MEN AMONG BOYS: THE CHARACTERISTICS, QUALIFICATIONS AND ACADEMIC IMPACT OF MALE KINDERGARTEN TEACHERS IN AMERICA

Jason Rose

A dissertation proposal submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education.

Chapel Hill
2009

Approved by:
Lynne Vernon-Feagans, Ph.D.
Oscar Barbarin, Ph.D.
Leigh Hall, Ph.D.
Kirsten Kainz, Ph.D.
Judith Meece, Ph.D.

ABSTRACT

JASON ROSE: Men among boys: The characteristics, qualifications and academic impact of male kindergarten teachers in America
(Under the direction of Lynne Vernon-Feagans, Ph.D.)

This research examined the influence of teacher gender and teacher gender-related characteristics on student reading achievement during the kindergarten year. Using a nationally representative sample of male and female kindergarten teachers and their students collected as part of the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99, analytic methods were designed to address two specific issues. First, to consider whether male and female kindergarten teachers also differ significantly at the population level in terms of any other potentially important characteristics, such as gender-associated differences in demographic characteristics, educational qualifications, employment experience, or instructional practice tendencies. Second, to test whether teacher gender itself, or any such identified teacher gender-associated differences, are significantly associated with student achievement outcomes pertaining to reading growth over the kindergarten year. This was tested by incorporating teacher gender and gender-associated characteristic differences as covariates in a series of multilevel models designed to test each of these identified covariates for associations with student reading levels over the kindergarten year. Additional sample-restricted models were also tested to more specifically determine whether any identifiable aspect of having a male teacher could be supported as significantly benefiting kindergarten boys in terms of measurable reading achievement.

Male kindergarten teachers were found to be younger (on average), less experienced, to have less formal training in early childhood education, and to be more likely to teach in either rural or urban schools and in half-day instructional settings. Population trends between male and female kindergarten teachers were not significantly different for any other areas of demographic characteristics or qualifications, and there was no evidence of any significant differences in instructional practices. Teacher gender was also not found to be significantly associated with kindergarten spring reading levels for either boys or girls. The most significant predictors of spring reading levels in this data were fall reading score, time lapse between assessments, student gender, student race/ethnicity, family SES, class type, and resident father status. Teacher gender did not, either alone or in interaction with student gender or other student characteristics, demonstrate any significant association with student end of year reading levels.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figures

CHAPTER ONE

INTRODUCTION

Rationale

The study of gender differences in education has been a prolific area of research over the past several decades with regards to a number of specific issues, including gender-related trends in psychological development and cognitive performance tendencies (Halpern, 1997; Maccoby & Jacklin, 1974), gender-related socialization effects on school readiness (Fabes, Martin, Hanish, Anders & Martin-Derdich, 2003) gender-related differences in academic self-competence (Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002; Lummis & Stevenson, 1990), gender differences in achievement and academic choice among middle grades and high school children (Marsh, 1998; Meece, Parsons, Kaczala, & Goff, 1982; Schweigardt, Worrell & Hale, 2001), and differential treatment of boys and girls by teachers in schools (Altermatt, Jovanovic & Perry, 1998; Duffy, Warren & Walsh, 2002).

However it has only been comparatively recently that increasing scholarly attention has specifically been paid to issues facing boys in the early years of school (Weaver-Hightower, 2003). Boys have since been found to be beginning school behind girls in reading achievement, as measured by both standardized (Chatterji, 2006; Guarino, Hamilton, Lockwood, Rathbun & RAND, 2006; Ready, LoGerfo, Burkam &

Lee, 2005) and unstandardized (Condron, 2007; Entwisle, Alexander & Olson, 2007) measures of achievement, and continuing to fall further behind over subsequent years. Young boys have additionally been found to face higher rates of negative behavior referrals (Wehmeyer & Schwartz, 2001) and expulsions during the first few years of school (Gilliam, 2005).

Despite the promising upswing in interest regarding the state of these problematic population trends among young boys, a commensurate body of research proposing scientifically-based methods for addressing these problems has not developed as quickly. As a result, research has fallen behind in this area to a competing rhetoric from other outlets about what is often termed a "boy crisis" (Barnett & Rivers, 2007; Kimmel, 2006) in education. Major publications (Scelfo, 2007), popular bestsellers (Sommers, 2000) and privately funded policy reports (Johnson, 2008; Maine Boys Network, 2007) have been allowed to largely form public awareness about how best to address the problems facing boys in school, while education researchers simply have not yet gathered enough evidence to be able to either support or refute many of their claims.

One such popularly espoused but largely untested claim is that boys would particularly benefit from having more exposure to male teachers in the schools, especially at the predominantly female taught lower grade levels (Scelfo, 2007; Sommers, 2000; Maine Boys Network, 2007; Johnson, 2008). Though this idea has generated renewed interest in recent years, it is by no means a new suggestion. Calls for more male teachers at the lower grade levels have been made repeatedly by vocal proponents – in response to a variety of different issues perceived to be threatening the well-being of boys in school – since teaching first became a predominantly female profession in America around the mid-nineteenth century (Kimmel, 2006; Sexton, 1969; Sugg, 1978; Williams, 1995). Despite the ongoing popular appeal of this idea, it has generated surprisingly little in the way of rigorous research about male teachers in the lower grades or whether their presence might uniquely benefit boys in any measurable way.

This study was designed to begin addressing this issue by informing it in two important ways: (1) expanding what is known about the men who teach at the lower grades levels (in this case, specifically kindergarten) in relation to their female colleagues, and (2) contributing to what is known about the measurable influence of teacher gender on student achievement gains in the critical subject area of reading/language arts for young boys during the first year of school.

Statement of the Problem

The idea that exposure to a male teacher at an early age in school would be particularly beneficial for young boys is one that scholars evaluating gender research might categorize as "intuitively compelling" (Ruble, Martin & Berenbaum, 2006), meaning that its plausibility and appeal are apparent but just exactly how it should work is much less so. One of the most immediate problems for researchers interested in testing whether or not increased exposure to a male teacher influence in the early grades of school can uniquely benefit boys at this age is that claims rarely define exactly how men are proposed to exert this unique influence over the boys in their classroom. It is not clear, for example, whether men are expected to actually teach in ways different from women, or whether boys are just expected to imitate or respond more positively to men when they display the same teaching behaviors as women. Allan (1994) addresses this confusion in part when he describes a "role modeling paradigm" (11), in which men teaching at the lower grades levels find themselves uncertain as to whether they are supposed to be modeling traditional masculine behaviors (strong, bold, i.e. "hypermasculine") or gender-atypical behaviors *as* masculine (nurturing, sensitive, i.e. "hypomasculine") for the benefit of their boys. Others have similarly suggested that teachers, administrators and parents alike tend to speak enthusiastically about the importance of male teachers at the early grades levels, using vague terminology such as "positive influence" and "strong male figure", but struggle to define exactly what such statements mean when pressed to elaborate (Sargent, 2001).

Similarly, it is often equally unclear at the student level what the specific benefits of having a male teacher are expected to be for boys, or how or when they are expected to appear. There is no consistent answer, for example, about whether these benefits are expected to be tangible, such as short-term, measurable academic gains (Costello, 2008; King & Gurian, 2006a), or more intangible such as long-term shifts in behavior or aspirations (Gamble & Wilkins, 1997).

A third challenge facing researchers interested in learning more about dynamics associated with male teachers in the lower grades is that there are so few of them. At the kindergarten level, men constitute only 2% of all teachers nationwide, compared to 98% being women. That translates to approximately 4,198 male kindergarten teachers dispersed throughout a population of 187,428 kindergarten teachers across the country (U.S. Department of Education, National Center for Education Statistics [NCES], 2004). This fact presents a significant obstacle for researchers interested in studying this population using any type of randomized experimental design on a local or regional level. Given this reality, it is not surprising that much of the research available today on male teachers of young children consists of small-scale, qualitative accounts which offer rich, contextual information about a handful of men but ultimately have little or no power to represent any larger population.

The fact that there have been so few rigorous large-scale direct efforts to better inform our knowledge about these men or their potential influence on the outcomes of the boys they teach reveals another, more basic, problem for researchers to address which is that they are simply understudied populations.

Research Questions and Hypotheses

Research Questions

This study addresses these research obstacles in several specific ways: (1) by using data from a nationally representative sample of all kindergarten teachers to provide a detailed profile of the demographic characteristics, qualifications and instructional practices of kindergarten teachers in America by gender; (2) by separating out identifiable practical differences between male and female teachers in an effort to more clearly distinguish between possible pathways of influence between teacher gender and student outcomes, and (3) by identifying one specific, measurable outcome in need of address at the student level and testing it for significant associations with teacher gender and teacher gender-related characteristics.

Using data from the base year teacher and child catalogs of the Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K), this research was directed by the following guiding research questions (*RQ*s):

RQ1: Who are the male kindergarten teachers in America, and how are they alike or different from their female colleagues (in terms of the following specific characteristics)?

RQ1a. What are the comparative distributions of other available demographic characteristics for male and female kindergarten teachers in America, and are there any significant differences in the distribution of these characteristics across teacher gender groups?

RQ1b. What are the educational and employment related qualifications of male and female kindergarten teachers in America, and are there any significant differences in distributional tendencies of these practical qualifications between groups that remain significantly associated with gender itself after accounting for the potential influence of other known demographic characteristics?

RQ1c. Are specific methods of reading/language arts instruction used differently in the classroom by male and female kindergarten teachers in America (as measured by self-reported the frequency of uses of specific instructional methods), and are there any significant differences in the reported usages of these specific teaching methods between groups that remain significantly associated with gender itself after accounting for the potential influence of other known demographic characteristics?

RQ2: Is teacher gender significantly associated with end-of-year reading achievement levels for kindergarten students, and specifically for kindergarten boys, as measured by the following criteria:

RQ2a. Does teacher gender itself account for a significant proportion of between-classroom variance in the end of year student reading achievement levels among all students after controlling for other known covariates, including kindergarten entry levels and any practical teacher gender-related differences identified by the results of *RQ1*?

RQ2b. Does teacher gender account for a significant proportion of between-classroom variance in the end of year student reading achievement levels of boys specifically, either by gender alone or in interaction with the additional student level variables of resident father status and behavior services status?

Working Hypotheses

The analytic steps detailed in Chapter 3 all begin working from the null hypothesis that no significant relationships between teacher gender and any given dependent variable will exist. This places the burden on the data to sufficiently disprove the null hypothesis at every step. However based on the review of existing literature presented in the next chapter, the following working hypotheses (*WH*) are proposed:

WH 1: Male kindergarten teachers are expected to be significantly different from their female colleagues in terms of the following specific characteristics:

WH1a. Male kindergarten teachers are expected to differ significantly from female kindergarten teachers on the demographic characteristics variable of *age*, with the male teachers being slightly older on average.

WH1b. Male kindergarten teachers are expected to differ significantly from female kindergarten teachers on the qualification variables of *total number of courses taken in early childhood*, *total number of courses taken in elementary education* and *total number of courses taken in reading instruction methods,* even after controlling for other demographic characteristics, with the male teachers reporting lower totals in each case.

WH1c. Male kindergarten teachers are expected to differ significantly from female kindergarten teachers on the instructional practice variables of *didactic instruction* and *student-centered instruction methods*, even after controlling for other demographic characteristics, with the male teachers reporting a higher frequency of use of didactic instruction and lower frequency of use of student-centered instruction.

WH2: Having a male teacher is expected to show a small, significant positive association with student reading achievement scores for kindergarten boys in the following ways:

WH2a. Teacher gender is not expected to account for a significant proportion of between-classroom variance in the end of year student reading achievement levels among all students after controlling for other known covariates including kindergarten entry levels and any practical differences between teachers known to vary significantly by teacher gender from the results of *RQ1*.

WH2b. Teacher gender is expected to account for a small but significant proportion of between-classroom variance in the end of year student reading achievement levels of boys alone when considered in interaction with additional student level variables of resident father status and behavior services status.

Definition of Terms

Akaike information criterion (AIC): The Akaike information criterion (AIC) is a measure of goodness of fit useful for comparing multiple statistical models and selecting the best available one. When comparing multiple statistical models by this criterion, the lowest AIC estimate indicates the best available model.

Base year: In reference to the ECLS-K data source used in this study, as well as findings in existing studies referenced in Chapter 2, the term "base year" refers to the first year of data collected for the Early Childhood Longitudinal Study – Kindergarten Cohort Class of 1998-99. ECLS-K base year data includes a nationally representative sample of students and teachers at the kindergarten grade level during the 1998/1999 academic year.

Instructional Practice(s): Term used in this study to refer collectively to a set of composite variables measuring teacher-reported use of specific instructional methods for teaching reading and language arts in the data. Instructional practice variables used in this study are derived from a composite of individual ECLS-K teacher questionnaire items identified by factor analytic methods in previous research (Hamilton & Guarino, 2004) to reliably measure use of the following seven instructional practices by teachers in the ECLS-K base year dataset, and include: *reading and writing activities*, *phonics*, *didactic instruction*, *comprehension*, *student-centered instruction*, *reading and writing skills*, and *mixed-ability grouping*. For a full list of individual teacher questionnaire items comprising each composite score, see Table 1.1.

Intraclass correlation (ICC): Measures the proportion of variance in the outcome that is between groups [in a multilevel regression model]. The ICC is also sometimes called the "cluster effect" (Raudenbush & Bryk, 2002, p.36).

IRT scale score: Item response theory scale score. ECLS-K direct child assessments utilized a two-stage assessment design, the first stage consisting of a 12 to 20 item routing test which determined the appropriate difficulty level of the second stage form administered. IRT scale scores are designed to allow for direct comparison of assessment scores regardless of which form children were administered by creating a common scale based on the items administered in the routing test as well as a core set of items shared on the second stage forms. Item response theory analyzes the pattern of correct, incorrect and omitted responses from each individual child's assessment using a model designed to account for the "difficulty, discriminating ability, and 'guess-ability'" (U.S. Department of Education, National Center for Education Statistics, 2001b, p. 3-2) of each item in order to determine the placement of each child's score along the common scale. Common scale scores produced by IRT are directly comparable across children because they represent an approximation of the scores each child would have received had all test items on all forms been administered to all children, while guarding against the possibility of score distortion by guessing or chance. IRT scale scores are considered the most appropriate scoring option for comparing student growth over time when using the ECLS-K direct cognitive assessment (NCES, 2001b).

Likelihood ratio test: The likelihood ratio test, as reported here by the “xtmixed” procedure [in Stata], is an overall test of the covariance parameters associated with all random effects in the model. In models with a single random effect … it is appropriate to use this test to decide if that random effect should be included in the model. (West, Welch, & Gałecki, 2007, pp. 85, 145).

Multiple imputation: Multiple imputation procedures for replacing missing data, “produces *M* complete data sets. Each data set is analyzed via OLS and the results are averaged to arrive at a point estimate of each regression coefficient. … in addition to producing realistic standard errors, multiple model-based imputation will produce unbiased inferences about the parameters generating the complete data as long as the model assumptions hold and the data ‘missing at random’ (MAR)” (Raudenbush & Bryk, 2002, p.338).

Taylor series method: In reference to the specific type of sample weighting procedure used in this study to produce population estimates from the sample data. The ECLS-K Base Year Public-Use data files provide independent full sample weights for child, teacher, and school level data, as well as the option to compute standard errors at each level of data by replication method or Taylor series method. The Taylor series method produces a linear approximation to the survey estimate of interest and is recommended by ECLS-K specifications as the option designed for compatibility with the software package Stata (NCES, 2001b).

Chapter Organization

Chapter 1 has been presented as an introduction to the current study including a rationale, a statement of the problem, a summary of the proposed research questions and working hypotheses, and a definition of key terms used throughout the following chapters. Chapter 2 contains a review of literature designed to present an organized consideration of existing knowledge about the specific populations and variables of interest to this study. Chapter 3 contains an explanation of the methodologies used to pursue answers to each of the guiding research questions for this study, including details about the data sources used and analytic methods employed at each step. Chapter 4 presents a straightforward explanation of the findings, and Chapter 5 concludes by considering these findings more thoroughly with regards to existing literature and implications for future research.

CHAPTER TWO

REVIEW OF LITERATURE

The information presented in this chapter is divided into four major sections: (1) a consideration of the broader theoretical context within which this study is situated, (2) an explanation of the scope and focus of literature presented in this chapter, (3) literature pertaining to the teacher level issues associated with *RQ1* and informing the working hypotheses associated with it, and (4) literature pertaining to the teacher and student level issues associated with *RQ2* and informing the working hypotheses associated with it.

