ABSTRACT

Title of Dissertation: PYGMALION OR PLEKHANOV IN THE CLASSROOM: THE SUBTLE ROLE OF SOCIAL CLASS IN TEACHER PERCEPTIONS

Laurent Ross, Doctor of Education, 2015

Dissertation directed by: Dr. Willis Hawley, Teaching, Learning, Policy and Leadership

The purpose of this study was to examine the role of teacher characteristics and school demographics in teachers' perceptions of children's cognitive abilities. Most researchers find that teachers' personal characteristics are not related to their perceptions of their children's cognitive abilities. In a 2011 study, Douglas Ready and David Wright find that socioeconomic characteristics of the classroom and the school have a stronger relationship to teacher biases than the personal characteristics of the teachers themselves.

This study used the National Education Association's KEYS database to examine the relationship between teacher perception and student achievement. This study compared data on teacher perceptions of the abilities of their classrooms with school-wide standardized testing results using ANOVA as well as univariate analysis to examine five research questions. These five questions focused on the accuracy of teachers' perceptions

of their students' abilities and how that accuracy varied across the race of the student, the grade level of the student, the characteristics of the teachers, and the socio-demographic characteristics of the school

Teachers, on average, are quite accurate in their perceptions of their students' cognitive abilities. However, that average accuracy hides wide disparities among different groups of teachers. The socioeconomic (SES) status of the parents of the students served by the school and the type of community it is located in had a strong relationship with the accuracy of teacher perceptions. Schools with lower-income students and schools located in more urban communities tended to overestimate student ability while schools with upper-middle income students and those located in small towns tended to underestimate student ability on average.

Teachers systematically perceived minority students to have greater cognitive ability than their standardized test score results would suggest while systematically underrating white students relative to standardized test score results.

Teacher personal characteristics were not significant predictors of teacher accuracy. The SES of the parents of the children served by the school and the location of the school's neighborhood were significant predictors. This suggested that it is environmental characteristics of the school, rather than individual teacher characteristics that had the most influence on teacher perception of student cognitive ability.

PYGMALION OR PLEKHANOV IN THE CLASSROOM: THE SUBTLE ROLE OF SOCIAL CLASS IN TEACHER PERCEPTIONS

by

Laurent Ross

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Education
2015

Advisory Committee:
Professor Willis Hawley, Chair
Dr. Edward Crowe
Dr. Jacques Nacson
Dr. John O'Flahavan
Dr. William Strein

Dedication

I dedicate this dissertation to my wonderful wife, Mercedes, for putting up with me, not only while I was writing this thesis, but throughout my doctoral studies and thirty-two years of marriage.

Acknowledgements

I would like to give a standing round of applause to my dissertation committee for their assistance and guidance in completing this dissertation. And, an additional round of applause and thanks is due for their patience as I labored mightily with the KEYS data set to find my results.

I am especially thankful to Dr. Willis Hawley, my committee Chair, for originally suggesting the topic of this dissertation and for keeping a steady hand on the tiller as I went through what seemed like thousands of iterations of change in my methodology and approach. Dr. Edward Crowe and Dr. John O'Flahavan were very helpful as a sounding board for all the methodological techniques I envisioned throwing at the data. Dr. Jacques Nacson provided me with direct access to the KEYS data and contractor resources to deal with all the data issues I confronted in writing this thesis. Dr. Sylvia Rosenfeld generously gave of her time and expertise in reviewing my original proposal and gave me numerous helpful pointers in refining both the proposal and the final version of this dissertation. Dr. William Strein graciously agreed to join the committee as a voting member on sudden notice and made a number of helpful comments for the final editing of the dissertation.

Table of Contents

Dedication ii

Acknowledgements iii

Table of Contents iv

List of Tables vi

Chapter I - Introduction 1

Background of the Study……..... 2

Statement of the Problem……...5

Purpose of the Study…….. 6

Significance of the Study…….....7

Definition of Terms…….....8

Theoretical Framework……12

Research Questions…….....15

Limitations……17

Assumptions……..18

Chapter II – Review of the Literature 19

Introduction……19

Experimental Studies of Teacher Perception…….19

Conditional Race Neutrality and Teacher Perceptions……..20

Ready and Wright on Teacher Perception…….25

Summary……...29

Chapter III - Methodology 31

Research Question 1……..35

Research Question 2……..41

Research Question 3……..45

Research Question 4……..46

Research Question 5……..46

Chapter IV – Presentation and Analysis of the Data 48

Introduction……48

Research Question 1……...53

Research Question 2……...64

Research Question 3……...86

Research Question 4……...92

Research Question 5……...98

Summary……..108

Chapter V - Summary, Implications, and Recommendations..... 112

Introduction……..112

Summary……..112

Accuracy of Teachers' Perceptions of Their Students' Performance…...……113

Accuracy of Teachers' Perceptions of Minority Students…….……..….114

Teacher Accuracy by Grade…..……...115

Relationship of Personal Characteristics of Teachers to Teacher Accuracy..116
Relationship of the Environmental Characteristics of the Schools Teachers Work in to Teacher Accuracy……...117
Discussion……..119
Bibliography 128

List of Tables

Table 1: Demographic Information for All Teachers and Schools They Teach In: Race, Education Level, Certification, Years' Experience, Socioeconomic Status of School, Community of School, and Size of School by Gender of the Teacher 34
Table 2: ANOVA Assumptions for Perceptual Match: Summary Table 39
Table 3: ANOVA Assumptions for Perceptual Match Minority: Summary Table 42
Table 4: ANOVA Assumptions for Perceptual Match White: Summary Table 44
Table 5: Comparison of Distribution of Students between Elementary and Secondary Schools in the KEYS Data Set and the Digest of Education Statistics 50
Table 6: Comparison of Racial/Ethnic Background of Students in the KEYS Data Set and the Digest of Education Statistics 51
Table 7: Comparison of Distribution of the Percentage of Students Who Were Eligible for Free/Reduced Price Lunch, English as a Second Language Services, and Special Education Services in the KEYS Data Set and the Digest of Education Statistics 51
Table 8: Comparison of Highest Degree Attained by Teachers Surveyed in the KEYS Data Set and National Statistics from the Digest of Education Statistics 52
Table 9: Comparison of Years' Experience of Teachers in the KEYS Data Set and National Statistics from the Digest of Education Statistics 52
Table 10: Comparison of Racial/Ethnic Background of High School Teachers in the KEYS Data Set and the Digest of Education Statistics 53
Table 11: Comparison of Gender Makeup of High School Teachers in the KEYS Data Set and the Digest of Education Statistics 53
Table 12: Means and ANOVA Summary Table for Perceptual Match by Gender of the Teacher 54
Table 13: Means and Welch's F Test Summary Table for Perceptual Match by Racial/Ethnic Background of the Teacher 55
Table 14: Means and Welch's F Test Summary Table for Perceptual Match by Race of the Teacher (Non-White/White 56
Table 15: Means and Welch's F Test Summary Table for Perceptual Match by Education Level of the Teacher 57
Table 16: Means and One-Way ANOVA Summary Table for Perceptual Match by Level of Teacher Certification 57

Table 17: Means and One-Way ANOVA Summary Table for Perceptual Match by Total Years' Experience of the Teacher 58
Table 18: Means and One-Way ANOVA Summary Table for Perceptual Match by Years of Experience in the System of the Teacher 59
Table 19: Means and One-Way ANOVA Summary Table for Perceptual Match by Years of Experience in the Building of the Teacher 59
Table 20: Means and One-Way ANOVA Summary Table for Perceptual Match by Total Years' Experience of the Teacher (< 2 Years', > 2 Years') 60
Table 21: Means and One-Way ANOVA Summary Table for Perceptual Match by Years of Experience in the System of the Teacher (< 2 Years',> 2 Years') 60
Table 22: Means and One-Way ANOVA Summary Table for Perceptual Match by Years of Experience in the Building of the Teacher (< 2 Years', > 2 Years') 61
Table 23: Means and Welch's F Test Summary Table for Perceptual Match by Socioeconomic Level of the Parents of the Students by the School Serves 62
Table 24: Means and One-Way ANOVA Summary Table for Perceptual Match by Type of Community the School Is Located In 63
Table 25: Means and One-Way ANOVA Summary Table for Perceptual Match by Size of the School 64
Table 26: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Gender of the Teacher 65
Table 27: Means and One-Way ANOVA Summary Table for Perceptual Match White by Gender of the Teacher 66
Table 28: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Racial/Ethnic Background of the Teacher 67
Table 29: Means and Welch's F Test Summary Table for Perceptual Match White by Racial/Ethnic Background of the Teacher 68
Table 30: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Racial/Ethnic Background of the Teacher (Non-White/White 68
Table 31: Means and One-Way ANOVA Summary Table for Perceptual Match White by Racial/Ethnic Background of the Teacher (Non-White/White 69
Table 32: Means and Welch's F Test Summary Table for Perceptual Match Minority by Education Level of the Teacher 70
Table 33: Means and Welch's F Test Summary Table for Perceptual Match White by Education Level of the Teacher 71
Table 34: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Level of Teacher Certification 72

Table 35: Means and One-Way ANOVA Summary Table for Perceptual Match White by Level of Teacher Certification 72
Table 36: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Total Years' Experience of the Teacher 73
Table 37: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Years of Experience in the System of the Teacher 74
Table 38: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Years of Experience in the Building of the Teacher 74
Table 39: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Total Years' Experience of the Teacher (< 2 Years', > 2 Years') 75
Table 40: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Years of Experience in the System of the Teacher (< 2 Years', > 2 Years') 75
Table 41: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Years of Experience in the Building of the Teacher (< 2 Years', > 2 Years') 76
Table 42: Means and One-Way ANOVA Summary Table for Perceptual Match White by Total Years' Experience of the Teacher (< 2 Years', > 2 Years') 76
Table 43: Means and One-Way ANOVA Summary Table for Perceptual Match White by Years of Experience in the System of the Teacher (< 2 Years', > 2 Years') 76
Table 44: Means and One-Way ANOVA Summary Table for Perceptual Match White by Years of Experience in the Building of the Teacher (< 2 Years', > 2 Years') 77
Table 45: Average Perceptual Match Score Minority and Perceptual Match Score White for Teachers by Amount of Experience (< 2 Years', > 2 Years') 78
Table 46: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Socioeconomic Status of the Parents of the Students the School Serves 79
Table 47: Means and Welch's F Test Summary Table for Perceptual Match White by Socioeconomic Status of the Parents of the Students Served by the School 79
Table 48: Average Perceptual Match Score Minority and Perceptual Match Score White for Teachers by Socioeconomic Status of the Parents of the Students the School Serves 81
Table 49: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Type of Community the School Is Located In 81

Table 50: Means and Welch's F Test Summary Table for Perceptual Match Minority by Type of Community the School Is Located In 82
Table 51: Average Perceptual Match Score Minority and Perceptual Match Score White for Teachers by the Type of Neighborhood the School Is Located In 83
Table 52: Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Size of the School 83
Table 53: Means and One-Way ANOVA Summary Table for Perceptual Match White by Size of the School 84
Table 54: Average Perceptual Match Score Minority and Perceptual Match Score White for Teachers by the Size of the School 84
Table 55: Summary Table of Perceptual Match Minority Compared to Perceptual Match White for the Values of the Independent Variables 86
Table 56: Means and Welch's F Test Summary Table for Perceptual Match by Type of School 87
Table 57: Means and Welch's F Test Summary Table for Perceptual Match Minority by Type of School 88
Table 58: Means and One-Way ANOVA Summary Table for Perceptual Match White by Type of School 88
Table 59: Two-Way ANOVA Summary Table for Perceptual Match White by Gender and Type of School 91
Table 60: Summary Table of Statistical Significance and Effect Size for Perceptual Match, Perceptual Match Minority and Perceptual Match White by Teacher Characteristic Variables 93
Table 61: Two-Way ANOVA Summary Table for Perceptual Match by Gender and Racial/Ethnic Background of the Teacher (Non-White/White 95
Table 62: Two-Way ANOVA Summary Table for Perceptual Match Minority by Gender and Racial/Ethnic Background of the Teacher 96
Table 63: Two-Way ANOVA Summary Table for Perceptual Match White by Gender and Racial/Ethnic Background of the Teacher 97
Table 64: Summary Table of Statistical Significance and Effect Size for Perceptual Match, Perceptual Match Minority and Perceptual Match White by Environmental Variables 99
Table 65: Two-Way ANOVA for Perceptual Match by Socioeconomic Status (SES) of the Parents of the Students Served by the School (Middle/Lower SES) and the Type of Community the School Is Located In 100
Table 66: Simple Main Effect for Socioeconomic Status (SES) of the Parents of the Students the School Serves (Middle Income/Lower Income) in Two-Way ANOVA for Perceptual

Match by SES of the Parents of the Students Served the School Serves and the Type of Community the School Is Located In 101
Table 67: Simple Main Effect for Type of Community in Two-Way ANOVA for Perceptual Match by Socioeconomic Status (SES) of the Parents of the Students the School Serves (Middle/Lower SES) and the Type of Community the School Is Located In 102
Table 68: Two-Way ANOVA for Perceptual Match Minority by Socioeconomic Status (SES) of the Parents of the Students the School Serves (Middle/Lower SES) and the Type of Community the School Is Located In 103
Table 69: Factorial ANOVA for Perceptual Match by Gender of the Teacher, Teacher Experience in the School Building (< 2 years', > 2 years'), Socioeconomic Status (SES) of the Parents of the Children Served by the School (Middle/Lower SES) and Type of Community School Is Located In 105
Table 70: Factorial ANOVA for Perceptual Match Minority by Gender of the Teacher, Teacher Experience in the School Building (< 2 years', > 2 years'), Socioeconomic Status (SES) of the Parents of the Children Served by the School (Middle/Lower SES) and Type of Community School Is Located In....107

Chapter 1: Introduction

> *…You see, really and truly, apart from the things anyone can pick up (the dressing and the proper way of speaking, and so on), the difference between a lady and a flower girl is not how she behaves, but how she's treated. I shall always be a flower girl to Professor Higgins, because he always treats me as a flower girl, and always will; but I know I can be a lady to you, because you always treat me as a lady, and always will.*
>
> –George Bernard Shaw, *Pygmalion*.

Most of us are familiar with *Pygmalion*, George Bernard Shaw's romantic comedy, with its sharp satire on the English class system, and early nod to feminism. In the play, phonetics professor Dr. Henry Higgins bets a colleague that he can make a lower-class flower girl pass for a lady at an embassy ball. Higgins believes the secret to transforming the cockney-accented Eliza Doolittle into a lady is to teach her to speak impeccably. Higgins does transform Ms. Doolittle and wins the bet, but ends up losing the lady (at least in the original version).

The title of Shaw's play is derived from the Greek myth of Pygmalion. Pygmalion was a Cypriot sculptor who created a statue of a woman out of ivory. His statue was so lifelike and beautiful that Pygmalion fell in love with it. He prayed to Venus to make the statue come alive and was granted his wish when he planted a kiss on its lips. Shaw's Higgins transforms Doolittle as well, into a high-bred lady, but in Shaw's feminist twist she becomes empowered to marry the person who saw her as a lady all along.

Fewer of us may be familiar with Georgi Plekhanov, the late 19th century Marxist philosopher. But in his famous essay "The Role of the Individual in History," Plekhanov posited that historical conditions ultimately determine the scope of the

individual's activity. He believed that even the greatest personalities could not escape the framework of history. Great leaders, like the great social ideas they create and express, arise from critical periods of history; it is not the leaders themselves who create great historical eras but the epochs themselves that provide the conditions in which the natural talents and genius of the leader can prevail.

So how do Pygmalion and Plekhanov intersect other than through alliteration? Both are, in essence, about self-fulfilling prophecies and both have applications for the classroom. Several researchers argue that teachers can have a profound effect on students by their *mere perception* of a child (as Professor Higgins did of Eliza Doolittle). However, Plekhanov might counter that while, yes, teachers *can* have a profound impact on students' lives by their very perception of a student, the most telling influence on a teacher's perception may be the very societal conditions in which that perception arises. To paraphrase myself, even the best-perceiving teachers cannot escape the framework of society. The teacher's perception of his or her students arises from a societal framework. The teacher does not singlehandedly create his/her own perception.

Background of the Study

U.S. society has a number of socio-demographic biases and those on the receiving end of the resulting stereotyping and discrimination suffer numerous consequences. Generally speaking, African-Americans, Latinos, and the poor suffer disproportionately from bias and discrimination, resulting in disproportionately high unemployment rates, disproportionately diminished housing opportunities, and disproportionately fewer educational opportunities. Stereotyping and discrimination

are not limited to the individual; they often become institutionalized. Governments, schools, and workplaces often unwittingly set policies that make existing discrimination worse, resulting in continued unequal outcomes and further perpetuation of stereotypes.

Teachers' perceptions of their students can have a number of strong influences on the children they teach. These perceptions often end up influencing a variety of educational placement decisions, including grouping by ability, grade retention, exposure to curricula, admission to selective programs or schools, assignment to English as a second language, and assignment to special education. These decisions, in turn, can have profound impacts on a student's future employment and other opportunities. Moreover, if teacher perceptions are biased in any way that impacts negatively on minority and poor children, the current educational inequality that exists in our nation may be made worse.

The issue of teacher perceptions is significant and worthy of study because many believe that it may have an influence on achievement differences between whites, African-Americans, and Latinos, as measured by standardized tests. The vast majority of elementary school teachers are white and female. When public-school enrollments reflect a higher percentage of non-white children compared to the overall population (half of whom are male), it is easy to imagine a teacher interpreting a child's skin color, family status, language, or mannerisms as indicators of academic ability.

The most important question then for researchers in the field of teacher perceptions is whether unequal teacher opinions of students in certain demographic groups stems (at least partially) from teacher perceptions themselves or from real cognitive differences (as measured by standardized tests) among socio-demographic groups. A large group of

experimental laboratory-based studies have found that teachers are biased in favor of white students (see Baron, Tom, & Cooper, 1985, for a meta-analysis of 16 laboratory-based studies). I label these researchers the experimentalists in teacher perception studies. The experimentalists argue that unequal teacher opinions of students in certain demographic groups stems from biases that teachers have against those groups.

There are a number of studies that posit that teachers' perceptions are indeed quite accurate and do not reflect systemic bias in the education industry. These studies argue that teachers' perceptions reflect real cognitive differences among socio-demographic groups and are reflective of the real-world classroom, rather than the theoretically racially harmonious classroom of the laboratory experiment (Lee Jusim and Samuel Meisels have made careers of defending the accuracy of teachers' perceptions; see Jussim, 1989; Jussim, Eccles & Madon, 1996; Jussim & Harber, 2005; Meisels, Bickel, Nicholson, Xue & Atkins-Barnett, 2001; Perry & Meisels, 1996). Nonetheless, a small number of researchers find systematic biases in teachers' perceptions (see Burkham, LoGerfo, Reading & Lee, 2007; Tach & Farkas, 2006). I label these investigators the naturalists in teacher perception research. Their studies focus on data from real teachers and students instead of teacher data on perceptions of imaginary students in laboratory experiments.

Douglas Ready and David Wright (2011) have come out with a groundbreaking study on teachers' perceptions which asserts that substantial biases pervade teachers' perceptions of students from different socio-demographic groups. They argue, however, this bias is accounted for more by characteristics of the classroom than by characteristics of the teachers themselves. In other words, teachers working in low-income classrooms are more likely to underestimate their students' abilities, regardless of that teacher's race,

income, or other background. So, *a la* Plekhanov, it is not so much the personal characteristics of the teacher that leads to biases but the context in which the teacher finds him/herself that can lead to biased perception. I label Ready and Wright as the environmentalists in teacher perception studies. The environmentalists use real data on teachers and students like the naturalists. However, they have found significant bias in teacher perception of certain demographic groups which they attribute to environmental factors.

Statement of the Problem

As important as the issue of teachers' perceptions of their students is, especially the role of socio-demographic factors, there is just not enough conclusive research on the matter. There are three main schools of thought in the realm teacher perception studies. Most of the experimentalists find that teachers are biased against African-American students (see Black & Cooper, 1985 for a meta-analysis of 16 experimental studies).

Most of the naturalists are fairly sure that teachers' perceptions of students are quite accurate (Jussim, 1989; Jussim, Eccles & Madon, 1996; Jussim & Harber, 2005; Meisels, Bickel, Nicholson, Xue & Atkins-Barnett, 2001; Perry & Meisels, 1996). However, a smaller number of the naturalists believe that teachers' consistently underestimate the abilities of lower-income students (Burkham, LoGerfo, Reading & Lee, 2007; Tach & Farkas, 2006). Two comprehensive literature reviews (Ferguson, 2003; Farkas, 2003) conclude that there is not enough evidence either for, or against, the hypothesis of teacher biases.

Ready and Wright's 2011 study, representing the environmentalists, is an important step forward. They use a sophisticated methodology to demonstrate that classroom biases

are affected more by classroom makeup than the characteristics of the teacher. However, while they use a nationally representative sample, their analysis examines only kindergarten students. While this was done by design ("…these children represent a unique analytic opportunity. From an institutional perspective, kindergarteners begin formal school with a relatively 'clean slate' (Ready & Wright, 2001, p. 340)), it fails to look at teachers in other grades. And, because the vast majority of kindergarten teachers are female, Ready and Wright do not examine the role of the teacher's gender in perception. While the Ready and Wright work is groundbreaking, I am interested in testing their findings at other grade levels and examining the role of teacher gender in teacher perceptual bias.

