this study, please fill in your lines below.
$\qquad$
$\qquad$
Print your name here
Sign your name here
Date
$\qquad$
$\qquad$

Name of Principal Investigator
Signature
Date


## Appendix K

Parent/Guardian Consent Form for Participants Under 18 Years of Age or Dependent Adults

## PARENT/GUARDIAN CONSENT FORM

Please read below with care. You can ask questions at any time, now or later. You can talk to other people before you fill in this form.

Study's Title:Using the clinical interview as a complementary assessment for male African American elementary students in order to determine their in-depth understanding of mathematics concepts.

Why is this study being done? This study is to see how much math these students can show that they know when they get to work on a different kind of test than the standardized test.

## What will happen while your child or dependent is in the study?

Your child will be interviewed by himself in a room that has been set aside for this purpose. We will walk with him to and from his classroom. In the interview, he will be asked to solve some math problems that are like those on the NJPASS. They will be
asked one at a time. He will be able to ask as many questions as he needs, and the interviewer will make sure he understands what is being asked. He will also be able to use materials like cubes and spinners to help him. The interview will be videotaped so that you can see exactly what happened.

Time:This study will take about 20-30 minutes of your child's time.

Risks:Your child might feel nervous but the risks to your child or dependent are no greater than those in ordinary life.

Benefits: Your child or dependent may benefit from this study because he will be able to show what he knows about the math problems. His teacher may benefit by getting more information about how much math he knows.

The societal benefits are that using more than one kind of test could help these students show how much math they know, and this would be useful to the students, the teachers, and the schools.

Who will know that your child or dependent is in this study? Your child or dependent will not be linked to any presentations. We will keep who your child or dependent is confidential by reporting the findings of the study using pseudonyms so that your child's identity cannot be identified to any outside person or group. The superintendent,
assistant superintendent, principal, and child's teacher will know that they are in the study.

## Does your child or dependent have to be in the study?

Your child or dependent does not have to be in this study. He is a volunteer! It is fine if he wants to stop at any time and not be in the study. He does not have to answer any questions that he does not want to answer.

Do you have any questions about this study? I can be contacted at P.O. Box 11504, Chattanooga, TN 37401 before and after the study. The home phone number there is 423-228-5240. During the study I will be at 1061 Ironbound Avenue, Plainfield, NJ 07060. The phone number there is $908-754-1817$. My cell phone number is $908-612$ 0302. My email address is ncrisp@comcast.net.

Do you have any questions about your rights? Phone or email the IRB Chair, Debra Zellner (reviewboard@mail.montclair.edu or 973-655-4327).

It is okay to use his data in other studies:
Please initial: ___ Yes No
I would like to get a summary of this study:
Please initial: $\qquad$ Yes $\qquad$ No
When the investigator is videotaping individual subjects, add the following two statements:
It is okay to videotape him while in this study:
Please initial:
Yes $\qquad$ No
It is not okay to videotape him while in this study:
Please initial: $\qquad$ Yes $\qquad$ No

The copy of this consent form is for you to keep.

If you choose to have your child or dependent in this study, please fill in the lines below.

| Name of Parent/Guardian    <br>     <br>  Signature  Date <br> Name of Parent/Guardian Signature   | Date |
| :--- | :---: | :---: |


| If you choose to be in this study, please fill in your lines below. |  |  |
| :---: | :---: | :---: |
| Print your name here | Sign your name here | Date |
| Name of Principal Investigator | Signature | Date |
| (if applicable) Name of Faculty Sponsor | Signature |  |

## Appendix L - Assent Form

## ASSENT FORM

Please read below with care. You can ask questions at any time, now or later. You can talk to other people before you fill in this form.

Who am I? I am Nicola Crisp. I used to teach $4^{\text {th }}$ in Scotch Plains at Evergreen School for 23 years before I retired last year.

Why is this study being done? This study is to see how children do on different kinds of math tests.

What will happen while you are in the study? You will be asked to do some math problems just like the ones you have been doing all year. It will be videotaped. I will let you, your parents, and teachers know how you did.

Time:This study will take about 20-30 minutes.

Risks:You may feel a little nervous until you see what you are asked to do. You may get a little worried if you forget how to do something, or if you are not sure of how to solve problem.

Benefits:This study will help you because you will be able to see how well you can do the problems and show how much math you know.

It will help your teacher because she will be able to see how much math you know.
Students, teachers, and schools will be helped by finding out that students can be helped by having more than one kind of math test to show what they know.

Who will know that you might be in this study? You, your parents, your teacher, the principal, the lady who will give you the problems, and I will know that you are here, but we won't tell anyone else.

## Do you have to be in the study?

You do not have to be in this study. We won't get angry with you if you say no. It is okay if you change your mind at any time and leave the study. Nothing bad will happen to you. This study will not have anything to do with your report card.

Do you have any questions about this study? I can be reached at 1061 Ironbound Avenue, Plainfield, NJ 07060 while the study is being conducted. My cell phone number is 908-612-0302. My phone number in Plainfield is 908-754-1817. My email address is ncrisp@comcast.net.

Do you have any questions about your rights? Phone or email the IRB Chair, Debra Zellner (reviewboard@mail.montclair.edu or 973-655-4327).

It is okay to use my data in other studies:
Please initial:
Yes $\qquad$ No

I would like to get a summary of this study:
Please initial: $\qquad$ Yes $\qquad$ No

When the investigator is videotaping individual subjects, add the following two statements:

It is okay to videotape me while I am in this study:
Please initial: $\qquad$ Yes $\qquad$ No

It is not okay to videotape me while I am in this study:
Please initial: $\qquad$ Yes $\qquad$ No

|  |  |  |
| :--- | :--- | :--- | :--- |
| Name of Research Participant | Signature | Date |
|  |  |  |


| Name of Witness | Signature | Date |
| :---: | :---: | :---: |
| Name of Principal Investigator | Signature | Date |
| (if applicable) Name of Faculty Sponsor | Signature | Date |


[^0]:    This Dissertation is brought to you for free and open access by Montclair State University Digital Commons. It has been accepted for inclusion in Theses, Dissertations and Culminating Projects by an authorized administrator of Montclair State University Digital Commons. For more information, please contact digitalcommons@montclair.edu.