Theoretical context

This study is situated within a larger discourse about the lack of male teachers at the early and middle childhood grade levels in America and whether that imbalance is disadvantaging to boys in any way. In order to contribute meaningfully to that discourse, it is first necessary to define the purpose of this study in terms relative to existing theory and knowledge about these issues. Because rigorous academic research directed specifically at this question is sparse, it is also important to acknowledge that participation in this particular discourse – at the current stage of what is known about gender differences associated with male teachers and students – necessarily involves also addressing this study’s goals with respect to a competing body of claims frequently put forward as knowledge in popular media outlets or other public forums.

The study of male teachers and male students at the lower grade levels constitutes a highly specific subfield of gender studies in education, which itself is subfield of gender studies in general. As such, researchers and other constituents often seek to understand gender differences at the school level as specifically applied cases of gender differences in society at large. Efforts to explain the source of any such education-related differences between males and females can, more often than not, be categorized fairly clearly as leaning to one side or the other of the more fundamental issue of “nature” vs. “nurture” in gender studies at large. Sabbe and Aelterman (2007) demarcate this corollary division in the current state of education research by assigning the terms “sex differences research” and “gender dynamics research” as the two theoretical umbrellas under which most traditional and current studies of gender in education can be categorized.

Sex differences research, as defined by Sabbe and Aelterman, consists of those studies which focus specifically on identifying observable differences between male and female teachers or students with respect to one or more specified outcome(s) or trait(s). A hallmark of sex differences research is a tendency to focus primarily on establishing the existence of such differences at the population level, but to largely or completely stop short of exploring the possibility of any social or situational gender role influences as further explanation of their findings. Though the authors note that sex differences researchers rarely articulate an overt endorsement of any specific theoretical explanation for differences found, a distinct degree of biological determinism is considered to be implied in that findings are generally “attributed to being a male or female and are presented as inherent characteristics of the group ‘men’ or the group ‘women’” (p. 524).

Methodologically, sex differences research tends to be comprised of studies involving large-sample data and more quantitative, instrumented analysis.

Gender dynamics research, alternately, is defined as consisting of those studies which focus on the role that gender (i.e., the socially-constructed identity, external expectations and other influences associated with biological sex characteristics) plays in decision-making or interpreting experience among teachers, students, or other educational constituents. The hallmark of this type of research is a more intimate study of the subjective experiences and perceived social determinants of individual agency governing the decisions made by research subjects, but with little or no additional effort to prove the broader generalizability of their findings beyond the specific participants in their research. Unlike the generally implied theoretical undertones of sex differences research, gender dynamics research is characterized as more likely to espouse an explicit theoretical perspective in advance and then to seek and interpret data exclusively through that lens. Methodologically, gender differences research tends to be comprised of studies involving smaller-sample data and more qualitative, interpretive analysis.

While these distinctions outlined by Sabbe and Aelterman provide useful a guide for moving forward and interpreting what is currently known or claimed about the situation of male teachers and students at the lower grade levels, as well as for situating the current study appropriately within this larger discourse, it should also be recognized that there are presumptions upon which these categories are built which have the potential to be misleading if not acknowledged. The most directly relevant of these in application to the current study being a general implication that the focus and structure of most research on gender issues in education is primarily reflective of the theoretical framework and intentions of the researcher(s) themselves. In some cases, the focus and structure of a study is likely to be determined more significantly by the question being asked and the data available answer it with.

For example, the majority of current studies specifically focusing on male teachers of young children are small-scale, qualitative investigations which would clearly fall within the parameters of gender dynamics research (Allan, 1994; King, 1998; Sargent, 2001; Williams, 1995), however it is likely that this is at least in part due to the fact that male teachers of young children are simply too rare in any particular geographic region to try to construct a sample with the kind of generalizability preferred in sex differences research. Conversely, the current study and several others cited in this chapter (Palardy & Rumberger, 2008; Strizek, Pittsonberger, Riordan, Lyter & Orlofsky, 2006; Xue & Meisels, 2004) would by all definitions outlined above be best described as sex differences research, but importantly share the common characteristic of using a pre-existing dataset in their analysis. In such cases, it is important to emphasize that by using a pre-existing dataset to answer population-level questions, a researcher is also limited in terms of the information available for interpretation of any resultant findings to the information collected in the original instrument.

This caveat is particularly important in understanding the theoretical framework of this study. Guided by the research questions detailed in Chapter 1, the goals of this study were to identify what *are* the significant differences between male and female kindergarten teachers and what *is* the significance of teacher gender or teacher gender-related characteristics for student achievement gains, as defined. However in exchange for the statistical power to pursue these answers at the population-level, it was accepted that the data would not be sufficiently capable of definitively also explaining *why*. For example, several studies of male teachers of young children have concluded that such men either are (Williams, 1995) or legitimately believe they are (Cognard-Black, 2004; King, 1998; Sargent, 2001) under constant suspicion regarding their physical interactions with students, and that the danger of a misunderstanding or false accusation is enough to actually force them to teach in more structured, didactic, distant ways (Cognard-Black, 2004; King, 1998; Sargent, 2001). While the consensus opinion of findings such as these strongly informed the working hypotheses of this study, the findings here will only be able to conclude whether or not male kindergarten teachers in fact *do* incorporate more didactic instructional practices than their female colleagues, but will not be able to confirm or deny what motivations or other explanatory factors may be underlying such a difference.

Still, the fact that some studies, due to the practical realities of research, do not fit perfectly within the categorical parameters of the sex differences vs. gender dynamics framework does not diminish the importance of those studies, nor does it necessarily diminish the usefulness of this categorical framework as an interpretive guide to the larger discussion. Such a frame of reference is particularly useful when applied to the claims of some popular media or private foundation outlets (such as those cited in Chapter 1) that have traditionally driven much of the public dialogue about male teachers and their importance to the success of male students. By doing so, some claims put forth by these outlets may be more easily identified as problematic for the ways they often tend to cite small-scale or individual case examples in support of general sex difference conclusions, or alternately apply population-level sex difference trends to expectations about individual motivations.

Academic research also benefits from the ability to organize current knowledge in this type of encompassing framework. For one thing it can help illuminate the fact that much of the specific knowledge cultivated by either camp has traditionally remained insulated from the perspective of the other. It is clear now that it will take a more concerted effort to bridge the knowledge base between sex difference and gender dynamics researchers if academic researchers, as an entity, are to effectively regain an influential stake in a public dialogue in which legitimate (e.g., concerns about unrecognized gender bias in over-referral of boys/under-referral of girls for out of classroom services (McIntyre, 1988)) , manufactured (e.g., concerns about elevated rates of homosexuality or pedophilia among male teachers of young children (Allan, 1994; Sargent, 2001) or about boys becoming either alienated from or "feminized" by school early on due to composition of the teaching force and the general "femininity of schools" (Biddulph, 2002; Sexton, 1969)), and insufficiently examined (e.g., male teachers provide unique benefits for male students in terms of behavior, motivation and/or achievement outcomes (Dee, 2006b)) concerns have become problematically intertwined.

With all of this in mind, the current study should be considered an investigation of the nature and interaction of specific sex differences between kindergarten teachers and kindergarten students, though one that is informed by current literature about both sex differences and gender dynamics, and that is ultimately intended to contribute to a more collaborative dialogue among academic researchers from both traditions to foster a more informed dialogue between research and the public.

Scope of Literature Review

In considering the evidence presented in this chapter, there are a few acknowledgements to be made regarding the scope of information included. The primary populations of interest in this study were male kindergarten teachers in America and their male students. Because existing studies based exclusively on members of either of these populations are too rare to explicitly inform all of the research questions posed in Chapter 1, it was necessary to make certain decisions about expanding the scope of information considered relevant. The first way this was done was to expand the grade level parameters to allow for studies which included male teachers and/or their male students between the levels of preschool and fifth grade, with primary emphasis given to studies focusing on kindergarten or grade levels below third grade whenever possible. Some research addressing issues of teacher gender differences at the middle and high school levels was also included under two specific conditions: (1) in consideration of a particularly germane theory which may have been originally postulated at or tested across multiple grade levels, or (2) in consideration of existing knowledge about the demographic characteristics, qualifications or instructional practice tendencies of male teachers in cases for which no other information was specifically available regarding male teachers at the lower grades levels alone. In instances such as these when information presented reflects findings from a broader range of or the entire (K-12) teacher population, rather than specifically from teachers at the kindergarten or lower grade levels alone, this distinction is clearly noted.

It is also important to note that the application of these expanded search parameters pertains only to research regarding the teacher level variables. Information presented with reference to gender differences in achievement among students remains restricted to findings at the elementary and early childhood levels only. This more restricted focus was necessary due to the wide variability known to occur in the specific nature of gender achievement differences across grades as students progress through school (American Association of University Women [AAUW], 2001; Ruble, Martin & Berenbaum, 2006).

For some issues about which there was particularly little or no directly applicable current empirical evidence to draw on, specifically relevant theoretical or historical perspectives were also considered. For example, in the absence of specifically available population demographics indicating the racial distribution of male kindergarten or early elementary grades teachers in recent years to use as an indication, historical research documenting participation trends in teaching among African American (Rury, 1989) and European American (Carter, 1989) teachers spanning several decades was used in an attempt to inform expectations of the current data prior to analysis. While such historical perspectives were used only as necessary here, and with respect to specific variables of interest, a more comprehensive background in the social, political, economic and religious influences that have reshaped the general landscape of teaching over the past century and a half (from what was once widely considered a male-oriented profession to what is now generally considered a female-oriented one) would be helpful for anyone interested in putting the findings of this study into greater perspective, but is also beyond the reasonable scope of this chapter or exploratory potential of this dataset. For a more thoroughly detailed chronology of this transformation, which scholars commonly refer to as the “feminization of education”, sources such as Blount (2000), Enoch (2008), Rury

(1989), Sedlak & Schlossman (1986), Sugg (1978), and Warren (1989) all provide excellent historical accounts.

A final consideration to note is that because male kindergarten teachers and students are the two primary populations of interest in this research, both the literature review information gathered here and the findings presented in subsequent chapters are presented reflecting that focus. This perspective is in no way intended to dismiss the broader acknowledgement that for nearly every issue facing male teachers or male students presented in this chapter, there is likely to be an equal (if not directly reciprocal) issue to consider facing female teachers and students. For example, if men are overly discouraged from considering teaching or working with young children as an appropriate career choice, so are women likely to be overly encouraged to move into those jobs (Williams, 1995). If young boys are more at-risk for negative outcomes associated with being labeled a behavior problem in school, so too might young girls be more at risk for negative long-term outcomes associated with over-identifying with the role of a docile, passive learner (Fennema & Peterson, 1985). These and other such issues are acknowledged in advance as being equally deserving of further examination, but remain presently beyond the scope of this research.

RQ1: Gender-related differences between male and female teachers

Gender

Longitudinal trends in the gender composition of American public school teachers over the last near century and a half are well documented at the aggregate (across all grades) level. Between 1869 and 2001, the percentage of all teachers who were male ranged from a high of 42.8% just before the end of the nineteenth century to a low of 14.1% around 1920 (Snyder, Dillow & Hoffman, 2008). Historical accounts of issues affecting public education policy during these times add helpful perspective to these statistics which, upon first impression, appear to have fluctuated dramatically over this time period. Such accounts explain that the transfer of the teaching profession from a predominantly male to a predominantly female occupation only began to take hold around the mid-nineteenth century at the lower grades levels (Gamble & Wilkins, 1997; Sugg, 1978; Rury, 1989), and then accelerated during and just after World War I, when more women were needed in the upper grades levels to replace men who had either gone off to war and never returned, or who had returned from the war but not to teaching (Williams, 1995). Aside from a few periods of major economic upheaval or other change, the percentage of male teachers over the last century has actually experienced only minor fluctuations, remaining generally between 20 and 35% of the total teacher population (Snyder, Dillow & Hoffman, 2008). By 2000 the percentage of male teachers making up the total teacher population was dropping closer to the lower end of that range at 25.1%, though a closer look revealed that male teachers among new hires (those within their first three years of teaching) were slightly on the rise (Shen, Wegenke & Cooley, 2003).

Among elementary school teachers (K-5) alone, the percentage of male teachers in most recent estimates drops somewhat lower to 16.2% (Strizek, Pittsonberger, Riordan, Lyter, & Orlofsky, 2006). And as of 1998, the percentage of males among all teachers at the kindergarten level alone was just 2% (NCES, 2004).

Age/Experience

Average age among all male teachers has increased steadily between 1961 (34 years) and 2001 (47 years), while the average age for females has fluctuated more, beginning at 46 years in 1961, dropping as low as 33 years in 1976, and steadily rising up again to an average of 45 years in 2001 (Snyder, Dillow & Hoffman, 2008). A steady rise in the median years of teaching experience from 8 years in 1976 to 14 years in 2001 (Snyder, Dillow & Hoffman, 2008) would seem to suggest that the increase in average age over this time period was at least in part due to teachers remaining in the job longer. At last report, the average age of male and female kindergarten teachers combined was about 41 years old (NCES, 2004). No disaggregated data identifying individual trends in either age or years of experience by gender among kindergarten teachers alone was officially available prior to this study.

Much of what may be additionally discerned about age and experience differences between male and female teacher at the lower grades levels comes from trends identified in multiple smaller-scale studies. At the elementary grades levels, male teachers have been found to be to be slightly older on average than their female peers (Smith, Kirchner, Taylor, Hoffman & Lemke, 1998), though it is unclear whether the age difference in this study is associated as well with years of experience. Another potential indicator is that men who teach in the elementary grades have been found to enter teaching relatively later in life than their female colleagues (Brookhart and Loadman, 1996; Gross & Trask, 1976). One reason for this may be that male teachers often report having spent a considerable amount of time pursuing other career options before deciding to enter teaching (DeCorse & Vogtle, 1997; Freidus, 1992; Sargent, 2001). Findings such as these may offer some indication about trends to expect among men teaching at the elementary or lower elementary grades levels, though no similar studies using a controlled, nationally representative sample are currently available.

Race/Ethnicity

Racial demographics of public school teachers overall have remained fairly consistent over the past few decades at about 90% White, 7-8% African American and 2-3% combined among other identifications (Shen, Wegenke & Cooley, 2003). Trends in racial demographics prior to the mid-twentieth century are more difficult to follow for a number of reasons, including the disadvantaged legal and social status of African Americans and subsequently other minorities throughout much of the country's history; and the purposefully clandestine nature of organized education among African American communities in the South before the Civil War (Perkins, 1989; Perlmann & Margo, 2001).

In examining what information is available, Rury (1989) notes that the participation of African American men and women in teaching across all grade levels was close to equal around the beginning of the twentieth century, and then rapidly shifted over the first half of the century to a point where African American women outnumbered African American men in teaching approximately 4 to 1. He suggests that "teaching, particularly at the elementary level, had become sex-typed as women's work " (1989, 36) among both blacks and whites. Still others have argued that, when only considering within-gender group comparisons of men (African American men or other non-white, racial minority men vs. European American men), a teaching career may have historically been viewed by men of minority populations as a more favorable opportunity for social and financial gain than it would have been for white males (Carter, 1989) and that teaching (across all grade levels), appears to have traditionally drawn a relatively larger percentage of men from lower and working class backgrounds and minority populations (Lemkau, 1984). However trends across all grade levels cannot be presumed to apply directly to specific grade levels, nor do historic trends necessarily provide accurate indications about such distributions today.

Considering the kindergarten level alone, racial demographics today appear only slightly more distributed than at the aggregate level, at approximately 84% White, non-Hispanic; 6% Black, non-Hispanic; 6% Hispanic; 3% Other, non-Hispanic (NCES, 2004). There is no indication from these numbers as to whether distributions among male kindergarten teachers alone proportionally reflect these aggregate distributions. Because male kindergarten teachers, or male teachers at any of the lower grades levels, are so relatively rare in the population, small-scale studies focusing on these men have been unable to place any type randomized controls on their selection of samples. As a result, many of these studies do not even report in any formal way the race/ethnicity distributions within their own sample, nor is such information useful for projecting anything about the larger population when they do.

Location

Considering again teachers across all grade levels (K-12), recent estimates place male representation ranging from a low of 16.7% of the teacher population in South Carolina to a high of 33% in Wyoming. Male teachers are relatively evenly distributed between urban (24.5%), suburban (24.6%), and rural (26.9%) areas throughout the United States (Shen, Wegenke & Cooley, 2003). Findings from this same source also estimate that men make up a slightly more significant percentage of teachers at schools where minority enrollment was below 19% of the student population than in schools where minority enrollment was over 20% (Shen, Wegenke & Cooley, 2003).

At the kindergarten level alone, the largest percentage of all teachers combined by far are located in suburbs/large towns (40%), with other areas reporting a range between small towns (9%), rural (14%), large cities (18%) and midsize cities (19%). Kindergarten teachers also appear to be evenly divided between those at lower poverty schools (51%) and those at higher poverty schools (49%) (NCES, 2004). There is currently no information available to indicate whether the physical distribution of male kindergarten teachers alone is consistent or inconsistent with any of these aggregate patterns.