Purpose of the Study

The purpose of this study is to extend the literature on teacher perceptual accuracy by comparing the hypotheses of the experimentalists, the naturalists, and the environmentalists with a national database. The experimentalists argue that teachers are biased against African-American students. The naturalists, for the most part, argue that teachers are actually quite accurate in their perceptions of students' cognitive abilities perception (see Burkham, LoGerfo, Reading & Lee, 2007; Ferguson, 2003; Farkas, 2003; Jussim, 1989; Jussim, Eccles & Madon, 1996; Jussim & Harber, 2005; Meisels, Bickel, Nicholson, Xue & Atkins-Barnett, 2001; Perry & Meisels, 1996; Tach & Farkas, 2006). The environmentalists (Ready & Wright, 2001) claim that teachers are, indeed, biased against lower-income students, but that those biases stem more from the demographics of the students and the neighborhoods their schools are in than from the personal characteristics of the teachers themselves. However, Ready and Wright looked only at

kindergarten teachers. Because the overwhelming majority of kindergarten teachers are female, Ready and Wright did not examine the role of gender in teacher perceptions of their students. This study extends the field of teacher perception research by using a national database to look at teachers in all grades and by examining the role of gender in teacher perceptual accuracy. Finally, I also looked at what school wide and classroom characteristics are (or are not) related to how accurately teachers perceive student achievement.

Significance of the Study

This study will contribute to a deeper understanding of how teachers' perceptions affect society. Almost all researchers agree that teachers will consistently underestimate the performance of low-income, African-American, Latino, and male students relative to upper-income, white, Asian, or female students. However, a large contingent of these researchers (the naturalists) believe that this bias simply reflects the real world, where upper-income, white, Asian, and female students all score relatively higher on standardized achievement tests than lower-income, African-American, Latino, and male students. These researchers posit that teachers are quite accurate in their perceptions of their students. The environmental analysis argues that teachers *are* biased, but the bias is explained more by their classroom's characteristics than by individual teacher characteristics.

This study used the National Education Association's (NEA's) "Keys to the Excellence of Your Schools" (KEYS) database to compare the hypotheses of the experimentalists, naturalists, and environmentalists. This study also tested the environmental hypothesis that teachers in lower-income classes and teachers in urban

school districts systematically underestimate how well their students will perform on standardized tests.

Confirmation of the environmental hypothesis would provide support for policies that attempt to create diverse student bodies, as well as policies that attempt to end the isolation of low-income, African-American, and Latino students. Confirmation of the environmental hypothesis might also engender more support for renewed efforts to ensure that all teaching candidates are exposed to practical courses, workshops, and internships where they can learn how to effectively work with children from a variety of socio-demographic backgrounds.

Definition of Terms

The operational terms I used in this study are as follows:

Bias: I used the Jussim et al. (1996) definition: "systematically evaluating two groups as differing on some criterion more or less than they actually differ" (p. 329). "…[T]eacher perceptions are biased only to the degree that they over-or-underestimate actual between-group differences" (Ready & Wright, 2011, p. 338). If teachers' perceptions are accurate, then they are, by definition, unbiased. If they are inaccurate, but random, then they are inaccurate and unbiased. If they are systematically biased, then teachers' perceptions are inaccurate and biased (Ready & Wright, 2011). Keep in mind that I cannot delve deeply into the psychological background of the teachers in the KEYS database, so bias cannot be defined conventionally here and has no moral implications. Rather, it is the systematically inaccurate perception of student performance by a group of teachers.

Teachers' perceptions: This refers to teachers' subjective evaluation of their students' performance. In the KEYS survey, teachers were asked how the students in their "target" class (defined as the class they spend the most time in, or, if they teach multiple classes of equal length, the first class they teach in the week) performed. They were also asked to evaluate the performance of their minority students and their white students. This analysis examined how teachers perceive all their students and compared how teachers perceive minority students in the school with how they perceive white students.

Teacher perceptual match: The difference between how teachers rated their target class and how the students in the highest grade in the school performed on standardized achievement tests. The methodology section describes this measure in more detail.

Teacher perceptual match minority: The difference between how teachers rated the minority students in their target class and how the minority students in the highest grade in the school performed on standardized achievement tests. The methodology section describes this measure in more detail.

Teacher perceptual match white: The difference between how teachers rated the white students in their target class and how the white students in the highest grade in the school performed on standardized achievement tests. The methodology section describes this measure in more detail.

Student performance: This refers to the average performance of students in an entire school on standardized achievement tests. In the KEYS survey, the survey administrator was asked about the average performance of all students in the highest grade in the school on standardized achievement tests. Administrators were also asked about the average per-

formance of minority students and white students in the highest grade on achievement tests.

Teacher racial/ethnic background: Teachers were able to report one of six different racial or ethnic backgrounds on KEYS:

1. American Indian/Alaska native;
2. Asian/Pacific Islander;
3. Black/African American;
4. Caucasian (not of Hispanic origin);
5. Hispanic/Latino; or,
6. other racial or ethnic background.

Teacher education level: Teachers were able to report one of six education levels on the KEYS survey:

1. high school degree;
2. two-year college diploma, degree, or certificate;
3. bachelor's degree;
4. master's degree;
5. education specialist or professional diploma; and,
6. doctoral degree (Ph.D., Ed.D., etc.).

Teacher certification: Teachers were asked if they have formal training in the subject area they teach in their target class. The possible responses were as follows:

1. certified in the subject I teach;
2. not certified, but have some formal training;
3. no formal training; or,
4. other.

Teacher experience: Teachers were asked how much full-time experience they had as an employee in the education sector. Teachers were asked to describe the length of

experience they had in the teaching profession, in their school system, and in their current building. The responses for all three questions were grouped as follows:

1. fewer than 2 years;
2. 2-5 years;
3. 6-10 years;
4. 11-19 years; or,
5. 20 or more years.

School socioeconomic status (SES): Survey administrators were asked to rate the socioeconomic status of the parents of the children their school serves. The possible responses were:

1. high income;
2. upper-middle income;
3. middle income;
4. lower-middle income; or,
5. lower income.

Type of community school is located in: Survey administrators were asked to determine what type of community the school was located in. The possible responses were:

1. urban;
2. suburban;
3. small city;
4. town; or,
5. rural.

School size: I calculated three categories of school size based on the number of students attending the school:

1. small, less than 350 students;

2. medium, between 350 and 749 students; or,
3. large, 750 or more students.

Theoretical Framework

Social psychology has a long tradition, dating back to the 1930s, of researching and debating humans' perceptions of one another and the role of accuracy and bias in those perceptions. Jussim, Eccles, and Madon (1989) do an excellent job summarizing this ongoing debate among social psychologists.

There are three reasons that a perceiver's expectations will be confirmed by the perceived person, labelled the "target" in social psychology literature. First: perceivers can create self-fulfilling prophecies, that is, their erroneous perceptions may result in the target actually fulfilling their perception. While their initial observation of the target was incorrect, the very fact that the perceiver feels a certain way about a target may influence the target's behavior so that the behavior actually comes to reflect the original, erroneous perception. A run on a bank is a typical example of perceptions influencing reality. A bank may be financially sound, but if its customers suspect differently, they will storm the bank in an effort to withdraw their deposits, thus causing the bank financial stress: a self-fulfilling prophecy. The customers' original judgments were erroneous, but the very fact they formed them, and then acted on them, resulted in those originally erroneous perceptions becoming quite accurate (Rosenthal & Jacobson, 1968).

Second: perceptions may be confirmed through perceptual biases. A perceiver may have an erroneous perception of a target and not have any significant effect on the target's behavior. However, the perceiver may perceive the target's behavior incorrectly and interpret it in ways that fit into his/her original perception (see, for example, Jussim, 1991).

A company's human-resources director may hold certain biases against African-Americans. If an African-American candidate applies for a job with that company, and is interviewed by that biased director, statements the applicant makes or the way in which they express themselves in the interview may "prove" to the director that the applicant is unfit for the job.

Third: perceptions may be accurate and reflect the social reality. Obviously, perceptions would be confirmed in this situation (Brophy, 1983; Jussim, 1991).

The role of perception is particularly important in the field of education. The seminal work in the area of perceptions and biases in education was Rosenthal and Jacobson's *Pygmalion in the Classroom* study (1968). By making teachers in a particular school believe that 20% of their students (whom Rosenthal and Jacobson had selected completely at random) were about to undergo an intellectual spurt — based on scores on a non-existent Harvard-University-designed test of "inflected acquisition"— the two researchers were able to influence how those students achieved. The randomly selected students performed significantly better than students who had not been selected. While there have been several criticisms of the *Pygmalion* study (see, for example, Elashoff & Snow, 1971), it was the first empirical study of the effect of teacher perceptions of their students.

Teacher perception has a powerful influence on children. It will influence their educational experience both subjectively *and* objectively. Teachers make profound decisions about their students based, at least partially, on their perceptions, decisions such as ability-group assignments, whether or not to recommend special education, or whether or not to retain a child (Burkham, LoGerfo, Ready, & Lee, 2007; Farkas, 2003; Page, 1987). If teachers' perceptions are influenced by socio-demographic biases, then existing,

documented educational inequality can worsen. Of course educational inequality leads to further economic and social inequality which many researchers point to as the strongest influence on educational inequality. Quite a vicious cycle, indeed.

Yet, teachers' perceptions may only reflect the existing social reality. African-Americans and Latinos do tend to score lower on standardized achievement tests than whites and Asians, and lower-income students tend to score lower than middle-and-upper-income students. To the extent that teachers' perceptions reflect actual socio-demographic trends in test performance and achievement, these are, by definition, accurate. Bias can be claimed only if teacher perceptions are inaccurate and the variation in that inaccuracy is systematic (Ready & Wright, 2011).

Some investigators theorize that biases can be explained by the socio-demographic characteristics of teachers themselves. Most teachers, at least elementary school teachers, are white, middle-income females. Some researchers posit that a predominantly white, female teaching corps may, on the average, misread cultural characteristics of minority students as indicators of academic ability (Delpit, 2006; Downey and Pribesh, 2004; Farkas, 2003; Lortie, 2002). Many laboratory experiments find that teachers tend to be biased against African-American students (Baron et al., 1985). I label these teacher perception researchers the experimentalists.

However, the majority of researchers that explore this link between teacher perceptions and actual student achievement scores find little or no evidence of teacher bias. In fact, they generally, with a few exceptions, find teacher perceptions to be quite accurate, when controlling for measured student achievement. I have labeled these researchers the naturalists because their studies are based on data on

real teachers and students not data on teacher perceptions of imaginary, idealized students in a laboratory setting.

The environmentalists, Ready and Wright (2011), however, using a nationally representative sample of kindergarten teachers and their students, find significant teacher bias in their perceptions of students. They argue that the variation in these biases can be explained more by the characteristics of a student's classroom than by the characteristics of his/her teacher, discounting the cultural-clash theories. According to Ready and Wright, teachers, regardless of their race and socioeconomic background, are more likely to underestimate the performance of children in classrooms where the average household income was low, but are not more likely to underestimate the academic performance of children in classrooms where the average household income was high.

Research Questions

There are five fundamental research questions that guide this study of teachers' perceptions of their students:

1. How accurate are teachers' perceptions of their students' performance when compared to standardized test results?
2. Do teachers' perceptions of minority students differ from their perceptions of white students?
3. How do teachers' perceptions vary by the student's grade level?
4. Do teacher characteristics explain the variation in the accuracy of teachers' perceptions?
5. Do school characteristics explain the variation in the accuracy of teachers' perceptions?

Questions 1 and 2 go to the heart of prior research on the topic of the accuracy of teacher perceptions of their students. The experimenters argue that teachers are biased against African-American students. The naturalists argue that teachers are quite accurate in their perceptions of their students, including their minority students. However, none of these studies has examined a national data set. And, Ready and Wright (2011), who do use a national data set, examine only kindergarten teachers. Questions 1 and 2 help me examine all grades on a national level. The hypothesis of the experimenters is that teachers, on the average, underestimate the ability of African-American students. The hypothesis of the naturalists is that teachers are accurate in their assessments of all students and, thus, are not biased in their perceptions. The environmentalists agree with the experimenters that teachers are biased, but argue that it has nothing to do with the personal characteristics of the teacher and everything to do with the environment that the teacher is practicing in.

Question 3 allows me to examine the grade level question in more detail. No previous study has examined all grades on a national level.

Question 4 examines the hypothesis posited by some of the early experimental researchers that whites, particularly females, are biased against minority children and will systematically underestimate their abilities. I also examined other teacher personal characteristics to determine any links they might have with teacher perception.

Question 5 directly examines Ready and Wright's environmental hypothesis, looking at the variables used in their model.

Limitations

The greatest limitations of this study arise from some of the deficiencies in the KEYS data set. The KEYS data set has many strengths. It is a particularly rich database for data on the demographics and opinions of teachers across the nation. This richness contributed to the research. However, some of its deficiencies limited my study. First, it does not include data on individual student achievement performance, nor on teachers' perceptions of individual students. The teacher perception data are the teachers' perceptions of their classrooms. Teacher perception of minority students data are the teachers' perceptions of the minority students in their classrooms. Similarly, teacher perception of white students are the teachers' perception of the white students in their classrooms. The student achievement data is an average of the test results of the students in the highest grade in the school. The student achievement data for minority students is an average of the test results of the minority students in the highest grade. I calculated the white student achievement data as I explain in more detail in the "Methodology" section. I compensated for the fact that the available student achievement data references only the highest grade in each school by limiting my analysis to teachers in that same grade in each school (again, see the Methodology section for details).

Second, the KEYS database is not constructed from a representative sample of all U.S. schools, rather, schools in the KEYS database have selected themselves to respond to the survey. Nonetheless, in the "Presentation and Analysis of the Data" section of this study I attempted to demonstrate that the KEYS data are fairly representative of schools in the United States as a whole.

Assumptions

Assumption 1: Data on teacher perceptions of an entire classroom can give us useful information on teachers' perceptions of their students.

Assumption 2: The KEYS data collection process was legitimate and provides useful data. This entails a number of further assumptions about that process, namely that:

1. the survey instrument was an accurate mechanism to capture opinions of teachers;
2. teachers' responses to the instrument reflected their true beliefs and opinions; and,
3. the data were collected in an ethical manner.

Assumption 3: Standardized student achievement tests are an accurate measure of student cognitive ability.

Assumption 4: The KEYS data set, while not a pure representative sample, does have sufficient characteristics that make it fairly representative of the public school teacher population of the United States as a whole.

Chapter 2: Review of the Literature

Introduction

This chapter presents the rationale for examining the relationship between teacher perceptions of their students' achievement levels and the students' performance on standardized tests.

The literature on biases and stereotyping in teachers' perceptions of their students is limited and rife with contradictions. This chapter will examine some of those contradictions. First I look at experimental research into teacher perceptions and how that research tends to focus on unconditional race neutrality. I then examine the concept of *conditional* race neutrality, which takes a more naturalistic look at the relationship between socio-demographic variables and teacher perceptions. Finally I examine a recent study by Douglas Ready and David Wright that finds interesting relationships between socio-demographic variables, characteristics of student classrooms, and teacher perceptions of their students' potential achievement.

Experimental Studies of Teacher Perception

The literature on bias and stereotyping in teacher perception of their students is filled with contradictory results. One reason is that researchers use different benchmarks to measure teacher bias.

Experimental research tends to use the benchmark of unconditional race neutrality. Expectations should be uncorrelated with race or ethnicity. By this benchmark, teachers who do not, on average, expect the same results from students of

different racial and ethnic groups, are biased. The typical experiment fabricates information about students and their race. Usually the experimenter takes care to prevent any correlation between race and other data in the fabricated data set. The researcher gives teachers this fabricated information and then asks them to predict how well the imaginary student will perform in the classroom, on tests, or against a specific measure of ability. Most of these experimental studies find that teachers have biased perceptions (Ferguson, 2003).

A typical study by DeMeis and Turner (1978) looked at 68 white, female teachers who were participating in summer classes at a university in Kentucky. The teachers listened to tapes of fifth-grade males responding to a question about their favorite television show. Accompanying each tape was a photograph of the student. The researchers asked the teachers to rate each student on personality, quality of response, as well as perceptions of current and future academic ability. The student's is a significant predictor of teachers' responses on all four categories.

In a meta-analysis of 16 experimental studies, Baron et al. (1985) find that in nine of the studies, teachers have higher expectations for white students compared to just one study that finds expectations biased in favor of African-American students. The meta-analysis concludes that teachers are biased towards white students.

Conditional Race Neutrality and Teacher Perceptions

The problem with unconditional racial neutrality is that African-Americans, whites, Latinos, Asians, and other racial and ethnic groups all, on average, do perform differently on the standardized tests that our society uses to measure school performance and cognitive ability. African-Americans and Latinos tend to score,

again, on average, lower than whites and Asians on standardized tests. Therefore, it is to be expected that teachers, on the average, would assume that African-Americans and Latinos will do less well than whites and Asians (see, for example, Jencks & Phillips, 1998, for an extensive look at the African-American-white testing-score gap; Lee & Burkham, 2002, for an in-depth look at the cognitive differences between racial/ethnic groups as children begin school). In fact, teachers do believe stereotypes about African-Americans and Latinos and we see them apply those biases in experimental situations. While these laboratory experiments are set up to create a perfect environment where every student of every ethnicity performs equally well on standardized tests, in the real world, teachers' perceptions might actually be quite accurate. The experimental data fail to persuade that teachers in the real world do not have accurate perceptions of their students (Ferguson, 2003).

Perhaps a more appropriate model might be to investigate the accuracy of teachers' perceptions of their real-life students, i.e., conditional racial neutrality, rather than unconditional racial neutrality in a laboratory. Numerous naturalistic studies find that teachers' perceptions of their students are quite accurate. While these studies find what, at first, appear to be biases in teachers' perceptions, these seeming biases actually reflect the real-world differences in their students' achievement scores. Thus, the teachers are biased in the unconditional sense, but because in the real world African-Americans and Latinos generally underperform white and Asian students on achievement tests, these perceptions, which would be biases in the unconditional sense, might indeed be objective.

In one study, Lee Jussim (1989) looked at all sixth-grade teachers in a public-school district in southeastern Michigan and most of their 634 students. Early in the school year (October), he asked teachers to evaluate each of their students on three factors:

1. math performance;
2. math talent; and,
3. math effort.

In October, and then again in March, Jussim asked students to judge themselves in the following areas:

1. math ability;
2. math effort;
3. time spent on math homework; and,
4. value they place on math.

Jussim also collected standardized test data and math grades for each student. He looks at each student's standardized math score at the beginning of sixth and seventh grades, as well as each student's fifth-and-sixth-grade math grades.

Jussim then uses path analysis to examine the relationship between teacher perceptions, student motivation, and student achievement. While he finds a small amount of teacher bias (e.g., incorrect perceptions of final achievement), and some self-fulfilling prophecies (e.g., some students did better because the teachers expected them to do better), by far the strongest variable explaining the variation in student achievement is the accuracy of the teachers' perceptions of their students.

In another study, Meisels, Bickel, Nicholson, Xue, and Atkins-Burnett (2001) examine the validity of a particular curriculum-embedded performance assessment, the Work Sampling System (WSS). The WSS depends heavily on teacher judgment.

Meisels et al. wanted to compare teacher judgments based on the WSS with other indicators of student performance. Drawing from a group of volunteers in the Pittsburgh public-school district, the researchers selected 17 teachers who were ranked in the highest quartile of WSS users by external examiners. The 345 students in these teachers' classrooms were each administered a standardized test.

Meisels et al. use four-step hierarchical linear modeling and ROC (receiver operating characteristic) curve analysis to determine if the WSS data make a unique contribution in explaining the variation in student achievement (as measured by the standardized test and controlling for the child's age, gender, socioeconomic status, race, and initial performance on the standardized test).

The investigators find that the WSS model, was indeed, a reliable predictor of student achievement (as measured by the standardized test employed). They also find that data from the WSS were reliable for determining a student's at-risk status. Meisel et al. conclude that teachers' judgments, when using the WSS, are reliable for assessing student performance.

In a literature review for the National Center for Education Statistics (NCES), Nancy Perry and Samuel Meisels (1996) conclude that teachers' perceptions of student performance were fairly accurate. NCES was interested in devising cost-efficient methods for collecting data on student achievement for its Early Childhood Longitudinal Study (ECLS). They were interested in knowing if teachers' perceptions of student achievement were accurate and how that could inform the development of measures of student achievement within ECLS.

Perry and Meisels, in their literature review, find that overall, teachers' perceptions of student achievement are accurate. They note that some researchers even find that teacher judgments were more accurate than standardized tests. While they find some teachers had difficulty understanding researchers' constructs, for example, "motivation", they find that teachers tend to "judge these constructs independently" (Perry & Meisels, p. 28) and do not have biases based on gender and behavioral characteristics of students. Perry and Meisels conclude that evidence for teacher bias is weak, and recommend that NCES further explore the validity of using actual teacher judgments as a measure of student achievement.

Ferguson's (2003) literature review also finds that teachers' judgments are accurate, once test score differences between African-Americans and whites are taken into account.

Tach and Farkas (2006) use data from the ECLS as well to examine reading-ability group placement and its results. They find that while teachers were more likely to assign African-American, Latino, male, and younger-aged first-graders to lower-level reading groups, most of the differences can be explained by controlling for reading test scores and social class. They find no statistically significant interactions between the student's and the teacher's race. However, they do find that children of lower socioeconomic status (SES) are more likely to be placed in lower-ability groups. Burkham, LoGerfo, Ready, and Lee (2007) also find similar results with kindergarten students using the ECLS data: male students, low-SES students, and children who entered kindergarten at a younger age are more likely to be held back in kindergarten.

Two major literature reviews (Ferguson, 2003 and Farkas, 2003) on teacher perceptions' and bias sum up the research by concluding that while there is some evidence of teacher bias, it is not very strong or convincing. They find that it is quite difficult to study this issue and the existing evidence is quite fragmentary. The crux is the difficulty in separating stereotyping from actual socio-demographic differences in student achievement.

Ready and Wright on Teacher Perception

A recent study (Ready & Wright, 2011), seeks to disentangle these issues. This study deserves particular attention both because it is the first teacher perception study to use a national database and because of its groundbreaking findings.