Education/Certification

Education level and amount of academic coursework have been found to have positive associations with teacher use of developmentally appropriate practices with young students (McCullen & Kazat, 2002) and student academic outcomes for older students (Wayne & Youngs, 2003) in prior research, so a closer examination of the achieved education and coursework levels of the men teaching kindergarten is an important area of information to better understand.

Distributions of highest education levels achieved are relatively consistent between men and women across all grade levels. At the ends of the spectrum, there are slightly more men than women who report holding less than a bachelor's degree as well as who report holding a doctoral degree. In between, there are more women than men who report holding a bachelor's degree, master's degree, or education specialist (Snyder, Dillow & Hoffman, 2008).

Among elementary grades teachers, men have been reported to hold slightly higher levels of post-graduate education than their female peers in other studies (Smith, Kirchner, Taylor, Hoffman & Lemke, 1998), though there is no indication of whether these findings were affected by any outlier men on the extreme upper end. Curiously, male elementary school teachers have also been reported elsewhere to be less academically-oriented, and less optimistic about the usefulness of coursework than their female colleagues (Brookhart & Loadman, 1996). One possible explanation for this potential contradiction between men's reported higher levels of post-graduate education in one study and reported lower regard for the usefulness of their academic coursework in another may be offered by the previously noted findings that male teachers at the elementary grades levels often report entering teaching after having pursued other career options (DeCorse & Vogtle, 1997; Freidus, 1992; Poll, 1979; Sargent, 2001). Consider that the respondents in Sargent's (2001) study of male early elementary grades teachers held college degrees in a wide variety of stereotypically male-dominated fields, including architecture, engineering, math, computer science, physical education, biology and chemistry.

At the kindergarten level alone, achieved education levels across all teachers ranged between: less than bachelor's degree (2%); bachelor's degree (63%); master's degree (29%); and education specialist or doctoral degree (6%). 84% of these kindergarten teachers reported being certified in elementary education, while 53% reported being certified in early childhood education and 5% were not certified in either area (NCES, 2004). Again however, it is not clear from this source whether distributions among male kindergarten teachers alone generally mirror these overall distributions.

Instructional practices

Teacher instructional practices at the elementary and lower grades levels have been well documented to vary highly from class to class and to be significantly associated with student achievement (National Institute of Child Health and Human Development [NICHD], 2002; Pianta, LaParo, Payne, Cox & Bradley, 2002; Wright, Horn & Sanders, 1997). Variations in instructional practice are known to be associated with a myriad of potential influences, including: class size, grade level, number of students with disabilities, number of students receiving free or reduced lunch, teacher's perceptions about their own influence, teacher's academic background (Buchanan, Burts, Bidner, White and Charlesworth, 1998), student conduct, social class, cumulative folder information (Dusek & Joseph, 1983), and interactions of both race and gender dynamics between students and teachers (Grant, 1985; Dee, 2005).

In considering studies regarding gender-related differences in teacher instructional practices, it is important to acknowledge that there has been a tendency among researchers in this area to not always distinguish clearly between specific instruction-oriented practices and general classroom interaction styles. While recognizing that there has traditionally been some overlap in the way that the two are operationalized at the lower grade levels, there has still been some degree of significant evidence that male and female teachers at these levels may demonstrate significant differences in both specific instructional methods (Good, Sikes & Brophy, 1973; Brophy, 1985; Freidus, 1992), and general classroom interaction styles (Fagot, 1981; Stake & Katz, 1982).

Among elementary grade teachers, male teachers have been found in specific cases to lecture more, be more direct and subject-centered during lessons, ask more follow up questions after incorrect student responses and provide less feedback after correct responses than female teachers. Female teachers in the same studies were found to be more interactive in style, offer more praise, and generally create environments in which students reported feeling more comfortable guessing answers when they were not sure (Brophy, 1985; Freidus, Good, Sikes & Brophy, 1973; 1992; Stake & Katz, 1982).

Somewhat contrastingly, at the Pre-K level, male teachers have also been found in at least one study to express more favorable comments about students, interact with children more at play time or recess, and give students more physical affection than their female counterparts (Fagot, 1981). In discussing the potential implications of these findings, Fagot (1981) states in summary that her answer to the question of whether male and female teachers actually differ in their teaching styles would be a “qualified yes.” The qualification she adds being that even though male and female teachers do differ significantly in some specific ways, they are overall more alike than different and that many differences that may seem attributable on the surface to gender characteristics (in her terms) are more likely to be attributable to background characteristics not intrinsically related to sex per se, but to gender group socialization differences. She suggests in her discussion that identifiable differences between men and women who teach young children are likely also due in part to a greater degree of selection effect among males teaching at this level than females. The general contention that male and female teachers are more alike than different overall has also been regularly supported in other reviews of teacher gender research (Sabbe & Aelterman, 2007).

Still, Fagot’s findings that male Pre-K teachers were more likely to give affection to students than female teachers stands in relatively stark contrast to more recent findings in direct studies of men who teach at the early grades levels. In response to questions about ways that their own teaching practices differ from those of the women they work with, the most pronounced differences that men in several studies seemed to be aware of were those associated with the need to avoid physical contact with their students (Allan, 1994; DeCorse & Vogtle, 1997; King, 1998; Sargent, 2001; Williams, 1995). Across these studies, there was no disagreement about this issue among any of the participants: these men unanimously agreed that they cannot be in physical contact with their students in the same ways that women can be and, if possible, that it is better to try not to touch their students at all. Blount (2000) suggests that these fears trace back to issues of social intolerance and ignorance that have historically conflated men who work in nontraditional occupations with homosexuality, and homosexuality with pedophilia. Some participants specifically acknowledged that this need to avoid contact with their students at all times resulted in their feeling forced to "teach differently than the women" (Sargent, 2001). Williams (1995) agrees, noting from her own direct observations that the male elementary school teachers she studied were confronted by suspicions of pedophilia and that their awareness of these suspicions often caused them to alter their work behavior to guard against sexual abuse charges. Despite Fagot's (1981) findings, the majority of more recent evidence would seem to indicate that any gender-related teaching differences, if they do occur, are more likely to be in the direction of increased didactic and structured instruction from male teachers in order to maintain these types of safe boundaries in the classroom.

A more explicit consideration of whether male and female teachers do tend to use different types of instructional practices along such lines could be particularly informative in addressing the early reading gaps facing boys. Considering that structured, direct instruction has been shown to be particularly necessary for struggling early readers in general (Foorman, Francis, Fletcher, Mehta & Schatschneider, 1998). A number of recent ECLS-K based studies have found significant associations between instructional practice and student achievement gains. Guarino, Hamilton, Lockwood, Rathbun & RAND (2006) , for one, found positive associations between increased teacher use of the instructional practice methods of writing skills, didactic instruction, phonics, and reading and writing activities and student reading gains. Extending the research across both kindergarten and first grade, Palardy & Rumberger (2008) similarly found significant positive associations between teacher reports of reading instruction, phonics instruction, silent reading, and writing from dictation and student reading gains. Xue & Meisels (2004) found that increased teachers use of integrated language arts and phonics instruction was positively associated with higher classroom mean achievement scores in direct cognitive assessments of language and literacy as well as indirect teacher ratings of achievement, though the impact of integrated language arts instruction was not as beneficial for students with lower school entry achievement levels.

Such studies all provide confirmation that significant variations in teacher instructional practices do exist across individual teachers and are significantly related to student achievement, however no similar studies have yet specifically addressed these variations by both teacher and student gender.

RQ2: Teacher gender and student achievement

Student achievement outcomes are well established to be associated with a complex combination of factors. Among some of the many variables known to have considerable influence are: poverty/SES status (Entwisle & Alexander, 1992; Arnold & Doctoroff, 2003), race/ethnicity (Barbarin, 2002, Hedges & Nowell, 1999), home learning environment and preschool attendance (Dubow & Ippolito, 1994; Vernon-Feagans, 1996; Christian, Morrison & Bryant, 1998), and school environment (Kainz & Vernon-Feagans, 2007). Gender, at both the student and teacher levels, has also been found to be related to achievement outcomes in a number of ways that may be either unique or part of an interaction with several other influences.

Student gender and achievement

Over the past few decades, the most commonly reaffirmed findings pertaining to gender differences and student achievement at the lower grades levels have been that: (1) girls generally demonstrate greater verbal ability than boys, (2) boys generally demonstrate greater visual-spatial ability, (3) boys generally demonstrate better performance in mathematics (Maccoby & Jacklin, 1974; Jacobson, 2001; U.S. Department of Education, National Center for Education Statistics [NCES], 2001a; Ruble, Martin & Berenbaum, 2006). More recently findings have indicated promising developments with regards to closing the gap between boys and girls in general mathematics achievement (Hyde, Lindberg, Linn, Ellis, & Williams, 2008). Yet there remains general consistency across studies for significant differences in measurements of language learning favoring girls at the younger grades levels (Chatterji, 2006; Guarino, Hamilton, Lockwood & Rathbun, 2006; Lummis & Stephenson, 1990; Ready, LoGerfo, Burkam & Lee, 2005), though it should also be emphasized that gender differences in both language and mathematics domains are known to range greatly across age-levels and across specific sub-skills within each domain (AAUW, 2001; Ruble, Martin & Berenbaum, 2006).

Nevertheless, when considering student outcomes during the first few years of school, a number of recent studies have continued to reaffirm evidence of reading-related achievement gaps favoring girls in the national population (NCES, 2004a). Guarino, Hamilton, Lockwood, Rathbun & RAND (2006) found that kindergarten girls made greater single-year achievement score gains than kindergarten boys in all reading skills areas measured. Using the kindergarten sample alone, Ready, LoGerfo, Burkam & Lee (2005) found significant support showing that girls entered kindergarten with stronger literacy skills and gained more than boys by the end of the year even after controlling for initial difference levels. Chatterji (2006) measured the same outcomes across kindergarten and first grade and confirmed the earlier conclusions that boys scored significantly lower than girls at kindergarten entry, while extending findings to conclude that boys continued falling further behind to the end of first grade even after controlling for the other significant correlates of race and poverty status identified in the data.

Similar or related patterns have been reported in studies measuring achievement by alternate instruments than standardized assessment. Boys have been found to receive lower grades than girls in reading as early as first grade (Entwisle, Alexander & Olson, 2007), to be placed in lower-ability reading groups (Condron, 2007), and to suffer in achievement outcomes related to negative teacher academic expectations (Bennett, Gottesman, Rock & Cerullo, 1993), increased out of class referrals for learning or behavior disability issues (Wehmeyer & Schwartz, 2001), and increased rates of expulsion (Gilliam, 2005).

Interactions between teacher and student gender

Whether or not teacher gender per se is related to these student gender differences is more difficult to consider. Earlier research testing for relationships between teacher gender and student academic outcomes was generally unable to identify any direct connection (Gold & Reis, 1982; Good, Sikes & Brophy, 1973). However with renewed interest has come new evidence in support of the idea that teacher gender may interact with student gender to influence different outcomes, specifically including positive academic (Dee, 2005; 2006a; 2006b) and social (Brutsaert & Bracke, 1994; Mancus, 1992) outcomes for boys taught by male teachers. Though certainly intriguing, the number of such findings are still extremely limited overall and any connections between teacher gender and student outcomes remain in need of much further testing before any larger conclusions should be drawn. In order to foster such inquiry, there first needs to be a clearer conceptualization of how any such proposed influence might actually work.

Toward this end, Dee (2005) offers a useful distinction between teacher effects which are proposed to be directly triggered by biological/physical/demographic traits, such as sex or race, ("passive teacher effects") and those teacher effects proposed to be more indirectly exercised through behaviors and interaction styles which may be socially or culturally connected to the physical traits, such as gender, ("active teacher effects"). Passive teacher effects are defined as being simply engaged by a demographic likeness between teacher and student (such as the role modeling influence male teachers are often presumed to naturally exert over boys at school), while active teacher effects are defined as those which operate through unintended biases played out in expectations and interactions between teachers and demographically similar or dissimilar students (Dee, 2005, 159). These distinctions turn out to be especially useful in that they function similarly to "sex differences" vs. "gender dynamics" framework of Sabbe and Aelterman considered earlier, here in application to interpreting teacher effects on students.

Passive effects

Among findings that may be considered evidence of passive teacher effects, the one most commonly associated with male teachers is that they provide a positive role models for young boys (Allan, 1994; Sargent, 2001; Sexton, 1969; Sommers, 1990). It is important to underscore the term "considered" here, because it is often unclear in claims whether the process is expected to involve any active participation on the part of the teacher or student, or to just exist as some sort of innate reaction that is triggered when young boys are in the presence of men. Either way, hard evidence for whether or not boys actually respond more positively in some way to male teachers is somewhat limited and decidedly mixed (Carrington, Francis, Hutchings, Skelton, Read & Hall, 2007; Brutsaert & Bracke, 1994; Mancus, 1992).

At the teacher level, at least, there is a noteworthy amount of evidence suggesting that role modeling expectations alone may actually influence male teacher behavior in a much more active way than was previously thought. Male teachers in the lower grade levels appear to be extremely cognizant of their expected role-modeling responsibilities (Allan, 1994; Carrington, Francis, Hutchings, Skelton, Read & Hall, 2007; Sargent, 2001). In direct interviews of men teaching at the kindergarten and early elementary grades levels many of these men spoke reverently of their perceived responsibility to be a role model for not just the boys in their class but all of the boys at their school. Some of these men even considering it literally a part of their job description, in the sense that it was an understood condition of their employment and one of the most important criteria on which they would be evaluated (Allan, 1994; King, 1998; Sargent, 2001). Still others (and sometimes even the same men) expressed confusion or anxiety about living up to all of the roles they felt expected to model. Among the role modeling expectations felt by men in Sargent's (2001) study were that they would: do 'nothing feminine', not show interest in art or poetry, be 'the men in their [student's] lives', display an interest in athletics, and be an authority figure to students on issues unrelated to school.

Teachers in King's (1998) study insisted that the responsibility of being a role model for boys was one of the most important aspects of their job (and that it was even more important if they happened to be the only male teacher at their school), despite the fact that they were altogether largely unable to identify what being a role model meant in terms of any specific actions. King described this as an "optimistic ritual approach", in which no one (teachers, parents, administrators) knew exactly what the men were doing to benefit boys, but believed that it was important they keep doing it. Perhaps as a result of this type of uncertainty, the men in this study reported an acute awareness of the conflict in simultaneously trying to demonstrate the example of a "real man" they perceived to be expected of them while at the same time trying to meet the age-appropriate needs of their students and perform the duties of a job typically performed by women. Allan (1994) refers to these opposing tensions as a "role modeling paradigm" and suggests these men may feel compelled to choose between an extreme hypermasculine (strong, athletic, bold) or extreme hypomasculine (nurturing, sensitive, gender-role atypical) persona, or are likely to at least struggle for some time to negotiate a comfortable balance between the two.

This tremendous amount of reflection at the teacher level suggests that what has been commonly conceptualized as a passive effect between male teachers and students may actually be as socially constructed as active effects. Related evidence also suggests that the processes involved with gender-related role modeling interactions may be just as subject to active interpretation at the student level. Controlled tests of a "like-sex hypothesis" – predicting that children of one sex would imitate models of the same sex in a variety of gender-typed and gender-neutral activities – have found boys to be likely to imitate masculine-typed behavior regardless of the sex of the model (Barkley, Ullman, Otto & Brecht, 1977). Other school-based studies have also found teacher gender to have little effect on academic motivation or engagement for the majority of young students when compared with the influence of other teacher characteristics, such as demeanor and fairness (Carrington, Francis, Hutchings, Skelton, Read & Hall, 2007). Still others have suggested that gender affiliation is a much more powerful influence on academic engagement at the horizontal peer level than the vertical teacher-student level (Dutro, 2001/2002; McCracken, 1973).

In addition to the problem of identifying clear mechanisms of teacher influence, another difficulty for researchers interested in studying role modeling as either a passive or active effect is that the intended student benefits are often either unclearly defined or otherwise defined in terms of intangible or delayed outcomes. While there is not much evidence that passive (or theoretically passive) effects such as role modeling have a significant effect on any short or immediate term outcomes, this is certainly not to say that there may not in fact be benefits for individual students or for outcomes in the longer term. For short term outcomes, however, passive gender effects appear unlikely to be the sole source of any measurable differences in student outcomes related to gender. Active effects, such as unintended but quantifiably evident variations or bias in teacher-student interactions, would appear instead to be the better source of potential teacher gender-related influence to examine.

Active effects

In studies conducted at the early elementary grades levels, teachers have been found to have lower academic expectations of boys when it comes to ratings of reading or language arts abilities (Beswick, Wilms & Sloat, 2005; Fennema, Peterson, Carpenter & Lubinski, 1990), and even more frequently to have lower social and behavioral expectations of boys (Bennett, Gottesman, Rock & Cerullo, 1993; Birch & Ladd, 1997; Dusek & Joseph, 1983; Stinnett, Bull, Koonce & Aldridge, 1999; Tomada & Schneider, 1997), each of which have been associated with performance on both standardized and unstandardized outcome measures (Beswick, Wilms & Sloat, 2005; Entwisle, Alexander & Olson, 2007; Harris & Rosenthal, 1986; Rosenthal & Jacobson, 1968). Entwisle, Alexander & Olson (2007) further found that for teachers who rated girls higher on an overall preference scale, the girls in their classrooms made greater single-year gains on standardized tests than the boys even after controlling for race, meal subsidy status, parent education, parent expectations of student, family configuration, school segregation status, and school-level parent education. While these studies do not test specifically for bias by teacher gender, considering that the teacher samples in all of these studies were comprised of predominantly or entirely female teachers, the findings of diminished expectations and outcomes for boys does open the door to the hypothesis that there may be at least some level of unintended gender bias being exercised by the teachers in these classrooms.

In studies of classroom interactions, evidence of student gender being related to specific teacher classroom behaviors has also been well documented. Findings include that teachers interact more with boys than girls, boys initiate more contact with teachers than girls, teachers initiate more contact with boys than girls, boys receive both more disciplinary contacts and praise contacts from teachers than girls, and teachers respond to requests for assistance from boys more (Brophy & Good, 1970; Fennema & Peterson, 1985; Garrahy, 2001). When making distinctions among specific kinds of attention, more detailed studies also extend these findings to elaborate that while boys do receive more overall attention, girls still receive more input from teachers about academic performance and boys receive far more input about behavior and classroom procedures (Brophy, 1985; Eccles & Blumenfield, 1985). Further findings add that elevated teacher interactions with boys are not equally distributed across all boys but rather are likely to consist of repeated interactions with one or more specific subset of boys, such as high-achieving, high-confidence or low-achieving boys (Fennema & Peterson, 1985; Sadker, Sadker & Klein, 1991).

Of course, any inferences made about the direct contribution of teacher gender to such imbalances in these studies simply because all or most of the teachers happen to be female must be considered speculative in the absence of experimental controls. However at least one study explicitly designed to gauge unintended gender bias (active effect) between teachers' stated beliefs and actual classroom behaviors has detected mismatch between how teachers believe they treat their students ("gender-blind") and how they actually interact in similarly gendered patterns as those found in the other studies (Garraghy, 2001). Differences in interaction styles may be an important source of active teacher effects given the evidence of positive associations between teacher-child relationship quality and student achievement both immediately (O'Connor and McCartney, 2007) and for years afterward (Hamre & Pianta, 2001), however the current research is not suited to account for other general teacher interaction styles apart from the specifically defined instructional practice tendencies measured.

Specific benefits for boys

One particularly salient factor motivating the current research was to also consider some of the specific claims found commonly in popular media about the benefits of having a male teacher for boys, and particularly those boys identified in certain ways as being most "in need of a male influence" – a characterization commonly ascribed to boys with documented behavior issues or who are living in a single-parent, female-headed household. As detailed earlier in this chapter, the men who teach at these levels are already particularly aware of these common beliefs and the fact that they often permeate the expectations others have of them. The idea that male teachers will not only benefit boys, but will especially benefit a specific subset of boys, has come to be ingrained in popular attitudes about education largely due to its' "intuitively compelling" nature and the dissemination of persuasive anecdotal or speculative reports supporting the idea in popular media outlets and private publications.

A sampling of quotes from a 2006 *Tuscon Citizen* article, for example, includes consistency on this issue across a variety of sources cited, including: an elementary school principal claiming that “[b]oys, more than girls, need a male role model, especially when they come from a one parent family”; a first-grade teacher explaining that “[s]ome kids do not have a male figure at home, … It's so obvious" (which the *Citizen* prefaces for the reader to mean that this teacher “can often tell a boy lacks a father at home because of the lack of discipline”); another first-grade teacher stating that “… young children can benefit academically and socially from having a male teacher”; a first-grade teaching assistant sharing her own (anecdotal) experience that “[b]eing a mom, I know. Mom is more lenient; fathers bring discipline. Students are more attentive when they have a male teacher.”; and from an associate dean of the college of education at a local university declaring that “[e]verybody agrees that having male role models in the classroom is important”, among other generally supportive opinions provided by parents and students (Kalaitzidis, 2006).

A similar report, aired on the ABC network’s *Good Morning America* and reprinted online at the ABC News website, offers the straightforward conclusion that “According to research, the presence of a male teacher in the classroom has an impact not only on boys' self-esteem but also on their academic performance.” The presented evidence for this claim consists primarily of an interview with the director of a non-profit organization dedicated to increasing the number of male teachers in schools, and what the report refers to as a “recent British study from the Training and Development Agency for Schools [which] found that the presence of a male teacher in the classroom for a year closed the achievement gap significantly between boys and girls, especially in English and social studies, subjects that girls tend to do better in than boys.” (Pleshette-Murphy, 2008). Efforts for this research to locate the referenced study on the Training and

Development Agency for Schools' website turned up nothing matching these specific claims, though it is worth noting that the Training and Development Agency for Schools is a government-associated staffing firm, primarily responsible for teacher workforce recruitment and development for all of England, and that several articles promoting the benefits of a teaching career for men were readily available on the site. Similar to reports found in American popular media outlets, the claims in these articles were supported primarily by anecdotal, first-person experience accounts of the problems of boys with no father at home and the need for a type of discipline and positive role modeling that male teachers are characterized as being uniquely able to provide.

Similarly reported claims based on limited evidence, found repeatedly in other news media outlets (Solochek, 2007), private foundation reports (Johnson, 2008), and popular bestsellers (Sexton, 1969; Sommers, 2000), reaffirm popular belief in the specific benefits of male teachers for boys in need of a special kind of "structure" or "discipline", or who need male teachers to supplement the male influence missing in a female-headed, single-parent home. It should not be surprising, considering there remains as of yet almost no controlled academic research on this subject to present in response to these claims, that quotes attributed to first-grade teachers, elementary school principals, college deans, and others – especially when legitimized by a known media news outlet or a published book – might be readily accepted as fact in a general public forum. This is why it is imperative for research to begin catching up to general belief in this area to provide a more constructively dispassionate understanding of what really is (or isn't) important for helping boys achieve better in school.

CHAPTER THREE

METHODOLOGY

The information presented in this chapter is divided into four sections: (1) a brief review of each study part and its goals, (2) a general introduction to the ECLS-K dataset and instruments used, (3) details of the analysis steps for *RQ1* and (4) details of the analysis steps for *RQ2*.

Review of study goals

The goals of this study were: (1) to present a detailed profile of the demographic characteristics, qualifications and instructional practice tendencies (related to the teaching of reading/language arts) of male kindergarten teachers at the population level in America and consider them in comparison with the respective female teacher population, and (2) to test whether or not teacher gender or teacher gender-related characteristics are related to end of year kindergarten reading levels for kindergarten students, and specifically boys, after controlling for other important student, teacher, and school level variables. Using data from the base year teacher and child catalogs of the Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K), these goals were pursued in two distinct study parts, each part corresponding with one of the stated research questions presented in the Chapter 1.

For *RQ1*, descriptive analyses of all variables of interest in the categories of demographic characteristics, qualifications and instructional practices were examined to produce comparative descriptive profiles of the male and female kindergarten teacher populations represented in the data. A series of correlation models were also used to test whether any significant differences found between male and female teachers remained significantly associated with teacher gender after accounting for the potential influence of other available demographic characteristics. Those variables identified as being significantly different between male and female teachers and uniquely associated with gender in the population were identified for inclusion as teacher-level differences to be tested in Part 2 analysis models.

For *RQ2*, a series of multilevel regression models were used to test whether teacher gender-related characteristics or interactions between teacher gender and student gender alone (after controlling for other identifiable differences) were significantly related to the kindergarten end of year student reading proficiency levels among all students. An additional series of models was conducted to test the same interactions again among the population of boys alone, and included additional testing of interactions of teacher gender with behavioral service and resident father statuses for boys' outcomes.

ECLS-K Overview

The Early Childhood Longitudinal Study – Kindergarten Class of 1998-99 (ECLS-K) is a nationally representative study conducted by the United States Department of Education, National Center for Education Statistics (NCES) and focusing on children's early school experiences. The ECLS-K employed a multistage probability sample design to select a nationally representative sample of children attending kindergarten during the 1998/1999 academic year.

Sample Design

The primary sampling units (PSUs) were geographic areas consisting of counties or groups of counties nationwide, identified first using an existing framework created for multipurpose government survey analysis. These original frames were derived from 1990 county-level population data, which the ECLS-K then updated with 1994 population estimates of five-year olds by race/ethnicity from the U.S. Census Bureau. After all necessary adjustments were made, the updated ECLS-K framework of PSUs contained a total of 1335 frames, of which 100 sample PSUs were selected for inclusion by a controlled random method (U.S. Department of Education, National Center for Education Statistics [NCES], 2001b). (See NCES, 2001b, Ch. 4 for full elaboration on statistical controls used in grouping PSU strata).

The second stage units sampled were schools within the selected PSUs. Schools were identified primarily using two national reference datasets from the U.S. Department of Education: the 1995-96 Common Core of Data (for public schools) and the 1995-96 Private School Universe Survey (for private schools). Schools run by the Department of Defense and the Bureau of Indian Affairs (and thus not represented in the Common Core of Data) were also added prior to sample selection. Each of these reference datasets was then limited to include only those schools located in the 100 sample frames selected in the first stage, and then again to include only schools in those frames offering traditional kindergarten, transitional kindergarten/first grade, or an age equivalent program (for ungraded schools).

Following an initial round of second stage selections, a variety of public and private resources in each PSU were consulted to identify any eligible schools not on current records, such as schools set to open and be offering kindergarten by 1998. After all necessary adjustments to update school eligibilities were made, the final totals for the second stage of the ECLS-K base year sample included 914 public schools and 363 private schools for a total of 1,277 schools (NCES, 2001b)

The third stage units sampled were students within the selected schools. Once school sites were selected in the second stage, ECLS-K staff obtained rosters of all kindergarteners enrolled at each school, including all kindergarten age students enrolled in non-mainstream programs for reasons of disability or language problems. A target sample of 24 kindergarteners was set for selection at each school though there was some minor variability allowed in special cases, such as for twins or Asian and Pacific Islander students purposefully oversampled at a given site. A total of 21,260 students were ultimately selected for participation in the ECLS-K base year sample child cohort (NCES, 2001b).

An independent sampling of kindergarten teachers at each site was also conducted during the base year of the study only. This means that teacher data in the base year was collected from all eligible kindergarten teachers at each selected site, including those not necessarily teaching a selected sample student. In subsequent years of ECLS-K data collection, teachers were only selected to participate as a result of having an active sample student placed in their class. Accordingly, only the base year sample of teachers is considered a controlled random sample and appropriate for use as a representation of the national population of kindergarten teachers at that time (NCES, 2001b).

Data and Instruments

Data for the ECLS-K was collected in three main parts: (1) direct students assessments, (2) computer-assisted personal and telephone interviews of parents/guardians, and (3) self-report questionnaires completed by teachers, school administrators, and authorized staff. Data considered in the current research comes primarily from the teacher self-report questionnaires and student direct assessments.

Teacher questionnaires consisted of three parts (A, B and C) and were designed to capture a comprehensive array of data on teacher backgrounds, practices and beliefs. Part A contained questions about the teacher's classroom composition and characteristics. Part B contained questions about the teacher's background, questions about classroom organization, instruction and evaluation methods, and questions about teacher' overall views on issues concerning students and school. Part C asked teachers more specifically about their perceptions of individual sample students in their classroom. Teacher data considered in the current research came only from Parts A & B, with no items from Part C.

Child data was collected by direct assessment in the fall and spring of the kindergarten year. The direct cognitive assessments contained items on reading, mathematics, and general knowledge. The reading assessment contained items at five progressive proficiency levels: (1) identifying upper- and lower- case letters of the alphabet by name, (2) associating letters with sounds at the beginning of words; (3) associating letters with sounds at the end of words; (4) recognizing common words by sight; and (5) reading words in context (NCES, 2001b).

Outcome measures of student achievement on these assessments were available in a number of different scoring formats in the dataset to allow researchers the flexibility of choosing the most appropriate scoring for the purpose of their analysis, including: number right scores, IRT (Item Response Theory) scores, standardized scores, criterion-referenced proficiency level and proficiency probability scores (NCES, 2001b). For the current study, student achievement is measured using IRT scores for the reading portion of the spring direct cognitive assessment. IRT scores are considered the most appropriate score type for measuring gains over time (NCES, 2001b).

RQ1: Gender-related differences between male and female teachers

Analytic Sample

The analytic teacher sample for *RQ1* included all available kindergarten teachers from the ECLS-K base year teacher file with non-missing data on the dependent variables of interest.

Variables

All variables associated with *RQ1* were conceptually grouped into three categories corresponding to the three areas of interest regarding the teacher population for this study: (1) demographic characteristics, (2) qualifications, and (3) instructional practices. Demographic characteristics include all variables in the data reflecting those personally descriptive traits which would otherwise be constant for each individual regardless of their employment as a kindergarten teacher (e.g. age, race). Qualifications include all variables of interest in the data that provide detail about either the educational background or practical job experiences of kindergarten teachers in the sample (e.g. total amount of education-related coursework, certification type, total years teaching experience, etc.).   Instructional practice variables are limited in scope here to only those practices pertaining to reading and language arts instruction, and are slightly different in nature from variables used in the other two categories in that they are composite scores derived from teacher responses to several individual questionnaire items each. For variables pertaining to teacher race/ethnicity and for some instructional practice individual survey items, recoding of original values was necessary for practical use and interpretation. In such cases an explanation of the recoded values are presented in the next chapter immediately preceding the appropriate results. (Further information about the creation of these composites is provided in Appendices A and C.)

All specific dependent variables considered within each of the three areas of interest are summarized below. (For a complete summary of all analysis variables used for *RQ1*, including design weights and original ECLS-K source information for variables described here, see Tables 1.1a and 1.1b.)

For demographic characteristics, four variables available in the data were considered: *teacher gender*, *teacher age (continuous)*, *teacher age (categorical)*, and *teacher race/ethnicity*.

For educational qualifications, eight variables from the ECLS-K base year teacher questionnaire were considered: *highest level of education received*, *total number of courses taken in early childhood education*, *total number of courses taken in elementary education*, *total number of courses taken in reading instruction methods*, *certification type*, *certification in early childhood education*, *certification in elementary education*, and *certification in another field of education*.

For job experience qualifications, six variables from the ECLS-K base year teacher questionnaire and ECLS-K base year administrator questionnaire were considered: *total* *number of years kindergarten teaching experience*, *class type*, *school type*, *school urbanicity*, *total student enrollment*, and *total school minority enrollment*.

For instructional practices, seven composite variables, constituting a known set of teacher practices measured by the individual items in the ECLS-K teacher questionnaire as established by factor analytic testing in prior research and recommended for use with studies of teaching practices in this dataset (Guarino, Hamilton, Lockwood, Rathbun, & RAND, 2006) were considered: *reading and writing activities*, *phonics*, *didactic instruction*, *comprehension*, *student centered instruction*, *reading and writing skills*, and *mixed ability grouping*. (Further information about the creation of these composites is provided in Appendices A and C.)

Analytic strategy

Response distributions were analyzed within both male and female kindergarten teacher populations separately, as well as in the combined teacher population for reference, for each of the dependent variables of interest. Descriptive statistics were reported in terms dictated by the form of the data for each variable, either as frequency distributions and corresponding population estimates or as mean scores, range and standard deviation calculations and corresponding population estimates. Projected population estimates were derived from the sample findings using the appropriate ECLS-K defined teacher-level sample weights. Sample weights were designed by the ECLS-K to adjust for differential selection probabilities and reduce bias due to nonresponse, allowing for inferences about the larger population to be made using the data (for more information on sample weighting procedures see “Taylor series method” in *Definition of Terms*, Chapter 1).