Ready and Wright start from the notion of conditional neutrality. They borrow from Jussim, Eccles, and Madon (1996) in defining bias as "systematically evaluating two groups as differing on some criterion more or less than they really do differ" (p. 329). Teacher perceptions will only be "biased to the degree that they over-or understate actual between group differences" (Ready & Wright, 2011, p. 338). Ready and Wright define three possible ways that teacher perceptions can be categorized:

1. accurate and unbiased – when teachers' perceptions of children's academic skills match objective measurements of those skills, i.e., achievement tests;
2. inaccurate and unbiased – when teachers' perceptions of children's academic skills do not match objective measurements of those skills, but the variation is random and non-systematic; and,

3. inaccurate and biased – when teachers' perceptions of children's academic skills do not match objective measurements of those skills and the teachers' perceptions vary systematically with socio-demographic variables.

Ready and Wright use data from the ECLS kindergarten cohort (ECLS-K) to investigate the possible presence of teacher bias in their perceptions of students, and possible explanations for any biases found. They specifically choose to look at only the kindergarten cohort in the data. They feel that kindergarten teachers, having no past records on which to base judgments, are more likely to use their previous experience with children from different socio-demographic groups to develop their first perceptions of those children's potential.

ECLS-K asks teachers to rate each child's skills based on their experience with the child in the fall and spring of their kindergarten year. While ECLS-K asks teachers to rate students on language and literacy, mathematics, and general knowledge, Ready and Wright look only at the language and literacy assessment. In the language and literacy section of the teacher assessment, teachers are asked to judge students in five areas:

1. speaking;
2. listening;
3. early reading;
4. writing; and,
5. computer literacy.

ECLS-K administers a cognitive/academic assessment test to individual children in the sample both in the fall and the spring. While these tests covered reading,

mathematics, and general knowledge, Ready and Wright use only the reading portion in their research. Data are also collected by ECLS-K on the characteristics of the children, the teachers, the classroom, and the school (National Center for Education Statistics, n.d.).

Ready and Wright use a three-level hierarchical linear model to examine teacher bias. They nest children within classrooms, within schools. Teacher assessments of each student's ability is the dependent variable in these models. The model's independent variables included:

1. ECLS-K literacy assessment
2. child characteristics:
    a. dummy variables for race/ethnicity (whites serving as the uncoded comparison group);
    b. socioeconomic status (composite of parents' income, education, and occupational prestige);
    c. age;
    d. gender;
    e. single-parent status;
    f. English as a second language (ESL);
    g. repetition of kindergarten;
    h. number of siblings at home;
    i. Latino-ESL interaction variable; and,
    j. Asian-ESL interaction variable.
3. classroom characteristics:
    a. dummy variables for race/ethnicity of teacher (whites serving as the uncoded comparison group);
    b. dummy variables for educational attainment of teacher (bachelor's degree serving as the uncoded comparison group);

c. dummy variable for teaching experience of teacher (1 = less than three years' experience);
d. aggregate teacher perception of all the children in the classroom in fall and spring;
e. average class SES;
f. dummy variable for minority concentration (> 70% non-white, non-Asian); and,
g. dummy variable for ESL concentration (> 20% ESL).

4. school characteristics:
   a. average SES;
   b. dummy variables for location (suburbs serving as the uncoded comparison group);
   c. dummy variables for sector (public schools serving as the uncoded comparison group); and,
   d. dummy variables for size (with medium-size schools, 350 to750 students, serving as the uncoded comparison group).

Ready and Wright find that, indeed, teachers tend to underrate the potential achievement of boys, African-Americans, Latinos, low-SES students, Latino ESL students, Asian ESL students, younger students, students from single-parent homes, and students who indicated they had siblings when they are first queried in the fall. However, when controlling for actual performance on the achievement test, nearly half the disparity is eliminated, and children from single-parent households are no longer found to be significantly underrated by their teachers. In other words, actual socio-demographic differences between groups of students explain half the original variance.

Teacher's perceptions become more accurate as the school year progressed. However, by the spring, controlling for actual performance, they still tend to

underestimate the potential of males, Latinos, low SES students, Asian ESL students, younger students, and students with siblings. Interestingly, in the spring, teachers who earlier had fairly accurate perceptions of the potential of kindergarteners who were repeating the level, begin to seriously underestimate their potential. The most serious teacher biases are for male students and low-SES students.

Ready and Wright then attempt to determine the source of these biases, particularly the tendency to underrate low-SES students. They find that the misperceptions of teachers are more related to the classes and schools in which they operate, than to individual characteristics of the teachers themselves. When controlling for the characteristics of their students, only one teacher characteristic is significantly related to teacher accuracy: new teachers tend to overestimate their students' potential. However, both the average class achievement score and the average class SES are found to be related to teacher accuracy to teachers of all experience levels. Teachers tend to overestimate the abilities of children in high-achievement, high-SES classrooms, and, conversely, underestimate the abilities of low-achievement, low-SES classrooms.

Using slopes-as-outcomes models, Ready and Wright find that these tendencies were more exacerbated in low-SES classrooms. There, teachers' inaccurate perceptions of low-SES students were even more pronounced (on average) than in high-SES classrooms.

Summary

The literature on teacher perceptions tends to be contradictory. Experimental studies often find a great deal of teacher bias in their perceptions of students, while

more naturalistic studies find less bias and greater teacher accuracy in perception. One explanation is that experimental studies measure bias unconditionally and fail to account for the real-world experience of most teachers where students of different socio-demographic groups have, on average, different levels of achievement. More naturalistic studies that take socio-demographics into account, usually find small, or no perceptual bias by teachers. While some naturalistic studies find biases, the research has been quite fragmentary.

By including an additional set of variables in their model, Ready and Wright *do* find significant bias in teachers' perceptions of their kindergarten students' potential, but conclude that this bias was less due to the characteristics of individual teachers and more a result of the characteristics of the schools and classrooms in which they taught. Ready and Wright conclude that teachers in classrooms where the average achievement is low and/or classrooms where the students have, on average, low socioeconomic status, are more likely to underestimate their students' potential, regardless of the teacher's own race, income, education, or experience level.

The Ready and Wright study is an important step forward in the literature. It is one of the first studies of a national sample, it demonstrates systematic bias in teachers' perceptions under the conditional race neutrality hypothesis, and finds that environmental factors are important predictors of bias.

This study extends the literature on teacher perceptual accuracy by examining the hypotheses of the experimental, naturalistic, and environmental researchers with a national database at all grade levels. Additionally, I add examine the role of gender as a factor in teacher perception.

Chapter 3: Methodology

Most of the naturalistic researchers in the field of teacher perceptions argue that teachers are quite accurate in their perceptions of their students. The environmentalists, Ready and Wright (2011), however, have established a strong statistical relationship showing significant bias in kindergarten teachers' perceptions of their students. They find that teachers systematically believe that male, African-American, Latino, and low-income students were more likely to have weak literary skills. However, if teacher perceptions are controlled for by the achievement scores of those same students, the overall variance is reduced by half, and teachers *are not* more likely to judge African-American students as weak in literary skills than white students. The remaining variance is explained more by the characteristics of the students' classrooms and school than by the characteristics of their teachers. Teachers in classrooms that have high numbers of lower-income students, and low-achieving students, are more likely to underestimate their students' performance.

Access to the National Education Association's (NEA) Keys to the Excellence of Your Schools (KEYS) database gave me the opportunity to see if I could verify this relationship between teacher perception and student achievement, and see if it holds true in grades above kindergarten. Additionally, I could examine the role of the teacher's gender in teacher perception.

KEYS was a "comprehensive, research-based, data-driven program for continual school improvement" (National Education Association, n.d.). NEA supported the KEYS initiative to foster school improvement. KEYS used a self-administered survey

of school staff and parents to identify areas where individual schools can improve teaching and learning. Over 1,800 schools have participated in the KEYS initiative.

The latest version of KEYS had 42 indicators of school quality, grouped into six broad groups:

1. shared understanding and commitment to goals;
2. open communication and collaborative problem solving;
3. continuous assessment for teaching and learning;
4. personal and professional learning;
5. resources to support teaching and learning; and,
6. curriculum and instruction.

In particular, teacher perceptions are addressed by three survey questions:

1. On average, what is the performance level of *all* students in your TARGET CLASS?
2. On average, what is the performance of *racial and ethnic minority* students in your TARGET CLASS?
3. On average, what is the performance of *Caucasian, not of Hispanic origin,* students in your TARGET CLASS?

The "target class" is defined on the survey instrument as the class in which the teacher spends the majority of his/her time, or, if he/she teaches multiple classes of equal length, the first class of the week that he/she teaches. The possible responses to all three teacher perception questions were as follows:

1. primarily high achieving;
2. primarily average to high achieving;
3. primarily average achieving;
4. primarily average to low achieving; and,
5. primarily low achieving.

My broad approach was to examine the various theories propounded by the three schools of teacher perception that I have identified: the experimentalists, the naturalists, and the environmentalists. I tackled the question of whether or not teachers are accurate in their perceptions of students and then took a look at their perceptions of minority students and white students. Ready and Wright (2011) had the first teacher perception study that used national data. All previous studies had used local or state data. However, Ready and Wright only used kindergarten teachers. This study extends the literature by looking at national data from all grades.

Methodologically, I took an unusual approach for analysis of variance (ANOVA). Typically in ANOVA, one starts with factorial ANOVA and then optionally proceeds to examine two-way and one-way ANOVAs of the independent variables. Because I was working with a limited data set in terms of the number observations (See Table 1 for a breakdown of independent variables by gender), I knew that it would be difficult to meet all the statistical assumptions (e.g., homogeneity of variance) for ANOVA if I expanded the number of independent variables to three or four in a model. Furthermore, I wanted to follow the development of the study in the field of teacher perception. Most early studies simply examine whether teachers are accurate in their perceptions of their students. The environmentalists, Ready and Wright (2011), have the most advanced model, examining the effect of several independent variables on teacher perception of students. I examined seven of those independent variables, the ones that were available on the KEYS database. I used those seven variables as the basis of my analysis, examining each of them (and variations) in one-way ANOVAs. I then examined selected ones in two-way ANOVA based on hypotheses from the

teacher perception literature. Finally, I examined a factorial ANOVA with four variables, straining the limits of the KEYS data set.

Table 1: *Demographic Information for All Teachers and Schools They Teach In: Race, Education Level, Certification, Years' Experience, Socioeconomic Status of School, Community of School, and Size of School*

Category	Description	N Males	% Males	N Females	% Females
Race					
	American Indian	9	1.7	13	0.8
	Asian	6	1.1	25	1.6
	Black	32	6.0	102	6.5
	Caucasian	444	83.1	1,331	85.2
	Hispanic	30	5.6	75	4.8
	Other	13	2.4	16	1.0
Education					
	H.S.	3	0.6	19	1.2
	A.A.	5	0.9	16	1.0
	B.A.	253	47.6	710	44.9
	M.A.	236	44.4	752	47.5
	Ed Spec	19	3.6	73	4.6
	Ph.D.	16	3.0	13	0.8
Certification					
	Certified	498	93.4	1,425	90.7
	Trained	20	3.8	105	6.7
	Not Trained	10	1.9	30	1.9
	Other	5	0.9	11	0.7
Total Exp					
	< 2 Years	38	7.2	84	5.4
	2 – 5 Years	80	15.1	268	17.1
	6 – 10 Years	125	23.5	317	20.3
	11 – 19 Years	113	21.3	440	28.1
	20+ Years	175	33.0	455	29.1
System Exp					
	< 2 Years	71	14.0	173	11.5
	2 – 5 Years	112	22.1	370	24.5
	6 – 10 Years	99	19.5	322	21.4
	11 – 19 Years	97	19.1	351	23.3
	20+ Years	128	25.2	292	19.4
Building Exp					
	< 2 Years	97	19.0	271	17.9
	2 – 5 Years	140	27.4	482	31.9
	6 – 10 Years	117	22.9	352	23.3
	11 – 19 Years	96	18.8	265	17.5
	20+ Years	61	11.9	141	9.3
SES (school)					
	Upper Middle	17	3.2	75	4.8
	Middle	160	30.0	457	29.1
	Lower Middle	171	32.1	482	30.7
	Lower	185	34.7	558	35.5

Category	Description	N Males	% Males	N Females	% Females
Community					
	Urban	127	23.5	352	22.2
	Suburban	108	20.0	340	21.4
	Small City	150	27.8	381	24.0
	Town	66	12.2	245	15.4
	Rural	89	16.5	269	17.0
School Size					
	Small	111	20.6	263	16.6
	Mid-Size	275	50.9	948	59.7
	Large	154	28.5	376	23.7
School Type					
	Elementary	252	49.8	990	65.6
	Middle School	171	33.8	381	25.3
	High School	83	16.4	137	9.1

I controlled for Type I error across the 64 different hypothesis tests at the $\alpha = .05$ level by using Holm's sequential Bonferroni adjustment.

Research Question 1

The KEYS data are problematic in that student-achievement data were collected at the school level and not the classroom level. Scores were also reported only for the highest grade in the school. Therefore, I knew each school's "score", but I did not know how that relates to individual classrooms. Even more problematic, the school score could be reported in four different ways by the KEYS facilitator (the school-level respondent):

1. percentile score;
2. standard score;
3. stanine score; and,
4. percentage of students at and above average performance.

Because each state has its own standardized scoring system, I decided to eliminate schools in category two above, i.e., those that reported standardized scores directly, entirely from the study.

I also limited the study to the teachers in the highest grade in each school. I determined which teachers to include in my study by examining data for each school individually. I noted how each school was labeled in the database (e.g., "elementary"). I then searched for the highest grade within the category which the school was labeled. If the highest grade was incongruous with the school label in the database, I looked up that school in the National Center for Education Statistics (NCES) Common Core database. If NCES had labeled the school differently than KEYS, I changed the school type label to match NCES' label and picked the teachers in the highest grade. If the school labels were similar, I picked the highest "correct" grade and ignored teachers in grades higher than that (e.g., in an elementary school, I would select 6th grade teachers even if there were 7th and 8th grade teachers listed under that school). If I could not find the school in the NCES Common Core database, I deleted all data for the school and its teachers from my database. In all, I deleted information on teachers from 55 different schools, schools which had I had singled out because of incongruous data on grade level and that also had no data in the NCES Common Core database. I did not include teachers whose target class was labelled as "mixed/combined classes."

Because of the complications of using the KEYS achievement data, I "standardized" the school test data into quintile data. I determined quintile cutoff points for each of the three remaining different ways that standardized test scores were reported. That way, I came up with a standardized measurement of test scores across the three types of reporting.

Next, I created a new variable, which I labeled "perceptual match", to measure how closely each teacher's perceptions of his/her individual classroom matched the school's

test quintile. Recall, teacher's perceptions were recorded on a scale of 1 – 5, with "1" meaning the teacher rated the class as high achieving and "5" meaning that the teacher rated the class as low achieving. I reversed these scores (1=5, 2=4, 3=3, 4=2, 5=1) and then subtracted the school's test quintile from the result to create my "perceptual match" variable. An example will clarify my methodology.

Suppose a teacher rated her class as average to high achieving. In the KEYS database, this teacher would have been originally assigned a value of 2 for that rating. I reversed that rating so that the teacher now had a value of 4 for her rating of her classroom. Suppose further that the school was in the 3rd quintile of test scores. The value for my new perceptual match variable would be equal to 4 – 3 = 1. Assuming that the standardized test score of the school is a valid measure of the cognitive abilities of the teacher's classroom, then this teacher has overestimated the ability of her classroom. If the school's test score had been in the 4th quintile, the teacher's estimate of her classroom's abilities would have been accurate, and her perceptual match score would have been 0. Conversely, if the school's score had been in the 5th quintile, this teacher would have received a perceptual match score of -1. She would have underestimated the ability of her classroom. A positive perceptual match score indicates that a teacher has overestimated his/her classroom; a negative score indicates that a teacher has underestimated his/her classroom. The perceptual match score also gives an intuitive measure by which to analyze my results. Each quintile represents 20 percentile points. So, for example, a perceptual match score of 0.5 means that the teacher overestimates his/her classroom by 10 percentile points.

I examined this new variable, perceptual match, by running a univariate analysis of it and then several one-way analyses of variance (ANOVAs) on various teacher characteristics to determine how accurate teacher perceptions of their classrooms were.

There are four assumptions associated with using ANOVA correctly. First, data should come from a random sample of the population. The KEYS data violate that assumption because schools select themselves to be included in the KEYS database. However, I will demonstrate in the Results chapter that the KEYS data are a relatively good reflection of the population data and, thus, can be assumed to approximate randomly sampled data. Second, observations should be independent. Perceptual match scores for one teacher should not influence the scores of other teachers. Perceptual match scores meet this second assumption (Huck, 2008).

Third, ANOVA assumes approximate normality of the data. I performed a Shapiro-Wilk test of normality for the distribution of perceptual match over the values of each independent variable I used in my ANOVA analyses. I then examined the skewness and kurtosis for the values of each variable that was found not to be normally distributed. If any of the distributions of perceptual match were found to be skewed or kurtosed, I then visually inspected its Normal Q-Q Plot for approximate normality.

Fourth, ANOVA assumes homogeneity of the variances. I used Levene's test for equality of variances to ensure that this assumption was met. When Levene's test revealed that this assumption was not met, I used the Welch's F test in place of an ANOVA (Huck, 2008).

Table 2, below, describes the results of my examination of the normality and homogeneity-of- variance assumptions for the variables I used in this analysis. As an

example of how to interpret this table, I will walk the reader through the results from Table 1 for the independent variable gender. A Shapiro-Wilk test of normality found that the distribution of perceptual match for male and female teachers was not normal ($p < .001$ for both genders). Examining the skewness and kurtosis of each distribution, I found that perceptual match for males was normally distributed but that the distribution for females was negatively kurtosed. With large sample sizes these tests can often be overly sensitive to deviations from normality. From visual inspection of the Normal Q-Q Plot I determined that perceptual match was approximately normally distributed for females. I found no extreme outliers in the data. There was homogeneity of the variances, as assessed by Levene's test for equality of variances ($p = .599$).

Table 2: *ANOVA Assumptions Testing for Perceptual Match: Summary Table*

Independent Variable	Value	Shapiro Wilk	Skewness	Kurtosis	Homogeneity of Variance
Gender					**.599**
	Male	<.001	none	none	
	Female	<.001	none	negative	
Race					**<.001**
	Native	.080	-	-	
	Asian	.059	-	-	
	Black	.007	none	none	
	Caucasian	<.001	none	negative	
	Hispanic	.001	none	none	
	Other	.120	-	-	
Race					**.010**
	Non-white	<.001	none	none	
	White	<.001	none	negative	
Education					**.037**
	High School	.025	–	–	
	Associate's	.007	none	none	
	Bachelor's	<.001	none	none	
	Master's	<.001	none	negative	
	Ed Specialist	.001	none	none	
	Doctorate	.010	none	none	
Certification					**.923**
	Certified	<.001	none	negative	
	Not certified but trained	.001	none	none	

Independent Variable	Value	Shapiro Wilk	Skewness	Kurtosis	Homogeneity of Variance
	No training	.043	–	–	
	Other	.002	none	none	
Total Experience					**.250**
	< 2 years	.001	none	none	
	2-5 years	<.001	none	none	
	6-10 years	<.001	none	none	
	10-20 years	<.001	none	none	
	20+ years	<.001	none	none	
System Experience					**.243**
	< 2 years	<.001	none	none	
	2-5 years	<.001	none	none	
	6-10 years	<.001	none	none	
	10-20 years	<.001	none	none	
	20+ years	<.001	none	none	
Building Experience					**.100**
	< 2 years	<.001	none	none	
	2-5 years	<.001	none	none	
	6-10 years	<.001	none	none	
	10-20 years	<.001	none	none	
	20+ years	<.001	none	none	
Total Experience					**.142**
	< 2 years	.001	none	none	
	>2 years	<.001	none	negative	
System Experience					**.086**
	<2 years	<.001	none	none	
	>2 years	<.001	none	negative	
Building Experience					**.427**
	<2 years	<.001	none	none	
	>2 years	<.001	none	negative	
Socioeconomic Status					**<.001**
	Upper Middle	<.001	none	none	
	Middle	<.001	none	none	
	Lower Middle	<.001	none	none	
	Lower	<.001	negative	none	
Community					**.136**
	Urban	<.001	negative	none	
	Suburban	<.001	none	none	
	Small City	<.001	none	none	
	Town	<.001	none	none	
	Rural	<.001	none	none	
School Size					**.103**
	Small	<.001	none	negative	
	Medium	<.001	none	none	
	Large	<.001	none	none	
School Type					**<.001**
	Elementary	<.001	none	negative	
	Middle	<.001	none	none	

Independent Variable	Value	Shapiro Wilk	Skewness	Kurtosis	Homogeneity of Variance
	High School	<.001	none	negative	

Research Question 2

My second research question was to determine if teachers' perceptions of minority students are different than their perceptions of white students. The methodology was identical to the analysis employed in my first research question. However, instead of examining how teachers perceived all the students in their classes, I looked at how they perceived the racial and ethnic minority students and the white students in their classes.

Because of the complications of using the KEYS achievement data, I "standardized" the school test data for ethnic and racial minority students into quintile data. I used the quintile cutoff points calculated for all students' data for each of the three remaining different ways that standardized test scores were reported (recall that I had eliminated all schools that reported direct standardized testing data as discussed above). This is similar to the calculations I used in standardizing the test data for all students. The difference is that while I used minority student testing data, I assigned quintiles based on the cutoff points for *all* students. That way, I came up with a standardized measurement of test scores for ethnic and minority students across the three types of reporting.

Calculating standardized test scores for white students was slightly more complicated because extensive data on white student test scores did not exist in the KEYS database. However, since the database does give complete information on the racial makeup of the student body and I have test score data for the entire student body and for minority students, it was possible, using simple algebra, to calculate a standardized test score for white students. I followed the same procedure I used in calculating test score quintiles for

minority students, using quintile cutoff points for *all* students to assign a quintile number for each school for white student standardized test scores.