With the exception of *teacher age (continuous)* and *total number of years kindergarten teaching experience*, all dependent variables of interest were categorical in nature and tested for associations with teacher gender using a reductive two-step process. First, all dependent variables were tested in series of independent two-way contingency table tests. Contingency table results described the overall distribution pattern of teachers' responses by gender in both sample counts and population estimates, and also provided a design-corrected test statistic to measure the likelihood of finding the same pattern of responses in a completely random sample where there was no association between teacher gender and the given variable of interest. If the test statistic failed to reject this hypothesis, the distribution results were reported but the variable was not tested for any further associations. For those variables with test statistics which did indicate a significant association with teacher gender, the second step was to test the variable further in a series of more specified regressions designed to control for the potential contributions of other demographic characteristics (*teacher age* and *teacher race*) to the observed pattern of responses. Depending on the nature of the categorical response data for a given variable (ordered vs. unordered), either ordered logistic regression or multinomial logistic regression techniques were used as appropriate. The results of these regressions were used to identify which variables demonstrated a significant unique association with *teacher gender* in the data. (See Figure 1 for a decision matrix illustrating this process.)

For the continuous dependent variables *teacher age (continuous)* and *total number of years kindergarten teaching experience*, the same process was followed with the only differences being that initial descriptive distributions were calculated as simple mean comparisons between male and female teachers (rather than contingency tables) and secondary regressions were conducted as general linear models (rather than logistic).

Following all comparison testing and regression checks, those variables found to be significantly different between male and female teachers in the population, and their differences to be uniquely associated with gender, were selected for inclusion as teacher gender-related differences to be controlled for in Part 2 analysis models.

RQ2: Teacher gender and student achievement

Analytic Sample

All eligible students in the ECLS-K base year child catalog with non-missing data were initially selected for analysis at this step with two restrictions: (1) language minority students who were unable to demonstrate English proficiency as measured by the Oral Language Development Scale (OLDS), and (2) students who transferred between either classrooms or teachers between fall and spring assessments of the kindergarten year. Prior to adjustment for missing data, this ECLS-K base year student sample contained 7,649 male kindergarten students (representing an estimated 1,668,479 male kindergarten students, or 51% of all kindergarten students in the national population) and 7,386 female kindergarten students (representing an estimated 1,584,775 female kindergarten students, or 49% of all kindergarten students in the national population).

Teachers were selected into the sample based on having any one or more of the analytic sample students in their classroom. This did not significantly change the teacher sample from its original size or proportions.

Variables

Analysis for *RQ2* consisted of a series of three-level hierarchical linear regression models on a dependant variable measure of student reading achievement at the end of kindergarten. The three-level structure for each model includes students (level 1), nested within teachers (level 2), nested within schools (level 3). Variables tested and used as covariates are summarized below according to their place in that structure. (For a complete summary of all analysis variables used for research question 2, including design weights and original ECLS-K source information for variables described here, see Tables 2.1a and 2.1b.)

Student Level Covariates (Level 1)

In the three-level structure of this analysis, level 1 variables were intended to account for important between student/within class sources of variance. Student achievement differences among younger students are known to be associated with a number of significant student level influences. Prominent among these include: poverty/SES status (Entwisle & Alexander, 1992; Arnold & Doctoroff, 2003), race/ethnicity (Barbarin, 2002, Hedges & Nowell, 1999), and home learning environment and preschool attendance (Dubow & Ippolito, 1994; Vernon-Feagans, 1996; Christian, Morrison & Bryant, 1998). Based on this information and other controls necessary to account for gains made within the year, the following level 1 covariates were included in the models: *fall reading IRT scale score*, *time gap between fall and spring assessments*, *student race/ethnicity*, and *family SES*. For extended analysis of the boys only sample as described previously, models applied to that sample also included level 1 covariates accounting for *resident father status* and *behavior/emotional services status*.

Teacher Level Covariates (Level 2)

In the three-level structure of this analysis, level 2 variables are intended to account for between class/within school variance. In this study, level 2 covariates included *teacher gender*, along with any identified teacher gender-related characteristics identified by the results from research question 1. The logic behind this approach was to distinguish between effects attributable in this data to *teacher gender*, specifically, and effects potentially associated with any identifiable practical differences in teacher qualifications or instructional methods that happen to be associated with teacher gender in order to better isolate and explain the nature of any potential associations that arise. The specific variables identified for inclusion as teacher-level covariates are presented with detailed explanation in the results chapter of this study.

School Level Covariates (Level 3)

In the three-level structure of this analysis, level 3 variables were intended to account for between school variance. Issues of minority segregation and poverty have both been shown to affect achievement at the school level (Kainz & Vernon-Feagans, 2007). In order to account for this, the two specified school level variables included were: *percent of students eligible for free lunch at school* (as a proxy for concentrations of poverty), and *percent minority student enrollment at school* (as a proxy for segregation),.

Dependent variable

The dependant variable in each model was end of year reading achievement level as measured by spring reading IRT scale score. IRT (Item Response Theory) scale scores measure student performance on a set of test questions with a broad range of difficulty by analyzing children's response patterns and omissions, in combination with information about the level of difficulty of each item and a formula designed to guard against the probability of any given student having guessed correctly on items they did not actually know, to produce a score that places the student on a common continuum of scores comparable to each other as if all students had answered all items on the same tests. IRT scores are considered the most appropriate option for measuring student growth over time on the ECLS-K direct cognitive assessment (NCES, 2001b).

Predictor variables

The predictor variable in the first (full) sample model was an interaction effect between teacher gender and student gender. Once the best-fit conditional model was identified containing all other significant covariates, this interaction effect was added to create the full model and test for any potential influence associated with an interaction of teacher gender and student gender teacher among all students detectable above and beyond the influence of the other established covariates.

The predictor variable in the second (boys only) sample model was an interaction effect between teacher gender and the additional "risk" factors of behavior services status or resident father status. Because of the restricted nature of this sample, an explicit interaction effect between student and teacher gender, specifically, was not necessary because there was no variability in student gender. The addition of *teacher gender*, alone, in this case was sufficient for identifying the effect of teacher gender for boys beyond other covariates.

Analytic strategy

Analysis pursuant to *RQ2* considered the influence of *teacher gender*, and teacher gender-associated variables identified by *RQ1*, on student reading achievement levels at the end of the kindergarten academic year. The first step required for this analysis was to prepare the full analytic student sample to be used by examining the nature of all relevant missing data in the original ECLS-K student data catalog and either imputing data or dropping cases as appropriate. Data for each individual variable of interest was considered missing at random unless otherwise noted in ECLS-K codebook documentation. For students with missing data on one or more variable of interest, multiple imputation procedures were used to impute best predicted responses using all other available data as predictors. Students with missing data for all variables of interest could not provide sufficient information for imputation and were therefore dropped from consideration. Complete information about the steps used in cleaning and preparation of the analytic sample are provided in Appendix B.

Once the analytic student sample was established with no missing data, descriptive statistics were run for the sample counts and population estimates across all analysis variables. These results were compared to the same descriptive statistics run on the original (unmodified) ECLS-K student sample to detect any significant differences caused by the dropped cases or imputed data in the analytic sample. Response distributions and means in the analytic sample were nearly identical to those in the original ECLS-K sample in every case. (The restricted (boys only) analytic sample was created simply by dropping all female students from the full analytic sample, so that the boys in both samples are identical with respect to any imputed data. See Tables 2.2 and 2.3 for descriptive comparisons of each analytic sample to the unadjusted ECLS-K sample.)

Once the analytic full sample was confirmed, outcome variance for student spring reading achievement scores was tested in both an unconditional two-level and unconditional three-level model design. Variance decomposition results for each model were compared using a likelihood ratio test to confirm whether a three-level model structure was indeed most appropriate for this data. A likelihood ratio test compares the model information of both structures and provides a test statistic based on the null hypothesis that the random effects associated with the extra level (in this case, level 2) in the three-level model should be omitted (West, Welch, & Gałecki, 2007). The results of the likelihood ratio test between model structures for the analytic sample were significant, indicating that this null hypothesis could be rejected and that a three-level variance model was appropriate for this analysis, $\bar{\chi}^2(1) = 80.49$, $p = 0.000$. See Table 2.4 for results of two and three level model comparisons and Table 2.5 for a descriptive summary of groupings in the three-level model.

For each analytic sample, spring reading achievement was first tested in an unconditional model to determine the total amount of variance in spring reading achievement scores associated with each level of the model. In other words, how much of the overall expected variance in scores is attributable to differences between students in classrooms, between classrooms in schools, and between schools – regardless yet of what the specific sources of those differences may be. This total amount of variance attributable to a specific level of the model is known as the intraclass correlation (Raudenbush & Bryk, 2002).

Following the establishment of the random effects components for each level in the unconditional model, conditional models were then sequentially constructed using the “building up” method, as described by West, Welch, & Gałecki (2007). Using the “building up” method, all of the covariates associated with each level of the model are added as fixed effects one group (level) at a time, tested, and either retained or rejected (individually) before then proceeding to add the next group of covariates associated with the next level. For example, if we believed that there were five important student level variables to account for as level 1 covariates, we would build the level 1 conditional model by adding all five of those covariates as fixed effects to the unconditional model. We would then examine the results to determine how much of the total level 1 random effect has been reduced by the addition of these controls to the model, and what the significance levels were for the contributions of each of the five covariates individually. If, for instance, three of the five covariates accounted for a significant reduction in the random variance of the model but the effects of the other two were not significant, then we would retain only the three significant variables as level 1 covariates in our models moving forward. This process is repeated two more times as level 2, and level 3 covariates are each added to the best conditional model preceding it.

Once the best fit conditional model was established, including all significant covariates at each level except for the predictor variable of interest, the final step was to test spring achievement scores in the full model. The full model simply added the predictor variable to the best fit three-level conditional model to determine whether it affects any further significant reduction in the remaining variance after all other known significant covariates are accounted for.

CHAPTER FOUR

RESULTS

RQ1:Gender-related differences between male and female teachers

Demographic characteristics

Results for all demographic characteristic distributions and comparisons are summarized in Tables 1.2 and 1.3. Personal demographic characteristic variables available for the teacher population in this data were gender, age and race. Gender was also the primary sorting variable across which age, race, and all variables of interest within further categories were considered. The analytic teacher sample contained a total of 72 male teachers and 2,934 female teachers, which projected by weighted estimates to represent a total of 4,197 male teachers and 183,230 female teachers in the national population.

The mean age of male teachers (37.56) was found to be younger than the mean age of female teachers (41.24), a difference that was significant, $F(1, 437) = 5.42$, $p = .02$. In order to gain a better understanding about what this mean age difference indicated with respect to overall age-related distribution patterns across the population, teachers' reported (continuous) ages were also recoded to values representing one of seven categorical age-range groupings. Each age-range grouping covered a five-year inclusive spans (meaning that, for example, in the group identified as ages "24-28", teachers reporting any continuous age between 24 years/0 months/0 days and 28 years/11months/31 days were included) and groupings were constructed across an overall span designed to ensure that all teachers in the sample were accounted for (see Appendix A for full recoding details). Frequency results presented in this and all subsequent sections are reported as weighted population estimates, followed by rounded (gender-specific) population percentages in parentheses. Unweighted sample counts and population totals are also provided in the associated summary tables.

A summary of teacher categorical age distributions is presented in Table 1.3. The comparative distributions found that: 1,157 (28% of) male and 26,638 (15% of) female teachers were between the ages of 24 and 28; 469 (13% of) male and 22,429 (13% of) female teachers were between the ages of 29 and 33; 358 (9% of) male and 21,491 (12% of) female teachers were between the ages of 34 and 38; 1,014 (24% of) male and 23,151 (13% of) female teachers were between the ages of 39 and 43; 591 (14% of) male and 33,814 (20% of) female teachers were between the ages of 44 and 48; 501 (12% of) male and 31,502 (18% of) female teachers were between the ages of 49 and 53; and 107 (3% of) male and 19,422 (11% of) female teachers were between the ages of 54 and 58.

In calculating contingency table distributions with survey data, Stata provides two versions of the chi-squared test of independence statistic: an uncorrected Pearson chi-squared statistic and an *F*-statistic which is the chi-squared statistic corrected for the survey design. Using the design-corrected *F*-statistic, test results supported that this pattern of age distributions indicates a significant association between age group and teacher gender in the population, $F(5.57, 2444.71) = 2.19$, $p = .046$, meaning that the comparative distributions do not demonstrate a pattern that we would reasonably expect to find if they had both been drawn at random from a population in which age group and teacher gender were not associated. These results provide a more informative explanation of the age differences between male and female teacher populations, indicating heavy concentrations of male teachers in the “24-28” and “39-43” year old groupings followed by steep, rapid declines in each group after these; while female distributions remain relatively consistent across all groups, with peaks in the “44-48” and “49-53” year old groupings. The fact that these differences were found to be significant is consistent with the significant difference found when comparing simple mean age distributions.

The ECLS-K teacher questionnaire item pertaining to teacher race/ethnicity originally offered five possible response options: (1) *Native American or Pacific Islander*, (2) *Asian*, (3) *Black or African American*, (4) *Native Hawaiian or Other Pacific Islander*, or (5) *White*. Because of the few number teachers identified in the individual race response category of *Native Hawaiian or Other Pacific Islander*, this data was suppressed in the public-use data catalog as potentially identifiable information. Due also to small sample sizes, the individual race response categories of *Native American or Pacific Islander* and *Asian* were combined to form a single response group in the preparation of the analytic teacher sample for this study. As a result, the original five categorical response options for teacher race were recoded and collapsed into four analytic categories: (1) *No response/Data suppressed*, (2) *NA/PI/As*, (3) *Black or African-American* and (4) *White*. The recoded categories of *No response/Data suppressed*, *Black or African-American* and *White* remain unchanged in their frequencies from original teacher questionnaire responses through this process. The created category *NA/PI/As* simply collapsed the responses of teachers who identified as *Native American or Pacific Islander* and those who identified as *Asian* into a single combined count. (Note that because teacher self-identification as Hispanic/non-Hispanic was asked as a separate, additional question in the fall kindergarten teacher questionnaire, it was not possible to accurately associate teachers' response to this question with their original identification in many cases due to suppressed data and was therefore not considered here.)

A summary of teacher *race/ethnicity* distributions is presented in Table 1.3. The comparative distributions found that: 379 (9% of) male and 6,049 (3% of) female teachers were identified as *No response/Data suppressed* respondents; 65 (2% of) male and 6,233 (3% of) female teachers identified as *NA/PI/As*; 257 (6% of) male and 12,025 (7% of) female teachers identified as *Black or African-American*; and 3,496 (83% of) male and 158,923 (87% of) female teachers identified as *White.* Using the design-corrected *F*-statistic, the distribution of race/ethnicity response identifications by gender in this data did not indicate a significant association between the two characteristics, $F(2.25, 989.62) = 2.123$, $p = .114$.

Qualifications (educational background)

Results for all educational background comparisons are summarized in Tables 1.4a and 1.4b. A summary of teacher response distributions for *highest level of education received* is presented in Table 1.4a. The comparative distributions found that: 1,388 (36% of) male and 53,147 (31% of) female teachers had achieved either a high school diploma, associate's degree or bachelor's degree as their highest level of education; 1,603 (42% of) male and 57,924 (34% of) female teachers had achieved up to some post-bachelor's degree coursework; 696 (18% of) male and 50,170 (30% of) female teachers had achieved up to a master's degree; 170 (4% of) male and 9,793 (6% of) female teachers had achieved up to an education specialist or other professional degree; and 0 (0% of) male and 469 (0.3% of) female teachers with a doctorate degree. Comparison testing of these distributions did not indicate evidence of a significant association between highest level of education received and teacher gender in the population, $F(3.47, 1486.74) = .605$, $p = .636$.

A summary of teacher response distributions for *total number of early childhood education courses taken* is presented in Table 1.4a. The comparative distributions found that: 537 (16% of) male and 15,464 (9% of) female teachers reported having taken zero early childhood courses; 288 (9% of) male and 11,332 (7% of) female teachers reported having taken one early childhood course; 839 (25% of) male and 17,882 (11% of) female teachers reported having taken two early childhood courses; 379 (11% of) male and 16,764 (10% of) female teachers reported having taken three early childhood courses; 440 (13% of) male and 13,585 (8% of) female teachers reported having taken four early childhood courses; 79 (2% of) male and 7,230 (4% of) female teachers reported having taken five early childhood courses; and 807 (24% of) male and 88,021 (52% of) female teachers reported having taken six or more early childhood courses. Comparison testing of these distributions did indicate evidence of a significant association between total number of early childhood education courses taken and teacher gender in the population, $F(4.50, 1950.22) = 2.97$, $p = .015$, so that further testing of this association was warranted. Results of further regression testing of significant educational background variables are presented together at the end of this section.