I created two new variables, perceptual match minority and perceptual match white. Perceptual match minority measures how closely each teacher's perceptions of the racial and ethnic minority students in his/her classroom matched the school's test quintile for minority students. Similar to how I created the variable perceptual match, I reversed the scores of how the teacher perceived the minority students in his/her classroom and then subtracted the school's minority test quintile from that number. I followed the same procedure for perceptual match white, substituting the teacher's perception of the white students in his/her classroom and the newly calculated standardized test performance of white students at the school into the equation. I then used simple univariate analysis to determine if there were statistically significant differences between the perceptual match score for minority students and the perceptual match score for white students.

Tables 3 and 4, below, describe the results of my examination of the normality and homogeneity-of-variance assumptions for the distributions of perceptual match minority and perceptual match white over the independent variables I used in this analysis.

Table 3: *ANOVA Assumptions Testing for Perceptual Match Minority: Summary Table*

Independent Variable	Value	Shapiro Wilk	Skewness	Kurtosis	Homogeneity of Variance
Gender					**.283**
	male	<.001	none	none	
	female	<.001	none	negative	
Race					**.567**
	Native	.054	-	-	
	Asian	.006	none	none	
	Black	.004	none	none	
	Caucasian	<.001	none	negative	
	Hispanic	<.001	negative	none	
	Other	.122	-	-	

Independent Variable	Value	Shapiro Wilk	Skewness	Kurtosis	Homogeneity of Variance
Race					**.103**
	Non-white	<.001	none	none	
	White	<.001	none	negative	
Education					**.007**
	High School	.145	-	-	
	Associate's	.165	-	-	
	Bachelor's	<.001	none	none	
	Master's	<.001	none	negative	
	Ed Specialist	.001	none	none	
	Doctorate	.193	-	-	
Certification					**.412**
	Certified	<.001	none	negative	
	Not certified but trained	.007	none	none	
	No training	.137	-	-	
	Other	.893	-	-	
Total Experience					**.378**
	< 2 years	.001	none	none	
	2-5 years	<.001	none	none	
	6-10 years	<.001	none	none	
	10-20 years	<.001	none	none	
	20+ years	<.001	none	none	
System Experience					**.150**
	< 2 years	<.001	none	none	
	2-5 years	<.001	none	none	
	6-10 years	<.001	none	none	
	10-20 years	<.001	none	none	
	20+ years	<.001	none	none	
Building Experience					**.257**
	< 2 years	<.001	none	none	
	2-5 years	<.001	none	none	
	6-10 years	<.001	none	none	
	10-20 years	<.001	none	none	
	20+ years	<.001	none	none	
Total Experience					**.822**
	< 2 years	.001	none	none	
	>2 years	<.001	none	negative	
System Experience					**.055**
	<2 years	<.001	none	none	
	>2 years	<.001	none	negative	
Building Experience					**.929**
	<2 years	<.001	none	none	
	>2 years	<.001	none	negative	
Socioeconomic Status					**.064**
	Upper Middle	.006	none	none	
	Middle	<.001	none	none	
	Lower Middle	<.001	none	none	

Independent Variable	Value	Shapiro Wilk	Skewness	Kurtosis	Homogeneity of Variance
	Lower	<.001	negative	none	
Community					**.091**
	Urban	<.001	negative	none	
	Suburban	<.001	none	none	
	Small City	<.001	none	none	
	Town	<.001	positive	none	
	Rural	<.001	none	none	
School Size					**.547**
	Small	<.001	none	none	
	Medium	<.001	none	none	
	Large	<.001	none	none	
School Type					**<.001**
	Elementary	<.001	none	negative	
	Middle	<.001	none	none	
	High School	<.001	none	none	

Table 4: *ANOVA Assumptions Testing for Perceptual Match White: Summary Table*

Independent Variable	Value	Shapiro Wilk	Skewness	Kurtosis	Homogeneity of Variance
Gender					**.231**
	male	<.001	none	none	
	female	<.001	none	negative	
Race					**.008**
	Native	.072	-	-	
	Asian	.002	none	none	
	Black	.010	none	none	
	Caucasian	<.001	none	negative	
	Hispanic	.026	-	-	
	Other	.245	-	-	
Race					**.019**
	Non-white	<.001	none	none	
	White	<.001	none	negative	
Education					**.289**
	High School	.025	-	-	
	Associate's	.066	-	-	
	Bachelor's	<.001	none	none	
	Master's	<.001	none	negative	
	Ed Specialist	<.001	none	none	
	Doctorate	.023	-	-	
Certification					**.829**
	Certified	<.001	positive	negative	
	Not certified but trained	.001	none	none	
	No training	.118	-	-	
	Other	.736	-	-	

Independent Variable	Value	Shapiro Wilk	Skewness	Kurtosis	Homogeneity of Variance
Total Experience					**.358**
	< 2 years	.004	none	none	
	>2 years	<.001	none	negative	
System Experience					**.010**
	<2 years	<.001	none	none	
	>2 years	<.001	positive	negative	
Building Experience					**.666**
	<2 years	<.001	none	none	
	>2 years	<.001	positive	negative	
Socioeconomic Status					**<.001**
	Upper Middle	<.001	none	none	
	Middle	<.001	none	none	
	Lower Middle	<.001	none	none	
	Lower	<.001	none	negative	
Community					**<.001**
	Urban	<.001	none	none	
	Suburban	<.001	none	negative	
	Small City	<.001	none	none	
	Town	<.001	positive	none	
	Rural	<.001	none	none	
School Size					**.438**
	Small	<.001	none	none	
	Medium	<.001	none	none	
	Large	<.001	none	none	
School Type					**.035**
	Elementary	<.001	none	negative	
	Middle	<.001	none	none	
	High School	<.001	none	none	

Research Question 3

My third research question looks at how grade level influences teacher perception. The Ready and Wright study, the sole research effort on teacher perceptions to look at a national database, only looked at kindergarten teachers and their students. KEYS gives me the opportunity to look at all grade levels. I used a one-way ANOVA to determine if the teacher's grade makes any difference in their perception of his/her students.

Research Question 4

My fourth research question examines the connection between teacher characteristics and the accuracy of their perceptions. I examined the ANOVAs that I ran for my first research question to see if any of the teacher characteristics had a significant influence on teacher perceptual match. I also ran 2-way and factorial ANOVAs to determine if there were any interaction effects amongst the teacher characteristics variables. I examined teachers' perceptions of all their students, their racial and ethnic minority students, and their white students. Because Ready and Wright have the most comprehensive model of teacher characteristics in their study of teacher perceptual accuracy, I wanted to examine the same independent variables that they do in their study. There are four variables in the KEYS database that correspond to the teacher characteristics that Ready and Wright used in their analysis:

1. teacher race/ethnicity;
2. teacher education level;
3. teacher certification in the target class subject matter; and,
4. teacher experience.

Research Question 5

Again, because Ready and Wright have the most comprehensive teacher perceptions model, I wanted to use the same environmental variables they had used. I used the same ANOVA methodology to determine if any of the following school characteristic variables from KEYS have a significant impact on teachers' perceptions:

1. school community type:
    a. urban;
    b. suburban;

c. small city;
d. town; and,
e. rural;

2. socioeconomic status of the parents of the students served by the school; and,
3. size of the school.

As my final step, I attempted to construct a multi-factorial ANOVA model that incorporated those variables that Ready and Wright had found to be significant in their analysis as well as the variable gender. The role of gender is very important in the theoretical literature around the question of teacher perceptions.

Chapter 4: Presentation and Analysis of the Data

Introduction

Teachers' perceptions of students can have a wide range of influences on the children they teach, often influencing a number of educational placement decisions, including grouping by ability, grade retention, exposure to curricula, admission to selective programs or schools, and assignment to English as a second language and/or special education classes. These placement decisions, in turn, can have profound impacts on children's lives. Moreover, if teacher perceptions are biased in any way that impacts minority children negatively, the current educational inequality afflicting our nation may worsen.

There are a number of studies that posit that teachers' perceptions are quite accurate and do not reflect systematic bias in education. These studies argue that teachers' perceptions reflect real cognitive differences between socio-demographic groups (see Jussim, 1989; Eccles & Madon, 1996; Jussim & Harber, 2005; Meisels, et al., 2001; Xue & Atkins-Barnett, 2001; Perry & Meisels, 1996). I label these researchers the naturalists. However, Ready and Wright's groundbreaking 2011 study argues that teachers do have substantial biases in their perceptions of students from different socio-demographic groups. According to Ready and Wright, this bias is accounted for more by the characteristics of the classroom than by the characteristics of the teachers themselves. Teachers working in low-income neighborhoods are more likely to underestimate their students' abilities than teachers working in higher-income neighborhoods, regardless of the teacher's race, income, or other background characteristics. So it is not so much the

teacher's background that leads to biases, it is more the context in which the teacher finds him/herself that leads to biased perception. I label Ready and Wright the environmentalists.

Ready and Wright, while important for being the first teacher perception study to use national data, and for proposing the environmental theory of teacher perception, only examined kindergarten teachers. This study looks at national data from all grades to examine the various theories of teacher perception, including Ready and Wright's environmental theory of perception. To that end, five research questions guided me:

1. How accurate are teachers' perceptions of their students' performance when compared to standardized test results?
2. Do teachers' perceptions of minority students differ from their perceptions of white students?
3. How do teachers' perceptions vary by grade?
4. Do teacher characteristics explain the variation in the accuracy of teachers' perceptions?
5. Do school characteristics explain the variation in the accuracy of teachers' perceptions?

This study used the KEYS database to investigate the research questions. The latest version of KEYS has 42 indicators of school quality grouped into six broad categories. My approach was to examine the relationship between teachers' responses to questions about their perception of the abilities of their "target" class with school-wide data on student performance on standardized tests.

I wanted to compare KEYS data with national statistics from the Digest of Education Statistics to determine if my actions had introduced substantial biases into the KEYS

database that remained after my edits. To my pleasant surprise, the KEYS data matched national statistics quite accurately (with the exception of one data point, described below).

The KEYS data I selected differ substantially from national statistics in the distribution of students between elementary and secondary schools. Using the Digest's definition, which assigns all ninth graders to secondary school and all students in eighth grade or below to elementary school, I found that the KEYS database substantially overweights elementary schools (Table 5).

Table 5: *Comparison of Distribution of Students between Elementary and Secondary Schools in the KEYS Data Set and the Digest of Education Statistics*

Type of School	*Digest of Education Statistics*	KEYS Database
Elementary	70.0%	80.3%
Secondary	30.0%	19.7%

I found smaller discrepancies when I compared KEYS data on student racial/ethnic background to national data (Table 6). White and African-American students were somewhat overweighted in the KEYS data while Latino and Asian Americans were somewhat underweighted compared to the data in the *Digest*. This is understandable in light of U.S. demographic trends where the percentage of Latino and Asian students has been rising, while the percentage of white and African-American students has declined. KEYS data, which in this version have been collected since 2000, would be expected to be slightly behind this trend as they contain collections of data from 2000 to the present, while the *Digest* statistics represent arguably more current data from 2012.

Table 6: *Comparison of Racial/Ethnic Background of Students in the KEYS Data Set and the Digest of Education Statistics*

Racial/Ethnic Group	*Digest of Education Statistics*	KEYS Database
American Indian/Alaskan	1.1%	1.1%
Asian/Pacific	5.0%	3.0%
Black/African	16.0%	17.0%
Caucasian	52.4%	57.3%
Hispanic/Latino	23.1%	19.6%

The percentage of students eligible for free or reduced-price lunches, learning English as a second language, and qualifying for special education services was essentially the same in the KEYS database as national statistics recorded in the *Digest* (see Table 7).

Table 7: *Comparison of Distribution of the Percentage of Students Who Were Eligible for Free/Reduced Price Lunch, English as a Second Language Services, and Special Education Services in the KEYS Data Set and the Digest of Education Statistics*

Category	*Digest of Education Statistics*	KEYS Database
Percentage of students qualifying for free/reduced price lunch	48.1%	47.9%
Percentage of students receiving English as a second language services	9.8%	8.9%
Percentage of students receiving Special education services	13.0%	12.8%

With one exception, KEYS data on teacher characteristics also matched up quite well with national data in the *Digest*. KEYS teachers appear to have more doctorates than the national average, and more KEYS teachers reported holding an educational specialist certification or professional diploma as their highest degree attained (Table 8). However, in general, KEYS educational attainment data match quite closely to national statistics.

Table 8: *Comparison of Highest Degree Attained by Teachers Surveyed in the KEYS Data Set and National Statistics from the Digest of Education Statistics*

Highest Degree Attained	*Digest of Education Statistics*	KEYS Database
American Indian/Alaskan	1.1%	1.1%
Asian/Pacific	5.0%	3.0%
Caucasian	52.4%	57.3%
Hispanic/Latino	23.1%	19.6%

The teachers in the KEYS database also seemed to have somewhat more experience than teachers nationally (Table 9).

Table 9: *Comparison of Years' Experience of Teachers in the KEYS Data Set and National Statistics from the Digest of Education Statistics*

Highest Degree Attained	*Digest of Education Statistics*	KEYS Database
< 3 years'	13.4%	10.0%
3-9 years'	33.6%	28.1%
10-20 years'	29.3%	31.7%
20+ years'	23.7%	30.3%

The 2012 *Digest* only reported on teacher race/ethnicity and gender for high school teachers. Tables 10 and 11 compare KEYS data with *Digest* data on this measure. In Table 9, comparing race/ethnicity of teachers, KEYS data match up fairly well with national statistics. However Table 10 demonstrates differences in the gender makeup of these two sets of statistics with the KEYS database being overweighted towards male teachers relative to national data.

Table 10: *Comparison of Racial/Ethnic Background of High School Teachers in the KEYS Data Set and the Digest of Education Statistics*

Racial/Ethnic Background	*Digest of Education Statistics*	KEYS Database
American Indian/Alaskan	0.5%	1.1%
Asian/Pacific	1.5%	1.1%
Black/African	7.0%	7.1%
Caucasian	84.3%	84.2%
Hispanic/Latino	6.6%	4.9%

Table 11: *Comparison of Gender Makeup of High School Teachers in the KEYS Data Set and the Digest of Education Statistics*

Gender	*Digest of Education Statistics*	KEYS Database
Male	42.0%	48.7%
Female	58.0%	51.3%

Except for the gender of high school teachers, the KEYS database numbers match up very well with national statistics published in the Digest of Education Statistics published by the U.S. Department of Education. I conclude that, aside from gender differences, the KEYS database will provide me with data that is representative of teachers nationwide.

Research Question 1

How accurate are teachers' perceptions of their students' performance when compared to standardized test results?

A univariate analysis of teachers' perceptual match revealed that, on the average, teachers are quite accurate in their perceptions of students standardized test scores. The average perceptual match score was -0.04, which means that, on average, teachers perceptions of their students were within less than one percentile point of the students' actual standardized test performance. However, a high standard deviation of 1.55 suggests that the average does not tell the full story on teacher perceptions.

I used a series of one-way ANOVA analyses to examine teacher perceptual match by several variables, specifically, the variables that Ready and Wright use in their 2011 model. In addition to the variables Ready and Wright use, I added the independent variable gender. Because Ready and Wright look only at kindergarten teachers, they have very few male teachers in their sample and ignore gender as an independent variable. The following are the independent variables that I used in one-way ANOVA analysis:

1. gender of the teacher;
2. race of the teacher;
3. education level of the teacher;
4. certification of the teacher;
5. years' experience of the teacher;
6. socioeconomic status of the parents of the children served by the school;
7. type of community of the school; and,
8. size of the school.

The first independent variable I examined was gender. Table 12 displays the results:

Table 12: *Means and ANOVA Summary Table for Perceptual Match by Gender of the Teacher*

Gender	N	M	SD
Male	492	0.17	1.57
Female	1,453	-0.12	1.54

Source	Df	SS	MS	F	η^2
Perceptual Match	1	31.89	31.89	13.32***	< .01
Error	1,943	4,652.13	2.39		
Total	1,944	4,694.02			

Note. *** $p <$. 001

The difference between male and female teachers on perceptual match scores is statistically significant. Both male and female teachers are, on average, accurate in their

perceptions of students' test-taking ability. However, male teachers tend to overestimate the ability of their charges. Their mean perceptual match score was 0.17, just over 3 percentile points. Female teachers tend to underestimate their students with a perceptual match score of -0.12, or slightly more than 2 percentile points. While this difference is statistically significant, a partial η^2 came out to less than .01, so the difference between the genders in terms of perceptual match, or how close their perception of their target class matches the schools' standardized test score results for the highest grade, is trivial.

I next examined the effect of the race of the teacher. Because the assumption of homogeneity of variances was violated ($p < .001$), I used the Welch's F test, a more robust test of differences between means. Table 13 displays the results:

Table 13: *Means and Welch's F Test Summary Table for Perceptual Match by Racial/Ethnic Background of the Teacher*

Racial/Ethnic Background	N	M	SD
American Indian	20	-0.05	1.23
Asian/Pacific Islander	30	0.63	1.10
Black/African American	126	0.18	1.76
Caucasian	1,636	-0.10	1.52
Hispanic/Latino	93	0.33	1.87
Other	28	0.07	1.54

Source	df1	df2	SS
Perceptual Match	5	85.27	3.83

Note. $p = .004$

The difference among the six races of teachers on perceptual match scores is not statistically significant at levels determined by a Holm's sequential Bonferroni adjustment ($\alpha = .001$).

I took another look at race by comparing white teachers to teachers of all other races and ethnicities. The assumption of homogeneity of variances was violated (p = .010). Therefore, I used the Welch's F test. Table 14 displays the results:

Table 14: *Mean and Welch's F Test Summary Table for Perceptual Match by Race of the Teacher (Non-White/White)*

Racial/Ethnic Background	N	M	SD
Non-white	297	0.25	1.69
White	1,636	-0.10	1.52

Source	df1	df2	F
Perceptual Match	1	387.69	11.13

Note. *** $p = .001$

There is no statistically significant difference between white and non-white teachers on perceptual match scores. While the p value for the Welch's F Test was .001, the Holms' sequential Bonferroni adjustment value was $\alpha = .0009$.

I next examined the effect of the of the teacher's level of education. The assumption of homogeneity of variances was violated (p = .037). I used the Welch's F test. Table 15 displays the results:

Table 15: *Mean and Welch's F Test Summary Table for Perceptual Match by Education Level of the Teacher*

Education Level	N	M	SD
High School	20	.35	1.93
Associate's	16	-1.13	1.41
Bachelor's	886	-.06	1.51
Master's	931	-.04	1.57
Education specialist	81	-.01	1.74
Doctorate	29	-.10	1.52

Source	df1	df2	F
Perceptual Match	5	71.02	1.99

Note. $p = .090$

There is no statistical difference between the six different levels of education teachers reported on the survey.

The next independent variable I examined was teacher certification. Table 16 displays the results:

Table 16: *Means and One-Way ANOVA Summary Table for Perceptual Match by Level of Teacher Certification*

Level of Teacher Certification	N	M	SD
Certified	1,792	-0.04	1.54
Not certified but trained	115	-0.10	1.62
No formal training	45	-0.16	1.66
Other	14	-0.21	1.48

Source	df1	SS	MS	F
Perceptual Match	3	1.18	0.39	0.92
Error	1,962	4,685.73	2.39	
Total	1,965	4,686.91		

Note. $p = .920$

The difference between the different levels of teacher certification was not statistically significant.

Next up for examination was the experience of the teacher. KEYS collected data on the total years of experience of the teacher as well as the number of years' experience the teacher had in the school system and the building they worked in. Ready and Wright find that teachers with less than three years' experience were more likely to overestimate their students' abilities. Tables 17 - 19 display the results:

Table 17: *Means and One-Way ANOVA Summary Table for Perceptual Match by Total Years' Experience of the Teacher*

Total Years' Experience	N	M	SD
< 2 years'	109	0 .00	1.48
2-5 years'	324	-0.11	1.48
6-10 years'	404	-0.17	1.55
11-19 years'	525	-0.16	1.55
20+ years'	585	-0.09	1.59

Source	df	SS	MS	F
Perceptual Match	4	27.51	6.88	2.87
Error	1,942	4,653.83	2.40	
Total	1,946	4,681.34		

Note. $p = .022$

Table 18: *Means and One-Way ANOVA Summary Table for Perceptual Match by Years of Experience in the System of the Teacher*

Total Years' Experience	N	M	SD
< 2 years'	227	0.08	1.47
2-5 years'	446	-0.13	1.55
6-10 years'	380	-0.06	1.54
11-19 years'	433	-0.11	1.54
20+ years'	392	-0.04	1.56

Source	df	SS	MS	F
Perceptual Match	4	8.01	2.00	.84
Error	I,873	4,444.67	2.37	
Total	1,877	4,652.68		

Note. $p = .497$

Table 19: *Means and One-Way ANOVA Summary Table for Perceptual Match by Years of Experience in the Building of the Teacher*

Total Years' Experience	N	M	SD
< 2 years'	338	0.11	1.61
2-5 years'	578	-0.03	1.49
6-10 years'	435	-0.11	1.49
11-19 years'	357	-0.21	1.59
20+ years'	183	-0.06	1.55

Source	df	SS	MS	F
Perceptual Match	4	19.53	4.89	2.07
Error	I,880	4,445.93	2.37	
Total	1,884	4,465.49		

Note. $p = .083$

None of the three experience variables produced statistically significant differences between different number of years' experience.