A summary of teacher response distributions for *total number of elementary education courses taken* is presented in Table 1.4a. The comparative distributions found that: 272 (8% of) male and 7,741 (5% of) female teachers reported having taken zero elementary education courses; 61 (2% of) male and 3,455 (2% of) female teachers reported having taken one elementary education course; 291 (8% of) male and 5,206 (3% of) female teachers reported having taken two elementary education courses; 127 (4% of) male and

6,096 (4% of) female teachers reported having taken three elementary education courses; 331 (9% of) male and 6,074 (4% of) female teachers reported having taken four elementary education courses; 144 (4% of) male and 6,738 (4% of) female teachers reported having taken five elementary education courses; and 2,396 (66% of) male and 135,611 (79% of) female teachers reported having taken six or more elementary education courses. Comparison testing of these distributions did not indicate evidence of a significant association between the total number of elementary education courses taken and teacher gender in the population, $F(5.14, 2215.89) = 1.65$, $p = .143$.

A summary of teacher response distributions for *total number of reading methods courses taken* is presented in Table 1.4a. The comparative distributions found that: 240 (7% of) male and 6,361 (4% of) female teachers reported having taken zero reading methods courses; 558 (15% of) male and 26,485 (16% of) female teachers reported having taken one reading methods course; 1,395 (38% of) male and 42,763 (25% of) female teachers reported having taken two reading methods courses; 555 (15% of) male and 32,209 (19% of) female teachers reported having taken three reading methods courses; 634 (17% of) male and 17,776 (10% of) female teachers reported having taken four reading methods courses; 93 (3% of) male and 8,303 (5% of) female teachers reported having taken five reading methods courses; and 164 (5% of) male and 36,641 (21% of) female teachers reported having taken six or more reading methods courses. Comparison testing of these distributions did indicate evidence of a significant association between total number of reading methods courses taken and teacher gender in the population, $F(4.86, 2091.51) = 2.44$, $p = .034$, and that further testing of this association was warranted. Results of further regression tests of identified variables are presented together at the end of this section.

Certification type and specific area of certification were also considered part of the educational background qualifications of teachers. A summary of teacher response distributions for *certification type* is presented in Table 1.4b. The comparative distributions found that: 189 (5% of) male and 5,551 (3% of) female teachers reported holding no official teaching certification; 630 (16% of) male and 16,429 (9% of) female teachers reported holding a temporary certification; 227 (6% of) male and 4,858 (3% of) female teachers reported holding some type of alternate certification; 1,005 (26% of) male and 39,188 (22% of) female teachers reported holding a standard teaching certification; and 1,783 (47% of) male and 109,408 (62% of) female teachers reported holding the highest level of certification available to them. Comparison testing of these distributions did not indicate evidence of a significant association between certification type held and teacher gender in the population, $F(3.65, 1590.30) = 1.467$, $p = .214$.

In terms of specific areas of certification, teachers were asked in three, separate Yes/No questions about whether they held certification in early childhood education, elementary education and/or some other field. A summary of teacher response distributions pertaining to each area is presented in Table 1.4b. For *certification in early childhood education*, 1,260 (33% of) male and 94,357 (53% of) female teachers did hold specific certification in this field. This difference was found to be significant, $F(1, 436) = 5.62$, $p = .018$, so that further testing of this association was warranted. Results of further regression testing are presented together at the end of this section. For *certification in elementary education*, 3,433 (90% of) male and 146,392 (83% of) female teachers did hold specific certification in this field. This difference was not found to be significant, $F(1, 435) = 1.93$, $p = .166$. For *certification in another field of education*, 937 (22% of) male and 48,290 (27%

of) female teachers did hold certification in some area other than early childhood or elementary education. This difference was not found to be significant, $F(1, 438) = .545$, $p = .461$.

From the initial series of two sample comparison tests, three variables of interest pertaining to teachers' educational background qualifications were identified as areas of simple significant difference between male and female teachers: *total number of early childhood education courses taken*, *total number of reading methods courses taken*, and *certification in early childhood education*. Each of these variables were further tested to find out if that significant association remained after controlling for other demographic characteristics of age and race. The specific regression technique most appropriate for each variable was determined by the nature of the data for each of the three dependent variables, and as such varied in some cases. Variables in this section required the use of two different techniques: ordered logistic regression (for ordered categorical response data) and multinomial logistic regression (for unordered categorical response data).

Results for these regressions are summarized in Table 1.8 and show that after controlling for age and race only *total number of early childhood education courses taken*, $\beta = -.88$, $t = -3.21$, $p = .001$, and *certification in early childhood education* $\beta = .82$, $t = 2.26$, $p = .024$, remained significantly associated with teacher gender. *Total number of reading methods courses taken* was only associated with teacher age once all other available demographic characteristics were considered.

Qualifications (job experience)

For the organization of this study, variables of interest categorized as part of a teacher's practical job experience included variables pertaining to both length of experience and location of experience (as described by measures of school type, school urbanicity and student enrollment at the sites teachers were currently employed at the time of the survey). Results for the mean and frequency distribution comparisons of these job experience variables are summarized in Tables 1.5 and 1.6.

Results for the mean comparison of *total years of kindergarten teaching experience* between male and female teachers is presented in Table 1.5. Male teachers were found to have fewer years experience on average (4.54) than female teachers (8.24). This difference was significant, $F(1, 438) = 25.87$, $p < .001$. Results of further regression testing of significant job experience variables are presented together at the end of this section.

The variable *class type* described the instructional day setting of the class primarily taught by the responding teacher (AM/PM/All day/Multi). A summary of teacher response distributions for *class type* is presented in Table 1.6. The comparative distributions found that: 1,376 (33% of) male and 28,699 (16% of) female teachers reported teaching AM kindergarten only; 669 (16% of) male and 11,585 (6% of) female teachers reported teaching PM kindergarten only; 423 (10% of) male and 29,564 (16% of) female teachers reported teaching both AM and PM kindergarten sessions; 1,729 (41% of) male and 112,502 (61% of) female teachers reported teaching in all day kindergarten classrooms; 0 (0% of) male and 748 (0.4% of) female teachers reported teaching both AM and all day kindergarten sessions; and 0 (0% of) male and 132 (0.1% of) female teachers reported teaching both PM and all day kindergarten sessions. Comparison testing of these distributions did indicate evidence of a significant association between class type and teacher gender in the population, $F(4.19, 1839.43) = 3.04$, $p = .015$, with male teachers much less likely than female teachers to be teaching in full-day kindergarten settings. Results of further regression tests are presented together at the end of this section.

The variable *school type* described whether a teacher was currently employed in a public or private school at the time of the survey. A summary of teacher response distributions for *school type* is presented in Table 1.6. The comparative distributions found that: 3,376 (80% of) male and 145,925 (80% of) female teachers were working in public schools; and 822 (20% of) male and 37,305 (20% of) female teachers were working in private schools. Comparison testing of these distributions did not indicate evidence of a significant association between school type and teacher gender in the population, $F(1, 439) = .011$, $p = .917$.

The variable *school urbanicity* described the municipal context in which the school that a teacher was currently employed in at the time of the survey was located. A summary of teacher response distributions for *school urbanicity* is presented in Table 1.6. The comparative distributions found that: 2,396 (57% of) male and 88,419 (18% of) female teachers were working in schools located in a central city (urban) area; 393 (9% of) male and 53,324 (29% of) female teachers were working in schools located in an urban fringe (suburban) area; and 1,408 (34% of) male and 41,487 (23% of) female teachers were working in schools located in a small town (rural) area. Comparison testing of these distributions did indicate evidence of a significant association between school urbanicity and teacher gender in the population, $F(1.79, 787.71) = 5.01$, $p = .009$. Results of further regression tests are presented together at the end of this section.

The variable *total school enrollment* indicated a categorical approximation of the total student population at the school in which the teacher was currently employed at the time of the survey. A summary of teacher response distributions for *total school enrollment* is presented in Table 1.6. The comparative distributions found that: 1,160 (29% of) male and 21,462 (12% of) female teachers were teaching in schools with a reported total enrollment of between 0 and 149 students; 410 (10% of) male and 31,128 (17% of) female teachers were teaching in schools with a reported total enrollment of between 150 and 299 students; 422 (10% of) male and 47,634 (26% of) female teachers were teaching in schools with a reported total enrollment of between 300 and 499 students; 1,232 (30% of) male and 50,633 (28% of) female teachers were teaching in schools with a reported total enrollment of between 500 and 749 students; and 823 (20% of) male and 30,876 (17% of) female teachers were teaching in schools with a reported total enrollment of 750 or more students. Comparison testing of these distributions did indicate evidence of a significant association between total school enrollment at school taught in and teacher gender in the population, $F(3.55, 1532.35) = 3.32$, $p = .013$, with male teachers much more likely than female teachers to be teaching in schools with total student enrollments of either less than 150 students or over 500 students. Results of further regression tests are presented together at the end of this section.

The variable *total school minority enrollment* indicated a categorical approximation of the percentage of students at the school in which the teacher was employed at the time of the survey who identified as members of a racial/ethnic minority group. A summary of teacher response distributions for *total school minority enrollment* is presented in Table 1.6. The comparative distributions found that: 1,142 (28% of) male and 52,115 (29% of) female teachers were teaching at schools with a reported less than 10% total minority student enrollment; 552 (14% of) male and 32,636 (18% of) female teachers were teaching at schools with a reported total minority student enrollment of between 10% and less than 25%;

314 (8% of) male and 32,407 (18% of) female teachers were teaching at schools with a reported total minority student enrollment of between 25% and less than 50%; 347 (9% of) male and 20,170 (11% of) female teachers were teaching at schools with a reported total minority student enrollment of between 50% and less than 75%; and 1,692 (42% of) male and 41,696 (23% of) female teachers were teaching at schools with a reported total minority student enrollment of over 75%. Comparison testing of these distributions did not indicate evidence of a significant association between total school minority enrollment at school taught in and teacher gender in the population, $F(3.12, 1337.6) = 2.14$, $p = .09$.

From the initial series of two sample comparison tests, four variables of interest pertaining to teachers' job experience background qualifications were identified as areas of simple significant difference between male and female teachers: *total years of kindergarten teaching experience, class type taught (AM/PM/All day)*, *school urbanicity*, and *total school enrollment*. Each of these variables were further tested to find out if a significant association remained after controlling for other available teacher demographic characteristics of age and race. The specific regression technique most appropriate for each variable was determined by the nature of the data for each of the four dependent variables, and as such varied in some cases. Variables in this section required the use of three different techniques: general linear regression (for continuous response data) ordered logistic regression (for ordered categorical response data) and multinomial logistic regression (for unordered categorical response data).

Results for these regressions are summarized in Table 1.9. and show that after controlling for age and race only *total years of kindergarten teaching experience*, $\beta = -1.99$, $t = -4.24$, $p < .001$; *school urbanicity*, Urban fringe location: $\beta = -1.29$, $t = -2.95$, $p = .003$; and the *class type* environments of, *AM Only*: $\beta = 1.176$, $t = 2.69$, $p = .007$; *PM Only*: $\beta =$

1.25, $t = 2.92$, $p = .004$; *AM + All day*: $\beta = -28.64$, $t = -48.32$, $p < .001$; *PM + All day*: $\beta = -28.72$, $t = -30.93$, $p < .001$ remained significantly associated with teacher gender. *Total school enrollment*, was only associated with teacher race/ethnicity once all available demographic characteristics were considered.

Instructional practices

As a reminder, unlike variables in previous categories which each corresponded directly to teacher or administrator responses to individual questionnaire items, the classroom instructional practice variables here each represent a composite score derived from several different questionnaire items regarding teachers' use of specific activities. These composite scores were constructed in a slightly modified adaptation of the instructional practice composite variables identified for use with the ECLS-K dataset by Guarino, Hamilton, Lockwood, Rathbun & RAND (2006). The only modification made for this study is that composite scores were produced as mean scores (rather than sums) to allow for more meaningful interpretation of results. For a full list of individual questionnaire items represented in each composite variable and their associated factor loadings as identified by Guarino, Hamilton, Lockwood, & Rathbun & RAND (2006), see Appendix C.

Prior to calculating composite means, individual questionnaire items were recoded so that categorical responses all used an equivalent numeric scale. This recoding was necessary because one set of individual items on the teacher questionnaire offered two different response options (1 = "Taught at a higher grade level", 2 = "Children should know already") for activities that teachers report never using as part of their instruction, while the other set of individual items offered only one (1 = "Never"). As a result, all of the remaining categorical options for each set of individual items were identical in their associated values, but were off by one in their numeric values. All individual items were recoded to correct this prior to creating composite scores, with the only practical change to the original format being that those items with two response options for activities never taught were collapsed to a single value. Recoded values for frequency of teacher reported use of activities across all items became: 0 = “Never”, 1 = “Once a month or less”, 2 = “Two to three times per month”, 3 = “One to two times per week”, 4 = “Three to four times per week”, and 5 = “Daily”. For complete details on the cleaning and construction of these composite scores in data preparation, see Appendix A.

A summary of teacher response distributions for each classroom instructional practice measure is presented in Table 1.7. The mean item response for instructional practices relating to the use of *Reading and writing activities* was 3.05 (or approximately within the range of “One to two times per week”) for male teachers and 2.92 (or approximately within the range of “Two to three times per month”) for female teachers, a difference that was not significant, $F(1, 422) = 0.96$, $p = .33$. The mean item response for instructional practices relating to the use of *Phonics* activities was 4.4 for male teachers and 4.49 for female teachers (both approximately “Three to four times per week”), a difference that was not significant, $F(1, 422) = 0.83$, $p = .36$. The mean item response for instructional practices relating to the use of *Didactic instruction* was 2.67 for male teachers and 2.52 for female teachers (both approximately “Two to three times per month”), a difference that was not significant, $F(1, 422) = 0.42$, $p = .52$. The mean item response for instructional practices relating to the use of *Comprehension* activities was 3.19 for male teachers and 3.63 for female teachers (both approximately “One to two times per week”), a difference that was not significant, $F(1, 421) = 2.03$, $p = .15$. The mean item response for instructional practices relating to the use of *Student centered instruction* was 2.28 for male teachers and 2.23 for female teachers (both approximately "Two to three times per month"), a difference that was not significant, $F(1, 421) = 0.19$, $p = .67$. The mean item response for instructional practices relating to the use of *Reading and writing skills* focused lessons was 2.26 for male teachers and 2.11 for female teachers (both approximately "Two to three times per month"), a difference that was not significant, $F(1, 421) = 0.33$, $p = .57$. The mean item response for instructional practices relating to the use of *Mixed ability grouping* activities was 2.85 for male teachers and 2.92 for female teachers (both approximately "Two to three times per month"), a difference that was not significant, $F(1, 421) = 0.05$, $p = .82$.

Because no classroom instructional practice variables were found to be significantly associated with teacher gender, no further testing of these variables was required.

RQ2: Teacher gender and student achievement

Research question 2 considered the associations of teacher gender and specific teacher gender-related characteristics (as identified by the results of *RQ1*) with end of kindergarten reading levels after controlling for other significant covariates at the student and school levels. These were tested using three-level hierarchical linear modeling techniques to account for the nested structure of the data, and models were applied to two different iterations of the student sample: the full analytic sample (boys and girls) and a restricted analytic sample (boys only). The full analytic sample included all students available in the ECLS-K base year dataset with the exception of students who either changed schools or teachers during the year or who did not pass the OLDS English proficiency test during the first round of child assessments. Unless otherwise noted, instances of missing data among the remaining students were considered missing at random and multiple imputation procedures were used to produce best estimate predictions of missing values using all other covariates of interest as predictors. The restricted analytic sample was created simply by dropping all female students from the full analytic sample, so that the boys in both samples are identical with respect to any imputed data. A summary of the overall mean and frequency distribution descriptive characteristics of both analytic samples, alongside (for comparative reference) the same descriptives as seen in the unadjusted ECLS-K student sample (prior to restrictions and imputation procedures used in this study), is presented in Tables 2.2 and 2.3.

As noted in Chapter 3, the appropriateness of using a three-level model (rather than a two-level model) with this data was confirmed prior to proceeding with any further analysis. Results of this model structure comparison test are presented in Table 2.4, and a descriptive summary of the nested sample sizes at each level of the three-level structure is also provided in Table 2.5.

Full analytic sample (all students)

All model results produced in analysis of the full analytic sample are presented in Table 2.6.

Unconditional model

Results of the fully unconditional three-level model are presented in the first output column of Table 2.6. The predicted grand mean score for spring reading achievement across all students in the sample was 32.43, meaning that students assessed across all sites at the end of kindergarten, on average, answered approximately just over 32 questions correctly on the reading subtest. Covariance parameters provided in Table 2.6 indicate estimates of the proportion of unexplained variance in student spring reading levels remaining at each level of the model. Variance decomposition estimates (also in Table 2.6) are calculated directly from the covariance parameter estimates and are additionally reported for clarity, as they provide a useful percentage estimate of the proportion of the total unexplained variance remaining in the model that is attributable to each level. Looking at the variance decomposition estimates for the unconditional model, it is clear that the vast majority of variance in spring reading levels was associated with individual (between students) differences (74.59%), followed by variance associated with school level (between schools) differences (19.89%), and a comparatively small amount of remaining variance associated with teacher level (between teachers within schools) differences (5.52%). In the final sections of Table 2.6, two different model fit estimates are provided. The significant χ^2 likelihood ratio test result confirms that, even though the remaining variance is largely located at level 1, there are significant enough amounts at the other levels to warrant the use of the leveled model structure. The AIC estimate (see "*Akaike information criterion*" in Chapter 1, *Definition of terms*) provides an ongoing measure of comparative fit to gauge improvement across models.

Conditional models

Results of the level 1 conditional model are presented in the second output column of Table 2.6. The five theoretically specified student-level covariates added to build level 1 of the model were *Fall reading IRT score*, *time lapse between assessments*, *student gender*, *family SES level*, and *student race* (incorporated for better interpretation of results as the four individual variables: *White*, *African American*, *Hispanic* and *Asian*). Of these, six were confirmed to be significant predictors of spring student reading levels in this data (*Fall reading IRT score, time lapse between assessments, student gender, family SES level, African*

American and *Asian*). In general, beginning kindergarten with a higher fall reading IRT score, having a longer time lapse between fall and spring assessments, being female, having a higher family SES level and being of Asian race/ethnicity identification were all found to be significantly positively associated with spring reading scores, while being of African American race/ethnicity identification was found to be significantly negatively associated with spring reading scores. Being of White or Hispanic race/ethnicity identifications were found to not be significantly associated with spring reading levels. The addition of these eight student level covariates (accounted for primarily by the six significant ones) reduced the covariance parameter estimate of the total amount of unexplained variance remaining at level 1 from 80.27 in the unconditional model to 29.89 in the conditional level 1 model. Covariance parameter and associated variance decomposition estimates showed that a significant amount of the variance associated with the level 2 and level 3 random effects components were also reduced by the inclusion of these variables, indicating that these six significant student level covariates account for a large percentage of all variance in spring reading IRT scores. A marked reduction in the comparative AIC estimate from the unconditional to the level 1 model indicates an improved fit. The six significant level 1 covariates were retained in the model moving forward.

Results of the level 2 conditional model are presented in the third output column of Table 2.6. The six empirically specified teacher-level covariates added to build level 2 of the model were *class type, teacher age*, *total years of kindergarten teaching experience*, *total number of early childhood education courses taken, certification in early childhood education*, and *teacher gender*. Each of these variables were identified by the results of *RQ1* as representing significant areas of teacher-level difference between male and female teachers in the data, along with *teacher gender* itself which is focal to this study and tested here for any main effect. Of these, only one was confirmed to be a significant predictor of spring student reading levels in this data (*class type*). Attending full-day, rather than half-day, kindergarten classrooms was found to be a significantly positively associated with spring reading IRT scores. Neither *teacher gender*, nor any of the teacher-gender related characteristics identified in previous analysis were found to have any significant effect on student spring reading scores. The addition of these covariates made only a small reduction to the covariance parameter estimate associated with level 2 of the model, from 5.44 in the conditional level 1 model to 5.33 in the conditional level 2 model, and the comparative AIC estimate indicated a slight improvement in the overall model fit resulting from their addition. Only the significant level 2 covariate *class type* was retained moving forward.

Results of the level 3 conditional model are presented in the fourth output column of Table 2.6. The two theoretically and one empirically specified school-level covariates added to build level 3 of the model were *school percent free lunch eligibility*, *school percent minority enrollment* and *school urbanicity*, respectively. None of these three school-level covariates were found to be significantly associated with spring student reading levels in this data. The addition of these covariates made no reduction to the variance associated with the school level covariance estimate from the conditional level 2 model. None of these school level covariates were retained moving forward, meaning that any remaining between-schools variance in the full model will be interpreted as an unspecified random effects component.

Full model

The full model represents the addition of the predictor variable, *Teacher Gender x Student Gender* (a cross-level interaction effect between student gender and teacher gender), to the best fit conditional model. In this case, the best fit conditional model was the conditional level 2 model retaining only the significant covariate of *class type*. Results of the full model are presented in the final output column of Table 2.6 and indicated that *Teacher Gender x Student Gender* was not significantly associated with spring student reading levels in this data. The addition of this covariate made no reduction to the covariance parameter estimate associated with level 2 from the conditional model, and only a very small reduction in the comparative AIC estimate.

Figure 2 depicts a path diagram of the final full model detailed above and in the final column of Table 2.6. Figures 3 and 4 present a comparative look at actual and predicted spring reading IRT scores by fall entry scores for both male and female students by teacher gender. It is clear from both figures that reading growth over the kindergarten year looks similar for boys and girls regardless of teacher gender, and that the other significant covariates in the conditional model described here are highly predictive of spring reading levels, without any additional significant influence contributed by of teacher gender. Further interpretation and discussion of potential implications for these findings are discussed in Chapter 5.

Restricted analytic sample (boys only)

A second series of models was also run to consider the associations of these same teacher-gender related characteristics, along with additional indicators of student behavioral service and resident father statuses, among the population of male students alone.

Unconditional model

Results of the fully unconditional three-level model are presented in the first output column of Table 2.7. The predicted grand mean score for spring reading achievement among boys alone was 31.56, slightly lower than that of the full student sample with female students included. Covariance parameters and variance decomposition estimates provided in Table 2.7 indicate again that the vast majority of variance in spring reading levels among boys was associated with individual (between students) differences (75.5%), followed by variance associated with school level (between schools) differences (19.62%), and a comparatively small amount of remaining variance associated with teacher level (between teachers within schools) differences (4.88%). In the final sections of Table 2.7, the two different estimates of model fit are again provided and confirm again that, even though this remaining variance is largely located at level one, there are significant enough amounts at the other levels to confirm the appropriateness of the leveled model structure; and the unconditional model AIC estimate provides an ongoing measure of comparative fit to gauge improvement across models.

Conditional models

Results of the level 1 conditional model are presented in the second output column of Table 2.7. The six theoretically specified student-level covariates added to build level 1 of the model were *Fall reading IRT score*, *time lapse between assessments*, *family SES level*, *behavior services status*, *resident father status* and *student race* (incorporated for better interpretation of results as the four individual variables: *White*, *African American*, *Hispanic* and *Asian*). Of these, five were confirmed to be significant predictors of spring student reading levels in this data (*Fall reading IRT score*, *time lapse between assessments*, *family SES level*, *resident father status* and *Asian*). Among male students specifically, beginning kindergarten with a higher fall reading IRT score, having a longer time lapse between fall and spring assessments, having a higher family SES level, having some type of resident father (biological or adoptive) living at home, and being of Asian race/ethnicity identification were all found to be significantly positively associated with spring reading scores. Being of White, African American or Hispanic race/ethnicity identifications or being identified as receiving emotional/behavioral services in school were not found to be significantly associated with spring reading levels among boys in this data. The addition of these nine student level covariates (accounted for primarily by the five significant ones) reduced the covariance parameter estimate of the total amount of unexplained variance remaining at level 1 from 82.04 in the unconditional model to 29.87 in the conditional level 1 model. Covariance parameter estimates showed that a significant amount of the variance associated with the level 2 and level 3 random effects components were also reduced by the inclusion of these variables, indicating that these five significant student level covariates account for a large percentage of all variance in spring reading IRT scores among kindergarten boys. A marked reduction in the comparative AIC estimate from the unconditional to the level 1 model indicates an improved fit. The five significant level 1 covariates were retained in the model moving forward.

Results of the level 2 conditional model are presented in the third output column of Table 2.7. The six empirically specified teacher-level covariates added to build level 2 of the model were *class type*, *teacher age*, *total years of kindergarten teaching experience*, *total number of early childhood education courses taken*, *certification in early childhood education*, and *teacher gender*. Each of these variables were identified by the results of *RQ1* as representing significant, practical areas of difference between male and female teachers in the data, along with *teacher gender* itself which is focal to this study. As with the previous models, the only level 2 covariate confirmed to be a significant predictor of spring student reading levels among the restricted sample of boys alone was *class type*. Attending full-day, rather than half-day, kindergarten classrooms was again found to be a significantly positively associated with spring reading IRT scores. Also reflecting the results obtained with the full student sample, none of the teacher-gender related characteristics identified in prior analysis were found to have any significant effect on student spring reading scores in the restricted sample. The addition of these covariates made only a small reduction to the covariance parameter estimate associated with level 2 of the model, from 5.69 in the conditional level 1 model to 5.43 in the conditional level 2 model, and the comparative AIC estimate indicated a slight improvement in the overall model fit resulting from their addition. Only the significant level 2 covariate *class type* was retained moving forward.

Results of the level 3 conditional model are presented in the fourth output column of Table 2.7. The two theoretically and one empirically specified school-level covariates added to build level 3 of the model were *school percent free lunch eligibility*, *school percent minority enrollment* and *school urbanicity*, respectively. Again, none of these three school-level covariates were found to be significantly associated with spring student reading levels among the boys alone. The addition of these covariates made almost no reduction to the variance associated with the school level covariance estimate from the conditional level 2 model. None of these school level covariates were retained moving forward, meaning that any remaining between-schools variance pertaining to boys' spring reading levels in the full model will again be interpreted as an unspecified random effects component.

Full model

The full model here represents the addition of the predictor variable, *Teacher Gender x Resident Father Status* (a cross-level interaction effect between teacher gender and resident father status), to the best fit conditional model. (Because *behavior services status* was not significantly associated with reading outcomes in the conditional model, a further interaction here was not tested). As it was with the full student sample, the best fit conditional model for the restricted sample was again the conditional level 2 model retaining only the significant covariate of *class type*. Results of the full model are presented in the fifth column of Table 2.7 and indicated that *Teacher Gender x Resident Father Status* was not significantly associated with spring student reading levels among boys specifically in this data. The addition of this covariate made no reduction to the covariance parameter estimate associated with level 2 from the conditional model, and only a very small reduction in the comparative AIC estimate.

Figure 5 depicts a path diagram of the final full model detailed above and in the final column of Table 2.7. Figures 6 and 7 present a comparative look at actual and predicted spring reading IRT scores by fall entry scores for male students specifically by teacher gender. It is once again clear from both figures that reading growth over the kindergarten year looks similar for boys with male and boys with female teachers, and that the significant covariates in the conditional model described here are highly predictive of spring reading levels, without any additional significant influence contributed by teacher gender. Further interpretation of these and all other findings presented in this chapter are discussed in Chapter 5.

CHAPTER FIVE

DISCUSSION

Main Findings

Male kindergarten teachers were found to be significantly different from female kindergarten teachers at the population level with regard to a small number of additional characteristics. Men tended to be younger (on average), less experienced, to have less formal training in early childhood education, and to be more likely to teach in either rural or urban schools and in half-day instructional settings. Population trends between male and female kindergarten teachers were not significantly different for any other areas of demographic characteristics or qualifications, and there was no evidence of significant differences in teaching styles used during reading/language arts instruction.

Teacher gender itself was not found to be significantly associated with spring reading levels in kindergarten for either boys or girls. The most significant predictors of spring reading levels in this data were fall reading score, time lapse between assessments, student gender (in the full sample only), student race/ethnicity identifications of African American (in the full sample only) and Asian, family SES, class type, and resident father status (in the restricted sample only). Teacher gender, either alone or in interaction with student gender or student characteristics, did not demonstrate any significant association with changes in predicted student end of kindergarten reading levels.

RQ1: Gender-related differences between male and female teachers

Demographic characteristics

Male kindergarten teachers were found to have a mean age of just under 38 years old, which was significantly younger than the mean age of female teachers at just over 41. This finding did not meet the expectation of the working hypothesis, which was based on large-scale findings that indicated the average age of male teachers across all grade levels is slightly older than the average age of female teachers (Snyder, Dillow & Hoffman, 2008), as well as on smaller, direct studies of male elementary school teachers which observed that at this level tended to be slightly older on average than female elementary school teachers (Smith, Kirchner, Taylor, Hoffman & Lemke, 1998) and tended to enter teaching relatively later in life than females (Brookhart & Loadman, 1996; Gross & Trask, 1976) after often having spent a considerable amount of time pursuing other career options (DeCorse & Vogtle, 1997; Freidus, 1992). However a broader look at the categorical age distributions of male and female teachers suggests that perhaps these latter findings aren't entirely contradicted here, as there was evidence of a significant spike among male teachers between the ages of "39-43". Additionally, extended historical accounts of male teachers in education show that teaching has traditionally been viewed by men as a stepping stone job useful for financing one's way through school or supporting oneself in young adulthood while pursuing another career path (Rury, 1989; Sugg, 1978). Perhaps the two age-range spikes among male kindergarten teachers found in this data (at age ranges "24-28" and "39-43") indicates some evidence that the early career and the second or late career male teacher described by these different bodies of research may both still characterize a large portion of the male kindergarten teaching population today, and that the first working hypothesis for this study miscalculated by failing to take further into account the historical accounts of early career male teachers.

Racial distributions among male and female kindergarten teachers was almost uniform across all categories, with both groups identifying as approximately 85% White and approximately 6-7% African American. This disproportionate ratio is similar to what has been reported for the racial composition of the overall public teaching force in the over the past few decades (Shen, Wegenke & Cooley, 2003) and most recent estimates of the aggregate kindergarten teacher population (NCES, 2004). Evidence put forth by historical researchers that teaching may have been viewed, historically, as a more attractive career option to racial minority men than white men (Carter, 1989) appear to no longer be the case at least in this data. One possibility is that teaching, at the kindergarten level specifically, is considered equally gender-role inappropriate to men across race/ethnicity groups and that distribution patterns may be different at the upper grades levels. However this would not explain why the same racial imbalance is mirrored among the female kindergarten teacher distributions. More likely, these findings suggest that kindergarten teachers are nearly as overwhelmingly racially homogeneous as they are overwhelmingly female, and that this may be an important area for future research.

Qualifications

As hypothesized, male and female teachers did demonstrate a significant difference in the number of early childhood education (ECE) courses taken. Male teachers were most likely to have taken between zero and two ECE courses (49%), with the remainder of men split approximately evenly between those having taken three to four courses (24%) and those having taken five or more (26%). For female teachers this distribution was almost entirely reversed, with only 26% having taken between zero and two ECE courses, 18% having taken three to four courses, and 56% having taken five or more. The association between these differences and teacher gender remained significant even after controlling for the potential contribution of age and race. These findings are consistent with a number of indicators provided by small-scale studies of male teachers and teacher candidates which suggested that these men were more likely to have majored in a more traditionally masculine undergraduate program before deciding to teach later (DeCorse & Vogtle, 1997; Freidus, 1992; Poll, 1979; Sargent, 2001) and that men were likely to have a lower regard for coursework related to a teacher education program (Brookhart & Loadman, 1996). However this reasoning may not entirely explain the differences found in male and female educational backgrounds, as the same reasoning proved not to be supported when it came to the working hypotheses about total number of elementary education and reading methods courses taken. One possible explanation is that early childhood education may be an especially gender-typed field, even compared to elementary education, though other factors unavailable in this data (such as teacher personal beliefs, variations in state certification requirements of kindergarten teachers, etc.) make explanations beyond speculation of this kind unjustifiable at this point.

Even though there was not sufficient literature available prior to this study to warrant a specific hypothesis about certification differences, it makes sense that, given the significant difference in total amount of ECE courses taken by male and female teachers, a significant difference was also found indicating a lower percentage of males (33%) than females (53%) who held specific certification in the area of early childhood education. While these differences are important to be aware of for considering potential ramifications regarding the use of developmentally appropriate practice and student outcomes at the early childhood level (McCullen & Kazat, 2002), what exactly this means in terms of male and female kindergarten teachers being differently qualified per se is more difficult to untangle. One major caveat in such discussions is that specific teacher credentialing requirements vary from state to state, meaning that while a PreK-2 certification (with a coursework emphasis on early childhood education) may be the standard requirement to teach kindergarten in one state, a more general K-6 certification (with coursework requirements focused more broadly on elementary education) may be regarded as the standard equivalent in another state. In cases such as this, it would be inaccurate to proclaim that men or women meeting the coursework and certification requirements expected of them by different governing bodies are in any definitive way more or less qualified than the other. However, pending a more specific examination that can account for such varying standards, these differences in coursework and certification may still provide some degree of tentative support for those existing studies which suggest that male teachers of young children are more likely to have pursued more traditionally masculine coursework prior to entering teaching.

There were no other educational background-related variables for which male and female teachers in this data demonstrated a significant difference at the population level.