I then recast the experience variables as binary variables to match Ready and Wright's data more closely, comparing new teachers with veteran teachers. I collapsed the top four

categories of each experience variable (2-5 years, 6-10 years, 11-19 years, and 20+ years) into a single category to create a variable that measured whether the teacher had less than two years' experience or more than two years' experience in each of the three categories (total experience, experience in the system, and experience in the building). Tables 20-22 display the results:

Table 20: *Means and One-Way ANOVA Summary Table for Perceptual Match by Total Years' Experience of the Teacher (< 2 Years', > 2 Years')*

Total Years' Experience	N	M	SD
< 2 years'	109	0.00	1.48
> 2 years'	1,838	-0.06	1.56

Source	df	SS	MS	F
Perceptual Match	1	0.34	0.34	0.14
Error	I,945	4,681.00	2.41	
Total	1,946	4,681.33		

Note. $p = .709$

Table 21: *Means and One-Way ANOVA Summary Table for Perceptual Match by Years of Experience in the System of the Teacher (< 2 Years', > 2 Years')*

Total Years' Experience	N	M	SD
< 2 years'	227	0.08	1.47
> 2 years'	1,651	-0.09	1.55

Source	df	SS	MS	F
Perceptual Match	1	5.49	5.49	2.32
Error	I,876	4,447.19	2.37	
Total	1,877	4,452.68		

Note. $p = .128$

Table 22: *Means and One-Way ANOVA Summary Table for Perceptual Match by Years of Experience in the Building of the Teacher (< 2 Years', > 2 Years')*

Total Years' Experience	N	M	SD
< 2 years'	338	0.11	1.61
> 2 years'	1,537	-0.10	1.52

Source	df	SS	MS	F
Perceptual Match	1	12.69	12.69	5.37
Error	I,883	4,452.79	2.37	
Total	1,884	4,465.49		

Note. $p = .021$

None of the three bifurcated experience variables produced statistically significant results.

The next independent variable I examined was socioeconomic status (SES) of the parents of the students served by the school. Ready and Wright find that the average SES of the school was a significant determinant of teacher bias in their perception of their students. They find that teachers in low-SES neighborhoods are more likely to underestimate their students than teachers working in upper-SES neighborhoods. I collapsed teachers in schools that served upper-income students into the same category as teachers in schools that served upper-middle-income students because of the small number of observations in the upper-income category. The assumption of homogeneity of variances was violated ($p < .001$). I used the Welch's F test. Table 23 displays the results:

Table 23: *Means and Welch's F Test Summary Table for Perceptual Match by Socioeconomic Status of the Parents of the Students the School Serves*

Socioeconomic Status	N	M	SD
Upper middle	97	-1.28	0.84
Middle	560	-0.51	1.44
Lower middle	626	-0.10	1.42
Lower	686	0.54	1.59

Source	df1	df2	F	ω^2
Perceptual Match	3	490.92	113.11	.15

Note. *** $p < .001$

I found a statistically significant relationship between SES and perceptual match. As the income of the parents increased, teachers were more likely to underestimate the cognitive abilities of their students as measured by standardized test scores. Conversely, teachers were more likely to overestimate the abilities of students with lower-income parents. In schools with upper-middle-income parents, teachers, on average, underestimated their students abilities by more than 25 percentile points (perceptual match = -1.28). In the schools with low-income parents, teachers overestimated their students by more than 10 percentile points on average. The Games-Howell Post Hoc test found that there was a statistically significant difference between all four levels of SES (upper-middle income, middle income, lower-middle income, and low income). There is a large effect size as well (partial $\omega^2 = .15$).

I next examined the type of community the school was located in (urban, suburban, small city, town, rural). Table 24 displays the results:

Table 24: *Means and One-Way ANOVA Summary Table for Perceptual Match by Type of Community the School Is Located In*

Type of Community	N	M	SD
Urban	427	0.67	1.64
Suburban	451	0.04	1.42
Small city	504	-0.32	1.48
Town	300	-0.59	1.41
Rural	311	-0.18	1.48

Source	df1	SS	MS	F	η^2
Perceptual Match	4	355.92	88.87	39.97	.07
Error	1,988	4,420.90	2.22		
Total	1,992	4,776.29			

Note. *** $p < .001$

The perceptual match score was significantly different between the different types of communities where schools were located. There was a medium effect size (partial η^2 = .07). As Ready and Wright found in their study, urban communities were significantly different than other communities. Teachers in urban communities had an average perceptual match score of 0.67, meaning that on average they overestimated their students' abilities by more than 13 percentile points. Ready and Wright found that teachers in urban areas underestimated their students. Teachers in towns tend to be the toughest in their perceptions of their students, averaging a perceptual match score of -0.59, underestimating their students' scores by nearly 12 percentile points. Teachers in urban areas were significantly different on their perceptual match scores from all communities.

I next examined the size of the school as an independent variable. I examined school size as a categorical variable, dividing schools into three groups, small (less than 350 students), mid-size (350-749 students), and large (750 students or more). Table 25 displays the results:

Table 25: *Means and One-Way ANOVA Summary Table for Perceptual Match by Size of the School*

Size of School	N	M	SD
Small	354	-0.06	1.62
Mid-size	1,156	-0.10	1.54
Large	483	0.11	1.50

Source	df	SS	MS	F
Perceptual Match	2	14.56	7.28	3.04
Error	1,990	4,761.73	2.39	
Total	1,992	4,776.29		

Note. $p = .048$

The perceptual match score was not statistically significant different between the different sizes of schools.

To sum up my one-way ANOVAs for all students: several personal characteristics of teachers had a statistically significant differences across the characteristic for perceptual match, but none of these had a significant effect size. Both the SES and the type of community the school was located in had significantly different distributions of perceptual match. Teachers in schools with low-income students and schools in urban communities were more likely to have higher perceptual match scores, signifying that they were more likely to overestimate students' abilities relative to teachers in schools with high-middle-income students and schools in less urban communities.

Research Question 2

Do teachers' perceptions of minority students differ from their perceptions of white students?

A univariate analysis revealed that, on the average, teachers are fairly accurate in their perceptions of minority students' standardized test scores. The average perceptual match

score for minority students was 0.24. The average perceptual match score for white students was -0.28. The difference between the two scores is statistically significant and has a small effect size ($z = 8.77$; Cohen's $d = 0.31$). Teachers, on average, are more likely to overestimate their minority students' cognitive abilities while underestimating white students' cognitive abilities.

However, high standard deviations of 1.67 and 1.63, respectively, for minority and for white students suggests that the average does not tell the full story on teacher perceptions of their minority and white students. I used a series of one-way ANOVA analyses to examine teachers' perceptual match for their minority students (perceptual match minority) and for white students (perceptual match white) by the same independent variables I used in my examination of all students, above.

The first independent variable I examined for teachers' perceptual match for minority and white students was gender. Table 26 displays the results for minority students, Table 27 for white students:

Table 26: *Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Gender of the Teacher*

Gender	N	M	SD		
Male	385	0.52	1.68		
Female	1,089	0.15	1.66		
Source	**df**	**SS**	**MS**	**F**	**η^2**
Perceptual Match	1	39.25	39.25	14.19	.01
Error	1,472	4,073.36	2.77		
Total	1,473	4,112.62			

Note. *** p <. 001

The difference between male and female teachers on perceptual match minority scores is statistically significant. Female teachers are, on average, accurate in their perceptions of

their students' test-taking ability. However, male teachers tend to overestimate the ability of their charges by over 10 percentile points. Their mean perceptual match minority score was 0.52. Female teachers also tend to overestimate their minority students with a perceptual match minority score of 0.15. But this only represents 3 percentile points. There is a small effect size (partial $\eta^2 = .01$).

Table 27: *Means and One-Way ANOVA Summary Table for Perceptual Match White by Gender of the Teacher*

Gender	N	M	SD
Male	413	.06	1.61
Female	1,171	.36	1.64

Source	df	SS	MS	F
Perceptual Match	1	28.66	28.66	10.82
Error	1,582	4,192.20	2.65	
Total	1,583	4,220.86		

Note. $p =$. 001

The difference between male and female teachers on perceptual match white scores is not statistically significant. While the p for the ANOVA was .001, the Holm's sequential Bonferroni adjustment value was $\alpha = .0009$.

Both male and female teachers tend to overestimate minority students while underestimating white students. Male teachers are quite accurate in estimating white students cognitive abilities (perceptual match white = -0.06) but tend to overestimate the abilities of minority students (perceptual match minority =0.52). The difference between the two means is significant and has a small effect size ($z = 4.95$; Cohen's $d = .35$). Female teachers tend to more accurately perceive the abilities of minority students (perceptual match minority = 0.15) while underestimating white students (perceptual match white = -0.36) There is a statistically significant difference between female teacher perceptions of

minority students and female teacher perceptions of white students that has a small effect size ($z = 7.40$; Cohen's $d = .31$).

Thus, gender is not a factor in how teachers perceive their students. Both male and female teachers tend to perceive minority students differently than they perceive white students. Male and female teachers tend to overestimate minority cognitive ability and underestimate the ability of white students.

I next examined the effect of the race of the teacher on perceptual match for minority and white students. Tables 28 and 29 display the results for minority students and for white students:

Table 28: *Means and One-Way ANOVA Summary Tables for Perceptual Match Minority by Racial/Ethnic Background of the Teacher*

Racial/Ethnic Background	N	M	SD
American Indian	17	0.53	1.62
Asian/Pacific Islander	29	0.31	1.69
Black/African American	93	0.41	1.68
Caucasian	1,218	0.17	1.65
Hispanic/Latino	85	0.65	1.84
Other	20	0.30	2.00

Source	df	SS	MS	F
Perceptual Match	5	23.75	4.75	1.71
Error	1,456	4,403.01	2.78	
Total	1,461	4,066.75		

Note. $p = .129$

The difference among the six races of teachers on perceptual match minority scores is not statistically significant.

I found one extreme outlier in the perceptual match white data for Asian-American teachers. However, since there were only 17 Asian-American teachers with valid data, I

left this outlier in the analysis. Because the assumption of homogeneity of variances was violated (p = .008), I used the Welch's F test for perceptual match white.

Table 29: *Means and Welch's F Test Summary Table for Perceptual Match White by Racial/Ethnic Background of the Teacher*

Racial/Ethnic Background	N	M	SD
American Indian	17	-0.18	1.51
Asian/Pacific Islander	30	-0.10	1.45
Black/African American	90	-0.06	1.96
Caucasian	1,340	-0.31	1.60
Hispanic/Latino	77	-0.22	1.77
Other	21	-0.86	1.68

Source	df	df2	F
Perceptual Match	5	70.68	0.88

Note. p = .499

There is no statistically significant difference among the six races of teachers on perceptual match white scores.

I took another look at race by comparing white teachers to all other teachers for minority students and white students. Tables 30 and 31 display the results:

Table 30: *Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Racial/Ethnic Background of the Teacher (Non-White/White)*

Racial/Ethnic Background	N	M	SD
Non-white	244	0.48	1.75
White	1,218	0.17	1.65

Source	df	SS	MS	F
Perceptual Match	1	19.38	19.38	6.99
Error	1,460	4,047.38	2.77	
Total	1,461	4,066.75		

Note. p =. 008

There is no statistically significant difference between white and non-white teachers on perceptual match minority scores (the Holms' sequential Bonferroni adjustment value was $\alpha = .001$).

Table 31: *Means and One-Way ANOVA Summary Table for Perceptual Match White by Racial/Ethnic Background of the Teacher (Non-White/White)*

Racial/Ethnic Background	N	M	SD
Non-white	235	0.20	1.78
White	1,340	0.31	1.60

Source	df	SS	MS	F
Perceptual Match	1	2.53	2.53	.95
Error	1,573	4,188.71	2.66	
Total	1,340	4,191.23		

Note. p =. 330

There is no statistically significant difference between white and non-white teachers on perceptual match white scores.

There is a statistically significant difference in how non-white and white teachers perceive minority students and how they perceive white students. Both non-white and white teachers tend to overestimate the cognitive abilities of minority students while underestimating the ability of white students. Non-white teachers had an average perceptual match minority score of 0.48 and an average perceptual match score white of -0.20. The difference is statistically significant with a small effect size ($z = 4.18$; Cohen's $d = .39$). Similarly, white teachers overestimate the ability of minority students while underestimating the ability of white students (perceptual match minority = 0.17; perceptual match white = -0.31). This difference is also statistically significant with a small effect size ($z = 7.44$; Cohen's $d = .29$).

Thus, the teacher's race is not a factor in how teachers perceive their students. Both non-white and white teachers tend to perceive minority students differently than they perceive white students. Non-white and white teachers tend to overestimate minority cognitive ability and underestimate the ability of white students.

I next examined the effect of the of the teacher's level of education on perceptual match minority and perceptual match white. The assumption of homogeneity of variances was violated (p = .007) for perceptual match minority but not for perceptual match white. I used the Welch's F test to examine perceptual match minority and a one-way ANOVA to examine perceptual match white for any effects of teacher's education. Tables 32 and 33 display the results:

Table 32: *Means and Welch's F Test Summary Table for Perceptual Match Minority by Education Level of the Teacher*

Level of Education	N	M	SD
High School	21	0.86	1.98
Associate's	11	-0.82	0.98
Bachelor's	684	0.20	1.66
Master's	679	0.30	1.63
Education Specialist	69	-0.16	2.02
Doctorate	22	0.50	1.68

Source	df	df2	F
Perceptual Match	5	56.68	3.67

Note. p = .006

The Welch's F test was not statistically significant.

Table 33: *Means and One-Way ANOVA Summary Table for Perceptual Match White by Education Level of the Teacher*

Level of Education	N	M	SD
High School	16	-0.31	1.66
Associate's	10	-2.10	1.73
Bachelor's	715	-0.30	1.61
Master's	756	-0.26	1.62
Education Specialist	74	-0.26	1.62
Doctorate	26	-0.38	1.94

Source	df	SS	MS	F
Perceptual Match	5	36.49	7.30	2.74
Error	1,591	4,238.9	2.66	
Total	1,596	4,275.08		

Note. $p =$. 018

There is no statistical difference between the six different levels of education teachers reported on the survey.

Teachers with bachelor's and master's degrees had significantly different perceptual match scores for minority students and white students. Teachers with a bachelor's degree had an average perceptual match score of 0.20 for minority students and an average perceptual match score of -0.30 for white students. This difference is statistically significant with a small effect size ($z = 5.79$; Cohen's $d = .31$). Teacher's with a master's degree had an average perceptual match score of 0.30 for minority students and -0.26 for white students, a statistically significant difference with a small effect size ($z = 6.55$; Cohen's $d = .35$). There was no statistically significant difference for teachers with a doctorate in their perceptual match scores for minority students and white students.

Thus, the level of a teacher's education is not a factor in how teachers perceive their students. Teachers with a bachelor's degree and teachers with a master's degree tend to perceive minority students differently than they perceive white students. However, both

groups tend to overestimate minority cognitive ability and underestimate the ability of white students.

The next independent variable I examined was teacher certification. Tables 34 and 35 display the results:

Table 34: *Means and One-Way ANOVA Summary Tables for Perceptual Match Minority by Level of Teacher Certification*

Level of Teacher Certification	N	M	SD
Certified	1,349	0.24	1.66
Not certified by trained	91	0.31	1.69
No formal training	36	-0.06	1.74
Other	9	0.11	2.42

Source	df	SS	MS	F
Perceptual Match	3	3.73	1.24	0.45
Error	1,481	4,234.41	2.79	
Total	1,484	4,138.14		

Note. $p = .721$

Table 35: *Means and One-Way ANOVA Summary Tables for Perceptual Match White by Level of Teacher Certification*

Level of Teacher Certification	N	M	SD
Certified	1,461	-0.26	1.62
Not certified by trained	93	-0.57	1.6
No formal training	36	-0.44	1.63
Other	9	0.22	1.99

Source	df	SS	MS	F
Perceptual Match	3	9.17	3.06	1.16
Error	1,595	4,223.79	2.64	
Total	1,598	4,214.79		

Note. $p = .325$

The difference between the different levels of teacher certification was statistically insignificant for perceptual match minority and perceptual match white.

Certified teachers have significantly different perceptual match scores for minority students and for white students. Certified teachers had an average perceptual match score of 0.24 for minority students and an average perceptual match score of -0.26 for white students. This difference is statistically significant with a small effect size (z = 8.16; Cohen's d = .31).

Teacher certification is not a factor in how teachers perceive their students. Certified teachers tend to perceive minority students differently than they perceive white students. Both certified and non-certified teachers tend to overestimate minority cognitive ability and underestimate the ability of white students.

Next up for examination was the experience of the teacher. Ready and Wright find that teachers with less than three years' experience were more likely to overestimate their students' abilities. Tables 36 - 38 display the results for perceptual match minority:

Table 36: *Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Total Years' Experience of the Teacher*

Total Years' Experience	N	M	SD
< 2 years'	88	0.16	1.73
2-5 years'	258	0.22	1.54
6-10 years'	312	0.28	1.67
11-19 years'	399	0.10	1.71
20+ years'	420	0.36	1.69

Source	df	SS	MS	F
Perceptual Match	4	15.81	3.95	2.79
Error	1,472	4,101.25	2.79	
Total	1,476	4,117.06		

Note. p = .225

Table 37: *Means and One-Way ANOVA Summary for Perceptual Match Minority by Years of Experience in the System of the Teacher*

Total Years' Experience	N	M	SD
< 2 years'	173	0.12	1.56
2-5 years'	357	0.20	1.69
6-10 years'	302	0.16	1.60
11-19 years'	317	0.12	1.76
20+ years'	272	0.47	1.66

Source	df	SS	MS	F
Perceptual Match	4	23.47	2.07	0.75
Error	1,416	3,928.06	2.78	
Total	1,420	3,951.54		

Note. $p = .077$

Table 38: *Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Years of Experience in the Building of the Teacher*

Total Years' Experience	N	M	SD
< 2 years'	256	0.20	1.69
2-5 years'	453	0.22	1.62
6-10 years'	338	0.19	1.64
11-19 years'	258	0.15	1.81
20+ years'	125	0.45	1.58

Source	df	SS	MS	F
Perceptual Match	4	8.30	2.07	0.75
Error	1,425	3,966.80	2.78	
Total	1,429	3,975.09		

Note. $p = .531$

None of the three experience variables produced statistically significant results.

I then recast the experience variables as binary variables to match Ready and Wright's data more closely. I collapsed the top four categories of each experience variable into a single category to create a variable that measured whether the teacher had less than two years' experience or more than two years' experience in each of the three categories (total

experience, experience in the system, and experience in the building). Tables 39 - 41 display the results for perceptual match minority and Tables 42 – 44 for perceptual match white (note that perceptual match white for system experience failed the homogeneity of variances test, so I used Welch's F for that variable):

Table 39: *Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Total Years' Experience of the Teacher (<2 Years', > 2 Years')*

Total Years' Experience	N	M	SD
< 2 years'	88	0.16	1.73
> 2 years'	1,389	0.24	1.67

Source	df	SS	MS	F
Perceptual Match	1	0.57	0.57	0.20
Error	1,475	4,116.49	2.79	
Total	1,476	4,117.06		

Note. $p = .652$

Table 40: *Means and One-Way ANOVA Summary for Perceptual Match Minority by Years of Experience in the System of the Teacher (< 2 Years', > 2 Years')*

Total Years' Experience	N	M	SD
< 2 years'	173	0.12	1.56
> 2 years'	1,248	0.23	1.68

Source	df	SS	MS	F
Perceptual Match	1	1.93	1.93	0.69
Error	1,419	3,949.60	2.78	
Total	1,420	3,951.54		

Note. $p = .405$

Table 41: *Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Years of Experience in the Building of the Teacher (< 2 Years', > 2 Years')*

Total Years' Experience	N	M	SD
< 2 years'	256	0.20	1.69
> 2 years'	1,174	0.22	1.66

Source	df	SS	MS	F
Perceptual Match	1	0.12	0.12	0.04
Error	1,428	3,974.98	2.78	
Total	1,429	3,975.09		

Note. $p = .838$

Table 42: *Means and One-Way ANOVA Summary Table for Perceptual Match White by Total Years' Experience of the Teacher (< 2 Years', > 2 Years')*

Total Years' Experience	N	M	SD
< 2 years'	93	-0.27	1.62
> 2 years'	1,497	-0.29	1.64

Source	df	SS	MS	F
Perceptual Match	1	0.05	0.05	0.02
Error	1,588	4,264.54	2.69	
Total	1,589	4,265.59		

Note. $p = .889$

Table 43: *Means and One-Way ANOVA Summary for Perceptual Match White by Years of Experience in the System of the Teacher (< 2 Years', > 2 Years')*

Total Years' Experience	N	M	SD
< 2 years'	183	0.14	1.58
> 2 years'	1,346	0.32	1.64

Source	df1	df2	F
Perceptual Match	1	238.83	1.97

Note. $p = .162$

Table 44: *Means and One-Way ANOVA Summary Table for Perceptual Match White by Years of Experience in the Building of the Teacher (< 2 Years', > 2 Years')*

Total Years' Experience	N	M	SD
< 2 years'	272	-0.20	1.69
> 2 years'	1,266	-0.32	1.62

Source	df	SS	MS	F
Perceptual Match	1	3.17	3.17	1.19
Error	1,536	4,101.63	2.67	
Total	1,537	4,104.80		

Note. p = .276

None of the six bifurcated experience variables (three each for perceptual match minority and perceptual match white) had statistically significant results.

Teachers with greater than two years' experience have significantly different perceptual match scores for minority students and for white students. There is no significant difference in how teachers with less than two years total experience or system experience view the cognitive abilities of minority students versus how they view the cognitive abilities of white students. However, teachers with less than two years' experience in their building, like more experienced teachers, are more likely to overestimate minority students' ability while underestimating white students' ability. Table 45 displays the average perceptual match scores for teachers by level of experience.

Table 45: *Average Perceptual Match Score Minority and Perceptual Match Score White for Teachers by Amount of Teaching Experience (< 2 Years', > 2 Years')*

Variable	Description	Perceptual Match Minority	Perceptual Match White	*Z*	Cohen's *d*
Total Exp	< 2 Years'	0.16	-0.27	1.70	–
	> 2 Years'	0.24	-0.29	8.68	.32
System Exp	< 2 Years'	0.12	-0.14	1.55	–
	> 2 Years'	0.23	-0.32	8.34	.33
Building Exp	< 2 Years'	0.20	-0.20	2.68	.23
	> 2 Years'	0.22	-0.32	8.06	.33

Teacher's level of building experience is not a factor in how teachers perceive their students. Both new and experienced teachers, as defined by tenure in the school building, tend to overestimate the cognitive abilities of minority students and underestimate the abilities of white students. There is no statistically significant difference in how teachers who are new to the school system or new to the profession perceive minority students and white students. Thus, years of experience is not a significant determinant of the accuracy of teachers' perceptions of their students.

The next independent variable I examined was the socioeconomic status (SES) of the parents of the students served by the school. Ready and Wright found that the average SES of the parents of the students served by the school was a significant determinant of teacher bias in their perception of their students. They found that teachers in schools that served low-income students were more likely to underestimate their students than teachers working in schools that served upper-income students. I collapsed schools serving upper-income students into the upper-middle-income category because of the small number of observations of teachers in schools serving upper-income students. Tables 46 and 47 display the results for perceptual match minority and perceptual match white:

Table 46: *Means and One-Way ANOVA Summary Table for Perceptual Match Minority by the Socioeconomic Status of the Parents of the Students the School Serves*

Socioeconomic Status	N	M	SD
Upper middle	60	-0.18	1.6
Middle	430	-0.17	1.71
Lower middle	499	0.26	1.53
Lower	496	0.59	1.70

Source	df	SS	MS	F	ω^2
Perceptual Match	3	145.71	48.57	18.00	.04
Error	1,481	3,995.90	2.70		
Total	1,484	4,141.61			

*Note.**** p < .001

Table 47: *Means and Welch's F Test Summary Table for Perceptual Match White by Socioeconomic Status of the Parents of the Students the School Serves*

Socioeconomic Status	N	M	SD
Upper middle	73	-1.21	0.82
Middle	479	-0.63	1.41
Lower middle	524	-0.26	1.52
Lower	525	-0.16	1.88

Source	df1	df2	F	ω^2
Perceptual Match	3	383.21	44.64***	.08

Note. *** $p < .001$

I found a statistically significant relationship between the SES of the parents of the students served by the school and perceptual match minority. As the income of the neighborhood increases, teachers are more likely to underestimate the cognitive abilities of their students as measured by standardized test scores. Conversely, teachers are more likely to overestimate the abilities of students in schools serving lower-income students. In schools serving upper-middle-income students, teachers, on average, underestimated their students abilities by more than 4 percentile points (perceptual match minority = -0.18). In

the schools serving the lowest-income students, teachers overestimated their minority students by more than 12 percentile points on average. The Tukey HSD Post Hoc test found that there was a statistically significant difference for perceptual match minority schools with low-income parents and schools with every other SES level of parents. It also found that schools with middle-income parents were different than schools with lower-middle-income parents. There was a small effect size (partial $\eta^2 = .04$).

I found a statistically significant relationship between SES and perceptual match white. As the SES of the parents of the students served by the school increases, the amount that teachers underestimate their white students increases. Conversely, teachers are more likely to overestimate the abilities of white students in schools serving lower-income students. In schools serving upper-middle-income students, teachers, on average, had a perceptual match score of -1.28 for their white students. In schools serving low-income students, teachers overestimated their white students with and average perceptual match score of 0.16. The Games-Howell Post Hoc test found that there was a statistically significant difference between all four levels of SES (upper-middle income, middle income, lower-middle income, and low income). There was a medium effect size (partial $\omega^2 = .08$).

Teachers consistently rate minority students higher vis-a-vis their standardized test scores than they rate white students at all SES levels (See Table 48). Teachers in schools serving upper-middle-income and middle-income students tend to underestimate their white students. They also tend to underestimate their minority students, just not as much. In schools serving lower-middle-income students, teachers tend to underestimate white

students and overestimate minority students. In schools serving low-income students, teachers overestimate both minority and white students.

Table 48: *Average Perceptual Match Score Minority and Perceptual Match Score White for Teachers by the Socioeconomic Status of the Parents of the Students the School Serves*

SES	Perceptual Match Minority	Perceptual Match White	*Z*	Cohen's *d*
Upper Middle	-0.18	-1.21	4.50	.83
Middle	-0.17	-0.63	4.39	.30
Lower Middle	0.27	-0.26	5.53	.35
Lower	0.59	0.16	3.80	.24

I next examined the type of community the school was located in. Tables 49 and 50 display the results for perceptual match minority and perceptual match white:

Table 49: *Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Type of Community the School Is Located In*

Type of Community	N	M	SD
Urban	334	.90	1.61
Suburban	317	.31	1.69
Small city	385	.13	1.64
Town	256	-.27	1.57
Rural	216	-.06	1.58

Source	df	SS	MS	F	η^2
Perceptual Match	4	237.41	59.35	22.53***	.06
Error	1,503	3,959.24	2.63		
Total	1,507	4,196.66			

Note. *** $p < .001$

Table 50: *Means and Welch's F Test Summary Table for Perceptual Match White by Type of Community the School Is Located In*

Type of Community	N	M	SD
Urban	328	0.34	1.90
Suburban	356	-0.05	1.51
Small city	429	-0.71	1.50
Town	265	-0.64	1.55
Rural	246	-0.26	1.41

Source	df	SS	F	ω^2
Perceptual Match	4	761.49	22.86***	.05

Note. *** $p < .001$

The perceptual match minority score was significantly different between the different types of communities that schools were located in. There was a small effect size (partial $\eta^2 = .06$). As Ready and Wright found in their study, urban communities were significantly different than all other types of communities. Teachers in urban communities had an average perceptual match minority score of 0.89, meaning that they overestimated their students' abilities by more than 18 percentile points. Ready and Wright found that teachers in urban schools underestimated their students. Teachers in towns tend to be the toughest in their perceptions of their students, averaging a perceptual match score of -0.27, underestimating their students' scores by more than 5 percentile points.

The perceptual match white score was significantly different between the different types of communities that schools were located in. There was a small effect size (partial $\omega^2 = .05$). As Ready and Wright found in their study, urban communities were significantly different than other communities. Teachers in urban communities had an average perceptual match white score of 0.34. Ready and Wright found that teachers in

urban areas underestimated their students. Teachers in small cities tend to underestimate white students the most, averaging a perceptual match white score of -0.71.

Teachers consistently rate minority students higher vis-a-vis their standardized test scores than they rate white students in all types of communities except rural communities (see Table 51).

Table 51: *Average Perceptual Match Score Minority and Perceptual Match Score White for Teachers by the Type of Community the School Is Located In*

SES	Perceptual Match Minority	Perceptual Match White	*Z*	Cohen's *d*
Urban	0.90	0.34	4.02	.31
Suburban	0.31	-0.05	2.92	.23
Small city	0.13	-0.71	7.62	.54
Town	-0.27	-0.64	2.72	.24
Rural	-0.07	-0.16	1.39	–

I next examined the size of the school as an independent variable. I transformed school size into a categorical variable (small schools, less than 350 students; medium-size schools, 350-749 students; and large schools, 750 students or more). Tables 52 and 53 display the results for minority students and for white students:

Table 52: *Means and One-Way ANOVA Summary Table for Perceptual Match Minority by Size of the School*

Size of School	N	M	SD
Small	270	0.33	1.69
Mid-size	888	0.14	1.66
Large	350	0.43	1.65

Source	df	SS	MS	F
Perceptual Match	2	22.59	11.29	4.07
Error	1,505	4,174.07	2.77	
Total	1,507	4,196.66		

Note. $p = .001$

Table 53: *Means and One-Way ANOVA Summary Table for Perceptual Match White by Size of the School*

Size of School	N	M	SD
Small	295	-0.21	1.66
Mid-size	941	-0.37	1.66
Large	388	-0.09	1.65

Source	df	SS	MS	F
Perceptual Match	2	22.97	11.48	4.32
Error	1,621	4,313.45	2.66	
Total	1,623	4,336.41		

Note. $p = .014$

There was no statistically significant difference in perceptual match minority and perceptual match white scores over different size schools. In other words, size does not matter.

Teachers rate minority students higher vis-a-vis their standardized test scores than they rate white students in all size schools (see Table 54). Teachers tend to underestimate white students' cognitive abilities and overestimate minorities' cognitive abilities, regardless of the size of the school.

Table 54: *Average Perceptual Match Score Minority and Perceptual Match Score White for Teachers by the Size of the School*

Size of School	Perceptual Match Minority	Perceptual Match White	*Z*	Cohen's *d*
Small	.33	-.21	3.82	.32
Medium	.14	-.37	6.71	.31
Lower Middle	.43	-.09	4.27	.31

To sum up my one-way ANOVAs for minority students and white students: of the personal teacher characteristics, only gender had a statistically significant difference for perceptual match minority. Male teachers overestimated minority students by 10

percentile points while female teachers overestimated them by only 3 percentile points. None of the other independent variables that represented the personal characteristics of teachers produced statistically significant results.

Both the SES of the school's parents and the type of community the school was located in had significantly different distributions of perceptual match minority and perceptual match white. Schools that served low-income students and schools in urban communities were more likely to have higher perceptual match scores, signifying that they were more likely to overestimate students' abilities relative to schools serving higher-income students and less-urban communities.

On almost all variables, there were statistically significant differences between the perceptual match score for all students and the perceptual match score for minority students. In every case, teachers had higher perceptions of minority students relative to their standardized test scores than they did of white students. Table 55 summarizes these results.

Table 55: *Summary Table of Perceptual Match Minority Compared to Perpetual Match White for the Values of the Independent Variables*

Variable	Description	Perceptual Match Minority	Perceptual Match White	*Z*	Cohen's *d*
Overall		0.24	-0.28	8.77	.31
Gender	Male	0.52	-0.06	4.95	.35
	Female	0.15	-0.36	7.40	.31
Race	Non-white	0.48	-0.20	4.18	.39
	White	0.17	-0.31	7.44	.29
Education	Bachelor's	0.20	-0.30	5.79	.31
	Master's	0.30	-0.26	6.55	.35
Certified		0.24	-0.26	8.16	.31
Total exp	> 2 years'	0.24	-0.29	8.68	.32
System exp	> 2 years'	0.23	-0.32	8.34	.33
Bldg exp	< 2 Years'	0.20	-0.20	2.68	.23
	> 2 Years'	0.22	-0.32	8.06	.33
SES	Upper middle	-0.18	-1.21	4.50	.83
	Middle	-0.17	-0.63	4.39	.30
	Lower middle	0.27	-0.26	5.53	.35
	Lower	0.59	0.16	3.80	.24
Community	Urban	0.90	0.34	4.02	.31
	Suburban	0.31	-0.05	2.92	.23
	Small City	0.13	-0.71	7.62	.54
	Town	-0.27	-0.64	2.72	.24
School Size	Small	0.33	-0.21	3.82	.32
	Medium	0.14	-0.37	6.71	.31
	Large	0.43	-0.09	4.27	.31

Research Question 3

How do teachers' perceptions vary by grade?

KEYS allowed me to examine teachers in all grades. I set out to run a one-way ANOVA of the distribution of perceptual match over three types of schools (elementary, middle school, high school). Because the assumption of homogeneity of variances was violated ($p < .001$), I used the Welch's F test. Table 56 displays the results:

Table 56: *Means and Welch's F Test Summary Table for Perceptual Match by Type of School*

Type of School	N	M	SD
Elementary	1,158	-0.16	1.53
Middle school	520	-0.06	1.45
High School	209	0.34	1.70

Source	df1`	df2	F	ω^2
Perceptual Match	2	527.38	8.19***	.01

Note. *** $p < .001$

The perceptual match score was significantly different between the types of schools. There was a small effect size (partial $\omega^2 = .01$). Teachers in elementary schools had an average perceptual match minority score of -0.16, meaning that on average they underestimated their students' abilities by more than 3 percentile points. Middle school teachers were even more accurate, underestimating their students by only 1 percentile point (perceptual match = -0.06). High school teachers overestimated their students by almost 7 percentile points on average (perceptual match = 0.34).

I next examined perceptual match minority and perceptual match white over school type. The assumption of homogeneity of variances was violated ($p < .001$) for perceptual match minority, so I used the Welch's F test to examine that distribution. Tables 57 and 58 display the results:

Table 57: *Means and Welch's F Test Summary Table for Perceptual Match Minority by Type of School*

Type of School	N	M	SD
Elementary	851	0.04	1.71
Middle school	395	0.28	1.43
High School	184	0.79	1.72

Source	df1`	df2	F	ω^2
Perceptual Match	2	475.16	15.13***	.02

Note. *** $p < .001$

Table 58: *Means and One-Way ANOVA Summary Table for Perceptual Match White by Type of School*

Type of School	N	M	SD
Elementary	931	-0.38	1.67
Middle school	429	-0.28	1.55
High school	185	0.21	1.67

Source	df	SS	MS	F	η^2
Perceptual Match	2	57.74	27.37	10.20***	.01
Error	1,542	4,136.66	2.68		
Total	1,544	4,191.40			

Note. *** $p < .001$

The perceptual match minority score was significantly different between the types of schools. There was a small effect size (partial $\omega^2 = .02$). Teachers in elementary schools had an average perceptual match minority score of 0.04, meaning that on average they overestimated their students' abilities by less than 1 percentile point. Middle school overestimated their minority students by 7 percentile points (perceptual match minority = 0.28) and high school teachers overestimated these same students by more than 16 percentile points on average (perceptual match minority = 0.79).

Similarly, the perceptual match white score was significantly different between the types of schools. There was a small effect size (partial $\eta^2 = .02$). Teachers in elementary schools had an average perceptual match white score of -0.38, underestimating their students by nearly 8 percentile points. Middle school also underestimated their white students by more than 5 percentile points with a perceptual match white score of -0.28. High school teachers overestimated their white students with an average perceptual match score of 0.21.

Teachers at all school types consistently (elementary school, $z = 5.31$; middle school, $z = 5.38$; high school, $z = 3.27$; all with small effect sizes, Cohen's d of, respectively, .25, .37, and .34) rate minority students higher vis-a-vis their standardized test scores than they rate white students. While high school teachers tend to overestimate all of their students they overestimate minority students significantly more than white students. Elementary and middle school teachers tend to overestimate their minority students and underestimate their white students.

I ran a two-way ANOVA to examine the interaction between teacher gender and type of school. I hypothesized that the significance of type of school may be due to gender because of the much greater percentage of female teachers in elementary schools than in high schools and because I had found that gender made a difference in perceptual match minority scores, with male teachers overestimating minority students more than female teachers. The distribution of perceptual match over teacher gender and school type was not normal. Examining the skewness and kurtosis of each distribution, I found that perceptual match for female elementary school teachers was negatively kurtosed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match was approximately

normally distributed. I found no extreme outliers in the data. The assumption of homogeneity of variances was violated ($p < .001$). Unfortunately, unlike with one-way ANOVA, there is no way to resolve lack of homogeneity of variances other than data transformation. I attempted reflect and square root, reflect and logarithmic, and reflect and inverse transformations, but none had an effect on the Levene test. Therefore, two-way ANOVA was closed to me for examining teacher gender interaction with school type. Nonetheless, with a simple visual analysis of the means for perceptual match by teacher gender and school type, I found evidence that the significance of school type was not due to gender. The difference between the perceptual match score for male and female high school teachers was not statistically significant ($z = 1.38$).

I attempted the same analysis with perceptual match minority and perceptual match white. The distribution of perceptual match minority over the gender and school type was not normal. Examining the skewness and kurtosis of each distribution, I found that perceptual match minority for female elementary school teachers was negatively kurtosed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match minority was approximately normally distributed. I found no extreme outliers in the data. The assumption of homogeneity of variances, as assessed by Levene's test for equality of variances, was violated ($p = .001$). I attempted reflect and square root, reflect and logarithmic, and reflect and inverse transformations, but none had an effect on the Levene test. Therefore, two-way ANOVA was closed to me for examining teacher gender interaction with school type. Nonetheless, with a simple visual analysis of the means for perceptual match minority by teacher gender and school type, I confirmed what I had found with the perceptual match data for all students: the type of school effect is not due to

gender of the teacher. The difference between the perceptual match minority score for male and female high school teachers was not statistically significant.

I attempted the same analysis with perceptual match white. The distribution of perceptual match minority over the gender and school type was not normal except for male high school teachers. Examining the skewness and kurtosis of each distribution, I found that perceptual match minority for female elementary school teachers was negatively kurtosed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match minority was approximately normally distributed. I found no extreme outliers in the data. There was homogeneity of variances, as assessed by Levene's test for equality of variances (p = .173). Table 59 displays the results:

Table 59: *Two-Way ANOVA Summary Table for Perceptual Match White by Gender and Type of School*

Source	df	SS	MS	F	*p*	η^2
Gender	1	9.05	9.05	3.41	.065	<.01
School type	2	28.59	14.30	5.39	.005	<.01
Gender*type	1	11.55	5.78	2.18	.114	<.01
Error	1,912	3,988.50	2.65			
Total	1,916	4,206.00				

This two-way ANOVA for perceptual match white confirms what I had seen with visual inspection of the distribution of perceptual match and perceptual match minority: gender is not a factor in the significance of school type. Only school type is statistically significant in this two-way ANOVA and neither of the variables has a significant effect size.

In summary, the type of school a teacher teaches in matters in his/her perception of students. Elementary and middle school teachers tend to underestimate white students

while high school teachers of both genders tend to overestimate both white and minority students. Elementary and middle school teachers tend to overestimate the ability of minority students. Finally, all three types of teachers tend to estimate the cognitive ability of minority students higher, relative to their test scores, than they estimate white students relative to their test scores.

Research Question 4

Do teacher characteristics explain the variation in the accuracy of teachers' perceptions?

Most theory and research into the accuracy of teachers' perceptions of their students has concentrated on the characteristics of individual teachers (e.g., race) to determine if any of these characteristics have a bearing on the accuracy of the teacher's perceptions. In the first and second results section, I looked at a number of characteristics of teachers to determine if any relationship exists between those qualities and teacher perceptual match both for all students and for minority students. Table 60 summarizes the results of those analyses.

Table 60: *Summary Table of Statistical Significance and Effect Size for Perceptual Match, Perceptual Match Minority, and Perceptual Match White by Teacher Characteristic Variables*

Teacher Characteristic	Statistical Significance for Perceptual Match	Effect Size for Perceptual Match	Statistical Significance for Perpetual Match Minority	Effect Size for Perceptual Match Minority	Statistical Significance for Perceptual Match White	Effect Size for Perceptual Match White
Gender	significant	trivial	significant	small	significant	trivial
Race	not significant	–	not significant	–	not significant	–
Race flag (white/non-white)	significant	trivial	not significant	–	not significant	–
Education	not significant	–	not significant	–	not significant	–
Certification	not significant	–	not significant	–	not significant	–
Total experience	not significant	–	not significant	–		
System experience	not significant	–	not significant	–		
Building experience	not significant	–	not significant	–		
Total experience flag (<2 yrs, >2 yrs)	not significant	–	not significant	–	not significant	–
System experience flag (<2 yrs, >2 yrs)	not significant	–	not significant	–	not significant	–
Building experience flag (<2 yrs, >2 yrs)	not significant	–	not significant	–	not significant	–

The only personal characteristic of teachers that had any significant effect on perceptual match was gender. Gender, while statistically significant for both perceptual match for all students and perceptual match for minority students, only had a significant effect size for perceptual match minority.

Much of the early literature on teacher perception of students has theorized that white female teachers, which make up the vast majority of teachers in the nation, may be having cultural misunderstandings with an increasingly African-American and Latino student population, particularly males. While there has been little empirical work that has demonstrated that connection, I undertook a two-way ANOVA of gender and race to determine if such a relationship does, in fact, exist.

I ran a two-way ANOVA to examine the interaction between teacher gender and teacher race. The distribution of perceptual match over gender and race of the teacher was not normal for almost all combinations of gender and race. Examining the skewness and kurtosis of each distribution, I found that perceptual match for white female teachers was negatively kurtosed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match was approximately normally distributed. I found no extreme outliers in the data. The assumption of homogeneity of variances, as assessed by Levene's test for equality of variances, was violated ($p < .001$). Unfortunately, unlike with one-way ANOVA, there is no way to resolve lack of homogeneity of variances other than data transformation. I attempted reflect and square root, reflect and logarithmic, and reflect and inverse transformations, but none had an effect on the Levene test.

I took a second look at this issue by comparing white teachers to non-white teachers. The distribution of perceptual match over gender and race of the teacher was not normal. Examining the skewness and kurtosis of each distribution, I found that perceptual match for white female teachers was negatively kurtosed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match was approximately normally distributed. I found no extreme outliers in the data. There was homogeneity of the variances, as

assessed by Levene's test for equality of variances (p = .013). The results of this two-way ANOVA can be seen in Table 61.

Table 61: *Two-Way ANOVA for Perceptual Match by Gender and Race of the Teacher*

Source	df	SS	MS	F	p	η^2
Gender	1	13.77	13.77	5.80	.016	<.01
Whiteteacher	1	23.25	23.25	9.80	.002	.01
Gender*whiteteacher	1	0.08	0.08	0.03	.856	<.01
Error	1,912	4,535.94	2.37			
Total	1,916	4,600.00				

While both gender and race (binary construct: 0 = non-white, 1 = white) are statistically significant, their partial η^2 demonstrates that the effect size is insignificant. Of note is that fact that the interaction effect between race and gender is not even statistically significant. This table shows that the interaction between race and gender is not an important factor in determining teachers' perceptual match for the student body as a whole.

It is more important to do this same test on teachers' perceptual match for minority students. The distribution of perceptual match minority over the gender and school type was not normal for white male and Asian, white, and Latino female teachers. Examining the skewness and kurtosis of each distribution, I found that perceptual match minority for white female teachers was negatively kurtosed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match minority was approximately normally distributed for gender and race of the teacher. I found no extreme outliers in the data. There was homogeneity of the variances, as assessed by Levene's test for equality of variances (p =

.057). Table 62 shows the results of this two-way ANOVA with perceptual match minority as the dependent variable.

Table 62: *Two-Way ANOVA for Perceptual Match Minority by Gender and Race of the Teacher*

Source	df	SS	MS	F	*p*	η^2
Gender	1	12.41	12.41	4.52	.034	<.01
Race	5	28.83	5.77	2.10	.063	.01
Gender*race	5	13.76	2.75	1.00	.415	<.01
Error	1,438	3,947.78	2.75			
Total	1,450	4,105.00				

Only gender is statistically significant for perceptual match minority. However, its partial η^2 of less than 0.01 demonstrates that the effect size is trivial. Of note is that fact that the interaction affect between race and gender is not even statistically significant. This table shows that the interaction between race and gender is not an important factor in determining teachers' perceptual match for minority students.

I took a second look at this issue by comparing white teachers to non-white teachers. The distribution of perceptual match minority over gender and race of the teacher was not normal. Examining the skewness and kurtosis of each distribution, I found that perceptual match minority for white female teachers was negatively kurtosed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match was approximately normally distributed. I found no extreme outliers in the data. There was homogeneity of the variances ($p = .649$). The results of this two-way ANOVA can be seen in Table 63.

Table 63: *Two-Way ANOVA for Perceptual Match White by Gender and Race of the Teacher*

Source	df	SS	MS	F	*p*	η^2
Gender	1	23.90	29.90	10.90	.001	.01
Race	1	19.83	19.83	7.23	.007	.01
Gender*whiteteacher	1	1.26	1.26	0.46	.498	.01
Error	1,446	3,965.91	2.74			
Total	1,450	4,105.00				

While both gender and race (binary construct: 0 = non-white, 1 = white) are statistically significant their partial η^2 value demonstrates that the effect size is insignificant (while indicated as ".01", the actual partial η^2 values for gender and race were .007 and .005). Of note is that fact that the interaction effect between race and gender is not even statistically significant. The interaction between race and gender is not an important factor in determining teachers' perceptual match for minority students.

In summary, I found very little evidence that personal characteristics of the teacher have any relationship with how teachers perceive their students' abilities relative to those students' performance on standardized tests. The only personal characteristic that had a significant relationship was gender. Female teachers tend to overestimate minority students by a little more than 3 percentile points. Male teachers, on average, overestimate the ability of minority students by more than 10 percentile points. Thus, my findings agree with those of most of the naturalists and the environmentalists who find that individual teacher characteristics are not related to teachers' perceptions of their students. Ready and Wright (2011) find that a teacher's level of experience was related to teacher perception, but I found no statistically significant relationship for any variation of my three experience variables (total, system, building). On the whole, my findings support the findings of most

of the naturalistic and environmental researchers in the field of teacher perceptions: the personal characteristics of teachers are not indicators of teachers' perceptions of their students.

Research Question 5

Do school characteristics explain the variation in the accuracy of teachers' perceptions?

While most theory and research into the accuracy of teachers' perceptions of their students has concentrated on the characteristics of individual teachers, the environmentalists find that it was classroom-context variables that are more important in these teacher perceptions. This section will look at the characteristics of the schools where the teachers work to determine if any relationship exists between those characteristics and perceptual match. Because Ready and Wright (2011) had the most comprehensive model of teacher perception, I examined the same independent variables they use in their study: the size of the school, the socio-economic status of the parents of the children served by the school, and the type of community the school was located in (urban, suburban, small city, town, rural).

In the first and second results sections, I looked at a number of characteristics of schools to determine if any relationship exists between those and teacher perceptual match both for all students and for minority students. Table 64 summarizes the results of those analyses.

Table 64: *Summary Table of Perceptual of Statistical Significance and Effect Size for Perceptual Match, Perceptual Match Minority, and Perceptual Match White by Environmental Variables*

School Characteristic	Statistical Significance for Perceptual Match	Effect Size for Perceptual Match	Statistical Significance for Perpetual Match Minority	Effect Size for Perceptual Match Minority	Statistical Significance for Perceptual Match White	Effect Size for Perceptual Match White
SES	significant	large	significant	small	significant	medium
Community	significant	medium	significant	medium	significant	small
School size	not significant	–	not significant	–	not significant	–

Both socioeconomic status (SES) of the parents of the students the school serves and the type of community the school was located in were statistically significant with a medium effect size. Ready and Wright had found both of these variables to be significant indicators of teacher perceptual match in their study.

I examined a two-ANOVA of SES and type of community to determine if there were any interaction effects. The distribution of perceptual match over the SES of the school's parents and the type of community the school was located in was not normal for almost all combinations of the two variables. Examining the skewness and kurtosis of each distribution, I found that perceptual match for schools with low- or middle-low income students in urban areas and suburban areas were negatively skewed. In addition, the distribution for suburban schools with low-income students was positively kurtosed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match was approximately normally distributed over SES and type of community. I found no extreme outliers in the data. The assumption of homogeneity of variances was violated ($p < .001$).

Unfortunately, unlike with one-way ANOVA, there is no way to resolve lack of homogeneity of variances other than data transformation. I attempted reflect and square root, reflect and logarithmic, and reflect and inverse transformations, but none had an effect on the Levene test of homogeneity of variances.

I took another pass at the data. I created two SES categories out of the four that remained in the data base (I had earlier combined the small number of observations in the upper-income category of SES into the high-middle-income category): middle income (high-middle income + middle income) and lower income (lower-middle income + lower income). I then examined the distribution of teachers' perceptual match over my new SES variable and my community type variable. The distribution of perceptual match over SES and type of community was not normal except for middle-income suburban neighborhoods. Examining the skewness and kurtosis of each distribution, I found that perceptual match for teachers in lower-income urban neighborhoods and in all suburban neighborhoods were negatively skewed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match was approximately normally distributed. I found no extreme outliers in the data. There was homogeneity of the variances ($p = .196$). The results of this two-way ANOVA can be seen in Table 65.

Table 65: *Two-Way ANOVA for Perceptual Match by Socioeconomic Status (SES) of the Parents of the Children the School Serves (Middle/Lower SES) and the Type of Community the School Is Located In*

Source	df	SS	MS	F	*p*	η^2
SES	1	177.39	177.39	84.94	<.001	.04
Community	4	81.81	20.45	9.79	<.001	.02
SES*community	4	88.02	22.00	10.54	<.001	.02
Error	1,959	4,090.93	2.09			
Total	1,969	4,734.00				

Bifurcated SES (lower/middle SES), type of community, and their interaction were all statistically significant and all had small effect sizes. Schools located in urban areas, suburban areas, and small cities all had statistically significant differences between lower income neighborhoods and middle income neighborhoods (see Table 66). In lower SES neighborhoods, there is a statistically significant difference among the five types of communities in perceptual match scores, with urban schools having the highest scores and towns having the lowest scores (see Table 67).

Table 66: *Simple Main Effect for Socioeconomic Status (SES) of the Parents of the Students the School Serves (Middle/Lower SES) in Two-Way ANOVA for Perceptual Match by SES of the Parents of the Children the School Serves and the Type of Community the School Is Located In*

Type of Community	df	SS	MS	F	*p*	η^2
Urban Contrast	1	94.27	94.27	45.14	<.001	.02
Error	1,959	4,090.93	2.09			
Suburban Contrast	1	108.30	108.30	51.86	<.001	.03
Error	1,959	4,090.93	2.09			
Small City Contrast	1	78.48	78.48	37.58	<.001	.02
Error	1,959	4,090.93	2.09			
Town Contrast	1	0.00	0.00	0.00	.973	.00
Error	1,959	4,090.93	2.09			
Rural Contrast	1	3.48	3.48	1.67	1.97	.00
Error	1,959	4,090.93	2.09			

Table 67: *Simple Main Effect for Type of Community in Two-Way ANOVA for Perceptual Match by Socioeconomic Status (SES) of the Parents of the Children the School Serves (Middle/Lower SES) and the Type of Community the School Is Located In*

SES	df	SS	MS	F	*p*	η^2
Lower income Contrast	4	299.01	299.01	35.80	<.001	.07
Error	1,959	4,090.93	2.09			
Upper income Contrast	4	13.40	13.40	1.60	.171	<.001
Error	1,450	4,105.00				

I then did the same analysis SES and type of community to for perceptual match minority to determine if there were any interaction effects when looking particularly at minority students. The distribution of perceptual match minority over SES of the parents of the students the school serves and type of community of the school was located in was not normal for almost all combinations of the two variables. Examining the skewness and kurtosis of each distribution, I found that perceptual match for schools with low-income students in urban and suburban areas were negatively skewed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match minority was approximately normally distributed. I found no extreme outliers in the data. The assumption of homogeneity of variances was violated ($p < .001$). I attempted reflect and square root, reflect and logarithmic, and reflect and inverse transformations, but none had an effect on the Levene test.

I took a second look at these data as I had with perceptual match for all students. I created two SES categories out of the four that remained in the data base: middle income (high-middle income + middle income) and lower income (lower-middle income + lower income). I then examined the distribution of teachers perceptual match for minority students over my new SES variable and my community type variable. The distribution of perceptual match minority over SES and type of community of the school was not normal

except for middle-income suburban neighborhoods. Examining the skewness and kurtosis of each distribution, I found that perceptual match minority for teachers in lower-income urban neighborhoods was negatively skewed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match minority was approximately normally distributed. I found no extreme outliers in the data. There was homogeneity of the variances (p = .023). The results of this two-way ANOVA can be seen in Table 68.

Table 68: *Two-Way ANOVA for Perceptual Match Minority by Socioeconomic Status (SES) of the Parents of the Children the School Serves (Middle/Lower SES) and the Type of Community the School Is Located In*

Source	df	SS	MS	F	*p*	η^2
SES	1	52.34	52.34	20.19	<.001	.01
Community	4	78.39	19.60	7.56	<.001	.02
SES*community	4	28.96	7.24	2.79	.025	.01
Error	1,475	3,824.38	2.59			
Total	1,485	4,219.00				

Bifurcated SES (lower income/middle income) and type of community were statistically significant but the interaction effect was not significant. Thus, for minority students, teacher perceptions vary by SES and type of community as I have reported above, but there is no interaction effect between these two variables.

I attempted a two-way ANOVA for perceptual match white by SES and type of community but could find no configuration of the data that prevented the homogeneity of variances assumption from being violated. I attempted to transform the data and created new consolidated versions of the independent variables, however none of these techniques yielded results that met the homogeneity assumptions.

Finally, I constructed a factorial ANOVA model with all the independent variables that Ready and Wright (2011) find significant in their model. I added gender as well. As I

have explained, Ready and Wright did not include gender of the teacher as an independent variable in their analysis because they were using a sample of kindergarten teachers and had very few male teachers in their sample. My first model, including gender, new teacher flag (< 2 years' experience, > 2 years' experience), SES, and type of community did not meet the assumptions of homogeneity of variances required by factorial ANOVA. I modified SES as I had before in my two-way ANOVAs, bifurcating it into middle income and lower income categories.

The distribution of perceptual match was not normal for many combinations of gender, new teacher, SES, and type of community. Examining the skewness and kurtosis of each distribution, I found that perceptual match for new and experienced female teachers in urban schools serving middle-income students was negatively skewed. I also found that the distributions of perceptual match for experienced teachers in schools serving lower-income students located in suburbs and towns were positively skewed. From visual inspection of the Normal Q-Q Plot I determined that perceptual match was approximately normally distributed. I found no extreme outliers in the data. There was homogeneity of the variances ($p = .039$). The results of this factorial ANOVA can be seen in Table 69.

Table 69: *Factorial ANOVA for Perceptual Match by Gender of the Teacher, Amount of the Teacher's Experience in the School Building (< 2 Years', > 2 Years'), Socioeconomic Status (SES) of the Parents of the Children the School Serves (Middle/Lower SES) and the Type of Community the School Is Located In*

Source	df	SS	MS	F	*p*	η^2
Gender	1	12.90	12.90	6.34	.012	<.01
New teacher	1	1.19	1.19	0.58	.445	<.01
SES	1	42.205	42.25	20.76	<.001	.01
Community	4	26.45	6.61	3.25	.011	.01
Gender*new teacher	1	1.01	1.01	0.50	.481	<.01
Gender*SES	1	0.67	0.67	0.33	.567	<.01
Gender*community	4	8.92	2.23	1.10	.357	<.01
New teacher*SES	1	2.67	2.67	1.31	.252	<.01
New teacher*community	4	3.34	0.84	0.41	.801	<.01
SES*community	4	47.66	11.92	5.86	<.001	.01
Gender*new teacher*SES	1	2.89	2.89	1.42	.233	<.01
Gender*new*community	4	14.43	3.61	1.77	.132	<.01
Gender*SES*community	4	4.39	1.10	.54	.707	<.01
New teacher*SES*community	4	7.03	1.76	.86	.485	<.01
Gender*new*SES*community	4	4.97	1.24	.61	.655	<.01
New teacher*SES*community	4	7.03	1.76	.86	.485	<.01
Error	1,795	3,652.83	2.04			
Total	1,835	4,378.00				

There are no additional interaction effects from this model above what I had already found in the two-way ANOVA above for SES and type of community. While gender was statistically significant, the effect size was insignificant.

I wanted to construct a similar model for perceptual match for minority students and perceptual match for white students. My first model for perceptual match minority, including gender, new teacher flag (< 2 years' experience, > 2 years' experience), SES, and type of community did not meet the assumptions of homogeneity of variances required by factorial ANOVA ($p < .001$). I modified SES as I had before, bifurcating it into middle income and lower income. I also had to modify community, collapsing small

city, town, and rural into a single category for non-urban in order to have homogeneity of variance, a prerequisite for factorial ANOVA.

The distribution of perceptual match was not normal for many combinations of gender, new teacher, SES, and type of community. Examining the skewness and kurtosis of each distribution, I found that none of the distributions of perceptual change minority were skewed or kurtosed. I found three extreme outliers in the data. Two were in a cell for new male teachers in schools located in a schools serving middle-income students located in towns. Because there were only six observations in that cell, I left the extreme outliers in the data set. I removed the other extreme outlier, in the cell for new female teachers working in urban schools serving middle-income students because there were 32 observations in the cell and its removal would not seriously affect the variance. There was homogeneity of the variances ($p = .039$). The results of this factorial ANOVA can be seen in Table 70.

Table 70: *Factorial ANOVA for Perceptual Match Minority by Gender of the Teacher, Amount of the Teacher's Experience in the School Building (< 2 Years', > 2 Years'), Socio0Economic Status (SES) of the Parents of the Children the School Serves (Middle/Lower SES), and the Type of Community the School Is Located In*

Source	df	SS	MS	F	*p*	η^2
Gender	1	1.26	1.26	.49	.485	<.01
New teacher	1	1.71	1.71	.66	.416	.01
SES	1	32.75	32.75	12.71	<.001	.01
Community	2	22.60	11.30	4.38	.013	.01
Gender*new teacher	1	4.13	4.13	1.60	.206	<.01
Gender*SES	1	0.92	0.92	0.36	.549	<.01
Gender*community	2	1.22	0.61	0.24	.789	<.01
New teacher*SES	1	0.02	0.02	0.01	.933	<.01
New teacher*community	2	3.00	1.50	0.58	.559	<.01
SES*community	2	9.02	4.51	1.75	1.74	.01
Gender*new teacher*SES	1	0.52	0.52	0.20	.654	<.01
Gender*new*community	2	5.01	2.50	0.97	.379	<.01
Gender*SES*community	2	0.30	0.15	0.06	.424	<.01
New teacher*SES*community	2	4.42	2.21	0.86	.424	<.01
Gender*new*SES*community	2	0.79	0.40	0.15	.860	<.01
Error	1,364	3,514.52	2.58			
Total	1,388	3,922.00				

There are no additional interaction effects from this model above what I had already found in the two-way ANOVA above for SES and type of community.

I next attempted to construct a similar model for perceptual match for white students. My first model for perceptual match white, including gender, new teacher flag (< 2 years' experience, > 2 years' experience), SES, and type of community did not meet the assumptions of homogeneity of variances required by factorial ANOVA ($p < .001$). I attempted to transform the data as well, with square root, reflect and square root, log, and reflect and log transformations, but the homogeneity assumption continue to be rejected.

I modified SES as I had before, bifurcating it into middle-income and lower-income SES. I also had to modify community, collapsing small city, town, and rural into a single category for non-urban. Again, assumptions of homogeneity of variances were not met and, again, data transformation did not help. The result was that I could not run this model for white students.

In summary of this section I found that the SES of the parents of the students served by the school and the type of community the school was located in were significant factors in determining perceptual match scores for all students and for minority students. In addition, I found that there was an interaction effect between SES of the parents of the students served by the school and type of community when looking at perceptual match for all students. Teachers in schools in urban and suburban areas as well as in small cities had different perceptual match scores based on the SES of the school neighborhood. Schools serving middle-income students had significantly lower perceptual match scores for all students than school serving lower-income students in these three types of communities. Perceptual match scores for all students in towns and rural areas did not differ significantly over the two SES categories. In addition, for schools serving low-income students, the differences between each of the five types of communities were all significant as well.

Summary

In this study, I used one-way, two-way, and factorial ANOVA to empirically examine my five research questions. I found evidence that teachers' perceptions of their students' performance is accurate when compared to school-wide standardized test results.

However, teachers consistently tend to overestimate the abilities of minority students compared to the abilities of white students.

My findings support the environmental hypothesis that the school environment, rather than individual teacher characteristics, is the primary indicator of variation in teachers' perceptions of their students, particularly minority students. The following findings support the environmental hypothesis:

1. Socioeconomic status of the neighborhood where the school is located:
    a. As SES increases, teachers' perceptual match scores decreases, i.e., they are more likely to underestimate their students; teachers in schools with the highest (high income + upper middle income) SES *under*estimate their students standardized test scores by almost 26 percentile points; teachers in schools with low income students *over*estimate their students by more than 10 percentile points.
    b. This relationship holds true for teachers' estimates of their minority students' abilities vis-a-vis standardized test scores: teachers in the schools with the highest SES students have low perceptual match scores, underestimating student ability by an average 4 percentile points; teachers in the schools with low-income students have high perceptual match scores, overestimating minority students by 11 percentile points.
    c. For all four of the SES categories, the teacher perceptual match for minority students is significantly higher than teacher

perceptual match for white students, meaning they had higher perceptions of minority students relative to their test scores than they had of white students.

2. Type of neighborhood the school is located in:
    a. Teachers in urban neighborhoods have significantly higher perceptual match scores than teachers in every other type of neighborhood, averaging 13 percentile points over students' standardized test scores.
    b. Teachers in small towns have significantly lower perceptual match scores than teachers in every other type of neighborhood, averaging 11 percentile points below their students' standardized test scores.
    c. Teachers in urban neighborhoods have significantly higher perceptual match scores for minority students than teachers in every other type of neighborhood, averaging nearly 18 percentile points over students' standardized test scores.
    d. There was a significant difference between teachers' perceptions of minority students and teachers' perceptions of white students in every type of community except rural communities.
3. Interaction between SES and type of community:
    a. Teachers in low-SES neighborhoods in urban and suburban areas and in small cities rated their students higher than

teachers in middle income neighborhoods in the same types of communities.

b. For teachers in low-SES neighborhoods, there is a significant difference in perceptual match among all five types of communities with urban areas having the highest scores and towns having the lowest.

I have only one finding that indicates that individual teacher characteristics may have an influence on their perception of student ability. I found that, for minority students only, the gender of the teacher does make a difference in how he/she rates his/her students relative to their standardized test scores. While both male and female teachers tend to have high perceptual match scores for their minority students, i.e., overate their minority students relative to their test scores, males overestimate the ability of their minority students an average of 10 percentile points, females by an average of 3 percentile points.

Chapter 5: Summary, Implications, and Recommendations

Introduction

Chapter IV presented and analyzed the data. This chapter is a summary of the study's results and a discussion of my findings and conclusions, including recommendations for further research.

Summary

Teachers' perceptions of their students have a profound impact on the children they teach. Teachers' perceptions often end up influencing a variety of educational placement decisions, including grouping by ability, grade retention, exposure to curricula, admission to selective programs or schools, assignment to English as a second language, and assignment to special education. In turn, these decisions can have agency on a child's adult life. If teacher perceptions are biased in any way that hurts racial and ethnic minorities in particular, they may be contributing to our nation's current educational inequalities.

Because the vast majority of elementary school teachers are white females and public schools have higher percentages of minority children than society as a whole, it is easy to imagine a teacher interpreting a student's cultural mannerisms as an indicator of academic ability. Researchers who study teacher perceptions seek to determine whether educational inequality stems from those perceptions, or from real cognitive differences among socio-demographic groups.

Accuracy of Teachers' Perceptions of Their Minority Students' Performance

How accurate are teachers' perceptions of their students' performance when compared to standardized test results? Experimental researchers argue that, in general, teachers underestimate the cognitive ability of racial minorities and that personal characteristics of teachers (e.g., race) can predict which teachers will underestimate minority students the most. Naturalistic researchers mostly find that this is not the case, that teacher perceptions match actual student performance. The environmentalists find bias in teachers' perceptions, but attribute that to environmental factors rather than teacher personal characteristics. This first research question looks at the questions of bias and accuracy.

Findings: Teachers' perceptions of their students' abilities, on average, were quite accurate. However, this accuracy varied widely depending on the SES and type of community the school is located in. These findings contradict the naturalist hypothesis that there is no bias in teachers' perceptions of their students. They also support the environmental hypothesis that characteristics of the school's setting, rather than teacher personal characteristics, are stronger indicators of teacher perception.

However, the findings were surprising to the extent they were the mirror image of the findings of the environmentalists and what I expected to discover. The environmentalists find that teachers *under*estimate poor and urban students. I found, on the contrary, that teachers *over*estimate poor and urban students while *under*estimating wealthier students in more rural areas.

Accuracy of Teachers' Perceptions of Minority Students

Are teachers' perceptions of minority students as accurate as their perceptions of white students?

Findings: Teachers significantly overestimate minority students' cognitive ability relative to their perception of white students' cognitive ability. On average, teachers overestimate minority students by nearly 5 percentile points relative to what they actually score on standardized tests. Teachers *under*estimate white students by more than 5 percentile points relative to their standardized test scores.

Of all the personal characteristics of teachers I examined for their distributions of perceptual match minority and perceptual match white, only the distribution of perceptual match minority across gender showed statistically significant differences with a small effect size. Male teachers overestimated minority students significantly more than female teachers did.

The distribution of perceptual match minority and perceptual match white across the socioeconomic status of the parents of the students served by the school as well as the type of community the school was located in had statistically significantly differences. Schools that served low-income students and urban schools were more likely to have higher perceptual match minority and perceptual match white scores than schools serving wealthier students and less-urban areas.

Almost across the board, there were statistically significant differences between the perceptual match minority score and the perceptual match white score (see Tables 60 and 64). In every case, teachers had higher perceptions of minority students relative to their standardized test scores than they did of white students. Be clear, I am not saying that

teachers had higher perceptions of their minority students than they did of all students. What I am saying is teachers overestimated (or underestimated less) their minority students relative to their standardized test scores than they did their white students. These findings contradict the findings of the experimentalists and the naturalists. The experimental researchers find that teachers perceive minority and white students in different ways. My findings support that finding. However, the experimentalists argue that teachers *under*estimate minority students relative to white students. And, indeed, they do: teachers in the KEYS database on average gave better ratings to white students than to minority students. However, relative to their standardized test scores, teachers perceived minority students better than they did their white students. My findings also contradict the naturalists' findings because I found that there is a significant difference between teachers' perceptions of minority students, relative to their test scores, and teachers' perceptions of white students relative to their test scores.

Teacher Accuracy by Grade

Do teachers' perceptions vary by grade?

Findings: Grade level makes a difference. Elementary and middle school teachers tend to be quite accurate in their estimate of their students' abilities while high school teachers tend to overestimate their abilities. Elementary school teachers are also quite accurate in their estimates of their minority students' abilities, while middle and high school teachers tend to overestimate their minority students' abilities. Elementary and middle school teachers underestimated white students while high school teachers overestimated them. Still, teachers at all school levels had higher perceptions of their minority students relative to minority standardized test scores

than they did of all of their white students relative to white student standardized test scores.

Relationship of Personal Characteristics of Teachers to Teacher Accuracy

Do teacher characteristics explain the variation in the accuracy of teachers' perceptions? With this research question I was examining some of the traditional theories of teacher perceptions: that certain teacher characteristics lead to more inaccurate perceptions of students' abilities, particularly those of minority students. Most naturalistic researchers in this field find that teacher characteristics have no significant impact on teacher perceptions. The environmentalists find that personal characteristics of teachers are not significant indicators of perceptual inaccuracy. Research question 4 examines the hypothesis that personal characteristics of teachers affect the accuracy of their perceptions of their students.

Findings: None of the personal characteristic variables which I used in my model had a statistically significant relationship with teacher perceptual match scores for all students, minority students, or for white students. The lone exception was that gender of the teacher was a significant predictor of teacher perceptions of minority students with male teachers, on average, overestimating minority students significantly more than female teachers did (female teachers also overestimated minority students on average, just significantly less than male teachers). There was also a statistically significant difference between the perceptual match score (for all students) for male teachers and female teachers, however the effect size was trivial. I explicitly tested the experimental theory that white female teachers systematically underestimate minority students' abilities and found no evidence for that idea with

these data. These findings are in keeping with most of the naturalist and environmentalist work in the field of teacher perception.

Relationship of the Environmental Characteristics of the Schools Teachers Work in to Teacher Accuracy

Do school characteristics explain the variation in the accuracy of teachers' perceptions? With this research question I was examining the environmentalist theory that characteristics of the school and the neighborhood it is located in, rather than the individual characteristics of the teachers themselves, have the most important influence on teachers' perceptions of their students. Several of my findings support the environmentalist hypothesis.

Findings: Characteristics of the schools that teachers teach in, rather than personal characteristics of the teachers themselves were the most important factor in determining teacher perceptual match for all students, for minority students, and for white students.

I found that teachers in schools that served higher-income students were more likely to *under*estimate student performance than teachers in schools that served lower-income students. A similar pattern was repeated with teachers' perceptions of their minority students and their white students. Teachers in schools serving upper-middle-income students were fairly accurate in their perception of minority students. Teachers in in schools serving low-income students significantly overestimated minority students. Similarly, as the income of the students served by the school increased, teachers were more likely to underestimate their white students. For all categories of SES, teacher perceptual match minority scores were significantly higher

than teacher perpetual match white scores. This means that, on average, teachers have higher perceptions of minority students relative to their actual standardized test scores than they do of white students relative to their standardized test scores.

Teachers in urban neighborhoods had statistically significant higher perceptual match scores for all students, for minority students, and white students than teachers in any other type of community. Teachers in urban areas overestimated all their students, their minority students, and their white students significantly more in any other type of community. Teachers consistently rated minority students higher, relative to their standardized test scores, than they rated white students in all communities except rural communities. In rural communities there was no difference between how teachers rated minority and white students.

I found an interaction effect between SES and type of community. There was a statistically significant difference in the way teachers in schools that served upper-income students and teachers in schools that served lower-income students perceived their students for teachers working in urban areas, the suburbs, or small cities. There was no difference in the perceptions of teachers who worked in schools serving upper-income children and schools serving lower-income students for schools located in towns and rural areas.

Furthermore, for teachers in schools serving lower-income children, there was a statistically significant difference in how teachers in each type of community perceived their students relative to their test scores. The general trend in these schools serving lower-income schools was that the more urban the environment that school was located in, the more likely teachers were to overestimate students relative to their

test scores; the more rural the school was the more likely the teachers were to underestimate their students relative to their standardized test scores.

Discussion

My findings support the environmentalists who posit that environmental characteristics are more important than personal characteristics of the teacher in determining teachers' perceptions of their students. However, there is one very big difference in the results: while the environmental thesis finds that teachers in low-income urban areas *under*estimate their students, I found that they are *over*estimating those students.

These results differ from all previous studies in the field of teacher perception research. Teacher perception research is premised on the concept that teachers systematically underestimate the abilities of low-income and minority students. Although the vast majority of naturalistic researchers find no evidence of systematic underestimation by teachers in their perceptions of low-income and minority students, the point of past research has always been to determine if teachers *under*estimate minority students compared to white students. Therefore, my findings that teachers on average *over*estimate minority students relative to their standard testing results was quite surprising.

One explanation for these surprising results may be that poor urban students are getting poor access to curriculum, that they are receiving "dumbed-down" curriculum. In KEYS, teachers are asked about performance of their target class. It is easy to see a teacher responding to the KEYS questionnaire in the context of the other students they have seen over the years in their classes and the context of the type of

curriculum offered in his/her school (or class). The teacher might respond to this question relative to the milieu in which they have worked. Thus, for example, a teacher in a low-income urban neighborhood is probably no comparing his/her target class to students nationwide, but to students that he/she *actually seen over the years* in courses that are, on average, not as challenging as courses offered in less-urban areas with higher average income. If, is as indeed is the case, lower-income students tend to perform, on average, less well than higher-income students, then the teacher will be exposed, on average, to a group of students who perform less well on standardized tests. The teacher in the KEYS survey response is rating his/her students to similar classes they have taught over the years in this lower-performing milieu. This will lead, again, on average, to overestimation of those students' cognitive ability as measured on standardized tests. Standardized test results put students on the national stage. Results are comparable across the country. This is an explanation for why teachers in low-income and urban areas may end up overestimating students' abilities relative to standardized test data.

Linda Darling-Hammond points out that urban students "face dramatic differences in courses, curriculum materials, and equipment" (2004, p. 617). Jonathon Kozol's book, *Savage Inequalities* (1991), is a litany of the contrasts in equipment, materials, and funding between poor urban school districts and their wealthier suburban counterparts.

Lutz Berkner and Linda Chavez (1997), found that only 53% of low-income students were prepared to attend college (compared to 68% of middle-income students and 86% of higher-income students). And, of that 53%, 60% received a low

college-preparedness ranking on Berkner & Chavez' four-point scale (i.e., only 21% of low-income students were "highly-qualified" or "very-highly qualified" to attend college). A more recent study by ACT (2014), found that only 19% of low-income students met three out of four benchmarks for college preparedness. And, low-income students were significantly less likely to meet any of the four benchmarks than other income groups.

Low-income urban schools are also less likely to offer advanced courses in mathematics and the sciences (Oakes, Joseph & Muir, 2004). The National Center for Education Statistics reports that low-income, African-American, and Latino students take far fewer advanced placement courses than white and Asian students (National Center for Education Statistics, 2007).

Robert Marzano (2003) in his meta-analysis of meta-analyses (or meta-meta-analysis), ranks a guaranteed and viable curriculum as the factor having the most impact on student achievement of all the school-level factors he analyzed. He notes that there is a difference between the intended curriculum, the implemented curriculum, and the attained curriculum. The state or school district issues a curriculum which they expect teachers to teach. This is the intended curriculum. What teachers actually teach from the intended curriculum is the implemented curriculum. What students actually learn from the implemented curriculum is the attained curriculum. Researchers label the gap between the intended curriculum and the implemented curriculum as one of many lost "opportunities to learn." The closer the intended curriculum and the implemented curriculum are together, the greater the

opportunity to learn. The further apart they are, the less opportunity children have to learn.

Bokhee Yoon, Leigh Burstein, and Karen Gold (n.d.) show that the opportunity to learn gap is quite large, that, indeed, there is a large difference between what states and school districts intend with their curriculum and what teachers actually implement in the classroom. Of course, children who do not have the opportunity to learn the content that they are expected to learn cannot learn it.

A second, related explanation may be the cognitive dissonance of the teachers. Leon Festinger introduced the concept of cognitive dissonance in 1957 (Festinger & Carlsmith, 1959). Cognitive dissonance occurs when a person believes something, "X", but as a result of some sort of pressure or reward, publically avows "not X." Festinger posits that human beings have an inner drive to hold all of our attitudes and beliefs in harmony. When a person has cognitive dissonance, he/she is pressured by their own desire for inner harmony to act in one of three ways:

1. change the inner attitude towards X so that it aligns with "not X";
2. acquire new information that explains the relation of X and "not X";
3. reduce the impact of the cognition so that "not X" is not as important.

Teachers may be reacting with cognitive dissonance to their low-income urban students' actual test scores by using strategies 2 and 3 above. Teachers may "know" that their students are better than what their test scores indicate (strategy 2). Or they may argue to themselves (and others) that standardized tests do not really measure the true cognitive ability of their students (strategy 3). In either case,

teachers will perceive low-income urban students to have better cognitive abilities than those students' standardized test scores would indicate.

This cognitive dissonance is compounded because low standardized test scores reflect poorly on more than the students. Low standardized test scores can also be a reflection on the ability of the teacher his/herself. As Robert Merton points out in his book *Social Theory and Social Structure* (1957), teachers (he uses bureaucrats in his book) sanctify goals, i.e., give them moral legitimacy. Many teachers, particularly in urban areas, view themselves as saviors of children, or, if not saviors, at least rescuers. These teachers often are on a moral mission. However, Martin Haberman (1995) writes that the best teachers tend to be nonjudgmental and are not moralistic. They do not consider themselves as saviors and they do not expect the school, or the school system, to change because of their actions. Rather, they enjoy interacting with children and getting children learn, despite the obstacles.

The teacher that views him/herself as the savior or rescuer faces cognitive dissonance from standardized test results. Thus, it is feasible that the teacher can solve his-her own internal conflict by grading (or perceiving) students as better than ther actual (in class) performance might merit. They not only "protect" their students from external criticism, they protect themselves from that same criticism that suggests they might not be as good teachers as they themselves believe. Like Merton's bureaucrat, they sanctify their work, creating a moral mission of saving and rescuing out of giving children the opportunity to learn.

As Ready and Wright (2011) point out in their study of the accuracy of teacher perceptions, teachers play the dual role of umpire and coach in American

society. Their perceptions of their students drive important decisions made by society, especially those made for educational purposes, i.e., academic placement. Because children's educational outcomes so often have profound influences on their success in life, it is important that education researchers fully understand if teacher perception, particularly teacher biases, have any significant influence on these outcomes.

The Ready and Wright study had been the most comprehensive study to date on teachers' perceptions of their students. It was the first study to use a national database to examine the question and it examined a host of independent variables, many of which had not been studied before in the field of teacher perception. This study adds to research in the field of teacher perception by expanding the Ready and Wright study in two ways:

1. I examined teachers in all grades, not just kindergarten; and,
2. I examined the role of gender as an independent variable.

My results contradicted almost all the work in the field of teacher perception. Unlike the naturalists, I found evidence of systematic teacher inaccuracy in their perception of students. My results contradicted the findings of the experimentalists as well. While, like the experimentalists, I found evidence of systematic inaccuracy in teachers' perceptions of students, I found that teachers were more likely to *over*estimate the abilities of lower-income and urban students. Similarly, while my findings coincide with the environmentalists in that I found that environmental factors rather than personal characteristics of the teacher explain variance in teachers' perceptions of their students, my results turned out to be the opposite of Ready and Wright. Like them, I found that the SES of the parents of the students served by the

school and the type of community the school was located in were significant determiners of teacher perception of students. Unlike Ready and Wright, I found that teachers in lower-income urban schools *over*estimate their students' abilities relative to their standardized test scores. And, across the board, for almost every independent variable used in this study, teachers perceived the performance of minority students to be higher, relative to their test scores, than the performance of white students.

Research has demonstrated that low-income, urban students are not as prepared for post-high-school life as well as wealthier, non-urban students (see, ACT (2014); Berkner & Chavez, 1997; Darling-Hammond, 2004; Jencks & Phillips, 1998; Kozol, 1991; and Oakes, et al, 2004). Marzano, in his meta-meta analysis, cites a guaranteed and viable curriculum as the most important factor for schools to work on to ensure student achievement.

If the connection I am seeing between curriculum and teacher perception is an accurate explanation for the contradictory results I have found, then curriculum reform, at the school level, at the district level, and nationally is important.

School leaders should be clear in identifying and in communicating what they feel is the most essential content of district, state, and national curriculums. School leaders should ensure that teachers teach that essential content, i.e., implement the curriculum.

Nationally, the creation of the Common Core State Standards is a big step forward for curriculum reform. It will allow the majority of teachers in the country to work in a single basic standard. The Common Core creates nationally consistent guidelines of what every child must know in order to be successful in their life after

high school. It will provide efficiencies by allowing states to work together on a single test of students' knowledge. It will provide efficiencies in professional by having a single standard that almost all teachers work from.

Finally, we have to improve teacher education in this country. As Haberman (1995) finds, the best teachers are those who are not on a moral crusade. They are those who have learned to listen to their students, listen to their students' parents, and listen to the community in which the school is located. Teachers have many negative conditions to overcome in their work. The best teachers realize that teaching is a puzzle to be solved and that the puzzle will never end. These teachers have the attitude that all children can learn and that it is the teacher's job to get all children to be interested in learning. Our teacher education system must inculcate these qualities in our new teachers. It is a dilemma of sorts. We want teachers to believe in what they are doing and to be excited about their work. But at the same time, we do not want them to be on a crusade.

I would urge education researchers to use my analysis as an impetus to look deeper into the environmental hypothesis proposed by Ready & Wright (2011): namely, that educational contexts have important effects on teacher perceptions of their students and therefore on the equity of educational outcomes.

I would extend the Ready and Wright study to higher grades, leveraging the same data that Ready and Wright used in their study, the National Center for Education Statistics' Early Childhood Longitudinal Study (The 1998-99 cohort which examines students who were in kindergarten in 1998-99 through 8th grade and the 2010-11 cohort which examines students who were in kindergarten in 2010-11

through the 5th grade).This extension of analysis to other grades would be particularly interesting in light of my findings from the KEYS data that showed that grade level was a significant factor in determining teacher perceptions of all their students and of their minority students.

Finally I would like to see my concept of perceptual match used with other data sets. It would be interesting to if one received the same results as Ready & Wright.

Bibliography

ACT (2014). *The condition of college & career readiness 2014: Students from low-income families*. Iowa City, IA: ACT.

Baron, T., Tom, D.Y.H., & Cooper, H.M. (1985). Social class, race and teacher expectations. In J.B. Dusek (Ed.), *Teacher expectancies*. Hillsdale, NJ: Lawrence Erlbaum

Berkner, L. & Chavez, L. (1997). *Access to postsecondary education for the 1992 high school graduates*. Washington, DC: National Center for Education Statistics. Retrieved from: https://nces.ed.gov/pubs98/98105.pdf

Yoon, B., Burstein, L., & Gold, K. (n.d.). *Assessing the content validity of teachers' reports of content coverage and its relationship to student achievement*. Los Angeles, CA: University of California, Los Angeles.

Brophy, J.E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology* 75, 631-661. doi: 10.1037/0022-0663.75.5.631

Burkham, D.T., LoGerfo, L.F., Ready, D.D., & Lee, V.E. (2007). The differential effects of repeating kindergarten. *Journal of Education for Students Placed at Risk*, 12, 103-136. doi: 10.1080/10824660701261052

Darling-Hammond, L. (2004). What happens to a dream deferred? The continuing quest for equal educational opportunity. In J.A. Banks & C.A. M. Banks, *Handbook of research on multicultural education*. San Francisco, CA: Jossey-Bass

Delpit, L. (2006). *Other people's children: Cultural conflict in the classroom*. New York, NY: New

DeMeis, D.K. & Turner, R.R. (1978). Effects of students race, physical attractiveness, and dialect on teachers' evaluations. *Contemporary Educational Psychology*, 3, 77-86. doi: 10.1016/0361-476X (78)90012-7

Downey, D.B., & Pribesh, S. (2004). When race matters: Teachers' evaluations of students' classroom behavior. *Sociology of Education*, 77, 267-282. doi: 10.1177/003804070407700401

Elashoff, J.D., & Snow, R.E. (1971). *Pygmalion reconsidered: A case study in statistical inference: Reconsideration of the Rosenthal-Jacobson data on teacher expectancy*. Worthington, OH: C.A. Jones.

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, 1119-1146. Retrieved from: http://www.tcrecord.org.proxy-um.researchport.umd.edu/library/Issue.asp?volyear=2003&number=6&volume=105

Ferguson, R.F. (2003). Teachers' perceptions and expectations and the Black-white test score gap. *Urban Education*, 38, 460-507. doi: 10.1177/0042085903038004006

Festinger, L. & Carlsmith, J.M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal Psychology,* 58, 203-210. doi: 10.1037/h0041593

Haberman, M. (1995). *Star teachers of children in poverty*. West Lafayette, IN: Kappa Delta Pi.

Huck, S.W. (2008). Reading statistics and research. New York, NY: Pearson.

Institute of Education Sciences, Digest of Education *S*tatistics 2012. Washington, DC: U.S. Department of Education. Retrieved from: http://nces.ed.gov/programs/digest/

Jencks, C., & Phillips, M. (1998). *The Black-white test score gap*. Washington, DC: Brookings Institution.

Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. *Journal of Personality and Social Psychology*, 57, 469-480. doi: 10.1037/0022-3514.57.3.469

Jussim, L. (1991). Social perception and social reality: A reflection-construction model. *Psychological Review* 98, 54-73. doi: 10.1037/0033-295X.98.1.54

Jussim, L., Eccles, J., & Madon, S.J. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. *Advances in Experimental Social Psychology*, 29, 281-388. Retrieved from http://www.rcgd.isr.umich.edu/garp/articles/eccles96f.pdf

Jussim, L., & Harber, K.D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review* 9(2), 131-155. doi: 10.1207/s15327957pspr0902_3

Kozol, J. (1991). *Savage inequalities: Children in America's Schools*. New York, NY: Crown.

Lee, V.E., & Burkham, D.T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.

Lortie, D.C. (2002). *Schoolteacher*. Chicago, IL: University of Chicago

Marzano, R.J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development

Meisels, S. J., Bickel, D. D., Nicholson, J., Xue, Y., & Atkins-Burnett, S. (2001). Trusting teachers' judgments: A validity study of a curriculum-embedded performance assessment in kindergarten to grade 3. *American Educational Research Journal*, 38(1), 73-95. doi: 10.3102/00028312038001073

Merton, R.K. (1957). *Social theory and social structure*. Glencoe, IL: Free.

National Center for Education Statistics (2007). *The condition of education*. Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubs2007/2007064.pdf

National Center for Education Statistics (n.d.*). ECLS-K base year data files and electronic codebook*. Washington, DC: National Center for Education Statistics. Retrieved from http://www.nber.org/~kling/surveys/Userguide.pdf

National Education Association (n.d.). *Continuous school improvement*. Retrieved from http://keysonline.org/

Oakes, J., Joseph, R., & Muir, K. (2004). Access and achievement in mathematics and science: Inequalities that endure and change. In J.A. Banks & C.A. M. Banks, *Handbook of research on multicultural education*. San Francisco, CA: Jossey-Bass

Page, R. (1987). Teachers' perceptions of students: The link between classrooms, school cultures and the social order. *Anthropology & Education Quarterly* 18, 77-99. doi: 10.1525/aeq.1987.18.2.04x0667q

Perry, N.E., & Meisels, S.J. (1996). *How accurate are teacher judgments of students' academic performance?* Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=9608

Ready, D.D., & Wright, D.L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: The role of child background and classroom context. *American Educational Research Journal*, 48(2), 335-360. doi: 10.3102/0002831210374874

Rosenthal, R.R. & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectations and pupils' intellectual development*. New York, NY: Holt, Rinehart and Winston

Smith, M.L., & Shepard, L.A. (1988). Kindergarten readiness and retention: A qualitative study of teachers' beliefs and practices. *American Educational Research Journal,* 25, 307-333. doi: 10.3102/00028312025003307

Tach, L.M., & Farkas, G. (2006). Learning-related behaviors, cognitive skills, and ability grouping when schooling begins. *Social Science Research* 35, 1048-1079. doi: 10.1016/j.ssresearch.2005.08.001