Three significant population-level differences between male and female teachers pertaining to actual job experience were also found. The mean number of years teaching experience at the kindergarten level for men (4.54 years) was just over half the mean amount reported by women (8.24 years). This was not a variable for which there was enough previously available evidence to establish a working hypothesis, but the outcome does appear to be sensibly in line with the age differences found between male and female teachers earlier. If we accept the conclusions of other studies pertaining to gender differences in paths towards teaching (DeCorse & Vogtle, 1997; Freidus, 1992; Poll, 1979; Sargent, 2001), which suggest that a greater proportion of females than males viewed teaching itself as their original career of choice, it would stand to reason that the total number of years experience would be less for men as a result of both those men who entered the profession early and moved on after a few years and those who began teaching later after spending several years pursuing other options.

Significant differences were also found in the size and location of schools in which male kindergarten teachers were located. The men in this study were most likely to be teaching in high population, urban schools. Over half of the men in the population (51%) taught in schools with a total student population of 500 or more, compared to 44% of the female teacher population working in such schools. However this significant simple difference was found to be no longer related to gender after accounting for the influence of teacher race/ethnicity. Similarly, more than half of the male teachers were teaching in central city (urban) areas (57%), followed next by those teaching in small town (rural) areas (34%) and just 9% in urban fringe (suburban) areas. Female teachers were relatively more evenly distributed across locations, with 48% working in central city schools, 29% in urban fringe schools, and 23% in small town schools. Even after controlling for age and race these differences were found to be significantly associated with teacher gender, with the largest difference being clearly the greater proportion of female teachers working in suburban schools. While there was again no directly applicable literature about the location of teachers by gender previously in order to establish expectations for these variables, there are a number of potential possibilities to consider regarding why these differences may exist and that may warrant in terms of future research. It may be, for example, that a relatively greater need for teachers in urban (due to higher student concentrations) and rural (due to fewer available candidates) compels schools in those area to cast a wider net in their teacher recruitment efforts than suburban areas. Another possibility may be that the relatively higher cost of living in suburban areas is a greater deterrent to teaching among those males who feel a level of gender role-related social pressure to produce a higher level of income than teaching could provide. Again, these are speculative possibilities for explaining the current results, and should not be considered interpretations based on the data but rather suggestions for possible further investigations of the data.

The final job experience variable found to be significantly different between the male and female populations was that the men were significantly more likely to teach only a half-day (AM or PM only) classroom (49%) than women (22%). The vast majority of women (78%) taught either all-day, combination half-day/all-day, or both AM and PM classrooms, compared with just 51% of men who did. Prior to this study there was no directly applicable information available about the gender distribution of kindergarten teachers in different instructional settings, however the findings of a significant difference in this category are particularly interesting when considered in combination with the difference found between male and female kindergarten teachers' total years of teaching experience. Considering that men report having about half the number of years experience as women on average, and additionally now that they are much more likely to only teach half-day settings, it is possible to hypothesize that if a metric were designed to combine the data from these two survey items into a single measure of the total amount of actual classroom teaching experience accumulated over time (measured in hours or minutes), that we could expect the mean differences in total accumulated experience between male and female kindergarten teachers to be that much more compounded.

Instructional Practices

There were no significant differences found between the reported use of instructional practices pertaining to reading/language arts instruction between male and female teachers. Male teachers, on average, reported using phonics-related activities most frequently in their instruction, with a mean reported instructional frequency of approximately three to four times per week. Reading and writing activities and comprehension-related activities were the next most used, each with mean reported instructional frequencies in the range of one to two times per week. Mixed-ability grouping practices, didactic instruction-based activities, student-centered instruction practices, and reading and writing skills lessons were each found to have mean reported instructional frequencies in the range of two to three times per month, with the mean reported frequency for reading and writing skills lessons being the lowest among that group. The mean reported instructional frequencies of female teachers did not differ significantly from the males for any of these practices and, accordingly, mirrored the same order from most frequently (phonics-related activities) to least frequently (reading and writing skills) used practices. The widest margin of difference between male and female teachers was in the mean reported use of comprehension-related activities, which female teachers reported incorporating more frequently than men, though the difference was not significant.

These findings do not support the working hypothesis that men would use more didactic and fewer student-centered instruction approaches. Observational evidence of the expected patterns of male teacher instructional practice, especially with reference to increased use of didactic instruction methods (Good, Sikes & Brophy, 1973; Brophy, 1985; Freidus, 1992), as well as subject and observer reports of male teachers being forced to operate in more structured, less interactive instructional modes out of fear of misinterpreted contact (Allan, 1994; DeCorse & Vogtle, 1997; King, 1998; Sargent, 2001; Williams, 1995) were not supported in this data. Findings indicated much stronger support for studies concluding that male and female teachers generally tend to be more alike in their classroom behaviors than they are different (Ruble, Martin & Berenbaum, 2006; Sabbe & Aelterman, 2007).

Still, considering the highly consistent agreement in the literature that men at this level are under significant social pressure to modify those aspects of their instructional behavior pertaining to didactic versus student-centered instructional styles, it is intriguing that this would be the area in which there appears to be the least amount of difference between the two groups. There may be several larger reasons for this evident discrepancy between male teachers experience and the apparent uniformity of teaching styles found here. For one, it may be that the men interviewed for the studies cited here were prone to the common habit of conflating specific instructional practices with general interaction styles. Given the fluidity of a kindergarten or early elementary grade classroom, it is easy to understand how the researchers observing these men might have a difficult time separating the two types of behavior as well, whereas the ECLS-K teacher questionnaire remains strictly focused on specific instructional activities alone. Another possibility for further consideration may be that the adoption of scripted curriculum programs by some schools in this data effectively minimized the amount of individual choice that teachers had over which specific instructional practices they would use to teach critically tested subject areas such as reading/language arts and mathematics. For the purpose of this study, however, the major conclusion to be drawn is that there is no evidence of any significant teacher gender difference in use of these specific instructional practices.

RQ 2: Teacher gender and student achievement

The second part of this study was designed to consider the potential influence of teacher gender and teacher gender-related characteristics on measurable student achievement levels in reading. This influence was considered first in the general kindergarten student population and then a second time in a restricted population of only kindergarten boys. The research questions guiding this part of the investigation focused on whether or not teacher gender, alone or interaction with student gender, made any significant reduction in the between-class variance of spring reading levels for either of these populations after accounting for the influence of other established covariates. In the restricted student sample, the interaction of teacher gender and the popularly cited risk factor of resident father status was also tested for significant association in the model.

Full analytic sample (all students)

In the full student sample there was no significant association between teacher gender and spring student reading scores among boys or girls. The only teacher gender-related characteristic that was significantly associated with student reading scores was *class type* (half-day vs. full-day instructional setting). Students attending full-day kindergarten performed significantly better on spring reading assessments than students attending half-day kindergarten. Clearly the association between *class type* as a teacher gender-related characteristic and student achievement is a precarious connection to put forward without an explicit reiteration of what the term “teacher gender-related characteristic” means in the context of this study. This simply means that a significant difference was established to exist between the proportion of male teachers and the proportion of female teachers in this sample who teach in half-day versus full-day settings, with female teachers being relatively much more likely to teach in full-day classrooms than men. The reasonable conclusion to be drawn here is not that this finding therefore constitutes a “gender effect”, but more likely that it is the increased instructional time that is the significant factor in student score differences. As an issue of gender research, the more appropriate question to take away from this finding would be whether or not there is a reason that male teachers seem more likely to teach in half-day environments, which clearly offer their students much less opportunity to achieve at the same levels as full-day kindergarteners, as opposed to pursuing what this means about any association between teacher gender and student achievement.

All of the other significant influences identified in this sample occurred at the student level, meaning that all of the remaining identifiable variance in student scores occurred at the individual level. *Fall entry score*, *time lapse between assessments*, *student gender*, *student race/ethnicity*, and *family SES level* were all confirmed to be highly significant influences. Student gender maintained a significant negative association with spring reading levels throughout the models, a finding that reaffirms the conclusions of other ECLS-K-based studies that boys spring reading levels were significantly lower than girls as a group, and confirms that teacher gender had no significant impact on the nature of that gap.

Restricted analytic sample (boys-only)

In the restricted student student sample there was again no significant association between teacher gender and spring student reading scores. Also as was found in the full student sample, the only teacher gender-related characteristic that was significantly associated with student reading scores was *class type*, presumably for the same reasons as previously raised. All of the remaining significant influences identified with this sample again occurred at the student level, with the student level covariates of *fall entry score*, *time lapse between assessments*, *student race/ethnicity*, and *family SES level* again accounting for the vast majority of variance in the model, as well as *resident father status* additionally demonstrating a significant association among the boys.

The addition of the interaction effect between *Teacher gender x Father status* to the model was not significant, nor did it diminish the significance of *resident father status* at all. This finding indicates that having any type of resident father living at home, biological or adoptive, was found to significantly impact boys' academic achievement levels to a degree that having a male teacher instead could not diminish or make up for. Even though the importance of resident father status was only tested with the boys-only sample, it is important to consider again that the underlying explanation of the significance of this variable may not be specifically gender-related. It may more likely be the case that this variable is operating as a proxy for two-parent households, in which case the more tangible benefits associated with school achievement may not have anything to do with the boy's or father's gender per se but with the benefits provided by a two-parent home, such as a potentially higher family SES level and increased opportunities for interactions with, supervision by and/or guidance from at least one of the two parents at any given time, as opposed to the financial and time-related burdens more commonly associated with single-parent households.

Overall, the findings of this study did not support any tested aspect of the popular claims that male teachers would be especially helpful for boys, and especially for without fathers at home or with ongoing behavior issues in school, at least in terms of narrowing the achievement gap in reading between boys and girls during the first year of school.

Limitations and Future Research

Throughout this study, the most apparent limitation has been the overwhelming difference in sample sizes between male and female teachers and, consequently, between students who have male teachers and students who have female teachers. While it is acknowledged that this is clearly not ideal and underscores the need to interpret any findings with careful consideration, the first point to reiterate in addressing this concern is that the stated goal of this study was to consider the phenomena of interest as they actually exist at the population level, something that had not ever been done before in addressing these specific questions and was much needed. By using the entire teacher and student populations available this study has provided an important perspective for informing discussions about male teachers and students in education, which until now have been largely driven by personal and local experiences being used to inform a national perspective.

At the same time, it is also important to acknowledge that strategically limiting samples such as this one in future research may be beneficial for more closely examining a number of specific questions that still remain. Future research may benefit from restricting the sample in ways that would minimize the teacher-level imbalance and potentially illuminate important patterns of association not detectable at the population level. For example, selecting a fixed site-level sample consisting of from only those schools in which a male teacher is located. This approach would require sacrificing the ability to project findings beyond those specific schools, but has the potential to provide greater insight into any possible interactions between teacher and student gender without as much risk of the presence of one group overwhelming the influence of the other in the data. Other areas of further research on male teachers and students also still remain in high need with regard to other possible methods of influence and outcomes commonly attributed to the benefits of the male teachers for boys. Defining and measuring “role modeling effects”, examining associations of teacher gender or teacher gender-related characteristics with other outcomes such as student motivation, long-term achievement, or short-term achievement in other critical subject areas such as mathematics, all remain highly speculative areas of interest in the public discourse that need to be further addressed by the academic literature.

Table 1.1a: Variables and data sources for Study Part 1 (1 of 2)

Variable Name (Analysis ID)	ECLS-K Instrument Source	ECLS-K Variable ID
Demographic Characteristics		
Teacher gender (GENDER)[a]	Fall-kindergarten teacher questionnaire B	B1TGEND
Teacher race (RACE)[a]	Fall-kindergarten teacher questionnaire B	B1RACE1, B1RACE2, B1RACE3, B1RACE4, B1RACE5
Teacher age (continuous) (B1AGE)	Fall-kindergarten teacher questionnaire B	B1AGE
Teacher age (categorical) (AGE)[a]	Fall-kindergarten teacher questionnaire B	B1AGE
Qualifications (Educational Background)		
Highest level of education received (B1HGHSTD)	Fall-kindergarten teacher questionnaire B	B1HGHSTD
Total number of early childhood courses taken (B1EARLY)	Fall-kindergarten teacher questionnaire B	B1EARLY
Total number of elementary courses taken (B1ELEM)	Fall-kindergarten teacher questionnaire B	B1ELEM
Total number of reading methods courses taken (B1MTHDRD)	Fall-kindergarten teacher questionnaire B	B1MTHDRD
Certification type (B1TYPCER)	Fall-kindergarten teacher questionnaire B	B1TYPCER
Certification in early childhood education (B1ERLYCT)	Fall-kindergarten teacher questionnaire B	B1ERLYCT
Certification in elementary education (B1ELEMCT)	Fall-kindergarten teacher questionnaire B	B1ELEMCT
Certification in other field of education (B1OTHCRT)	Fall-kindergarten teacher questionnaire B	B1OTHCRT
Qualifications (Job Experience)		
Total years of kindergarten teaching experience (B1YRSKIN)	Fall-kindergarten teacher questionnaire B	B1YRSKIN
Class type (KGCLASS)	ECLS-K created composite	KGCLASS
School type (S2KPUPRI)	Spring-kindergarten school administrator questionnaire	S2KPUPRI
School urbanicity (KURBAN)	ECLS-K created composite	KURBAN
Total school enrollment (S2KENRLS)	Spring-kindergarten school administrator questionnaire	S2KENRLS
Total school minority enrollment (S2KMINOR)	Spring-kindergarten school administrator questionnaire	S2KMINOR
Classroom instructional practices		
Reading and Writing Activities (ReadWriteActMean)[b]	Spring-kindergarten teacher questionnaire A	A2SILENT, A2INVENT, A2CHSBK, A2COMPOS, A2JRNL, A2DICTAT, A2DOPROJ, A2PUBLSH
Phonics (PhonicsMean)[b]	Spring-kindergarten teacher questionnaire A	A2LERNLT, A2PRACLT, A2NEWVOC, A2PHONIC, A2CONVNT, A2RCGNZE, A2MATCH, A2WRTNME, A2RHYMNG
Didactic Instruction (DidacticMean)[b]	Spring-kindergarten teacher questionnaire A	A2BASAL, A2WRKBK, A2WRTWRD, A2READLD

[a]Recoded variable. See Appendix A for full detail of data preparation procedures. [b] Recoded composite variable. See Appendices A and C for full details.

Table 1.1b: Variables and data sources for analysis Study Part 1 (2 of 2)

Classroom instructional practices (cont.)		
Comprehension (ComprehensionMean)[b]	Spring-kindergarten teacher questionnaire A	A2PREPOS, A2MAINID, A2PREDIC, A2TEXTCU, A2ORALID, A2DRCTNS
Student Centered Instruction (StudCentInstMean)[b]	Spring-kindergarten teacher questionnaire A	A2RETELL, A2SKITS
Reading and Writing Skills (ReadWriteSkillMean)[b]	Spring-kindergarten teacher questionnaire A	A2SYLLAB, A2PNCTUA, A2COMPSE, A2WRTSTO, A2SPELL, A2VOCAB, A2ALPBTZ, A2RDFLNT
Mixed Ability Grouping (MixedAbilGrpMean)[b]	Spring-kindergarten teacher questionnaire A	A2MXDGRP, A2PRTUTR
Survey Design Weights		
Base year, teacher-level full sample weight (B1TW0)	ECLS-K designed sample weight	B1TW0
Base year, teacher-level stratum nesting weight (B1TTWSTR)	ECLS-K designed sample weight	B1TTWSTR
Base year, teacher-level PSU nesting weight (B1TTWPSU)	ECLS-K designed sample weight	B1TTWPSU

[a]Recoded variable. See Appendix A for full detail of data preparation procedures. [b] Recoded composite variable. See Appendices A and C for full details.

Table 1.2: Simple mean comparison of teacher (continuous) age by gender

	Mean	S.D.	95% Confidence Interval for Mean		Adjusted Wald Test	
			Lower	Upper	*F*	Pr > *F*
Age						
Male	37.56	1.59	34.44	40.67		
Female	41.24	.25	40.74	41.74		
					5.42	0.02*

* p < .05.

Table 1.3: Frequency distributions of teacher demographic characteristics by gender

	Gender							
	Male		Female		Combined			
	Sample	Pop.	Sample	Pop.	Sample	Pop.	*F*	Pr > *F*
Age								
24 - 28	18	1,157	430	26,638	448	27,795		
29 - 33	11	469	389	22,429	400	22,898		
34 - 38	7	358	335	21,491	342	21,849		
39 - 43	15	1,014	362	23,151	377	24,165		
44 - 48	8	591	513	33,814	521	34,406		
49 - 53	10	501	499	31,502	509	32,003		
54 - 58	3	107	323	19,422	326	19,529		
Total	*72*	*4,197*	*2,934*	*183,230*	*3,006*	*187,427*		
							2.19	0.046*
Race/Eth.								
No response	9	379	142	6,049	151	6,427		
NA/PI/AS[a]	3	65	116	6,233	119	6,298		
White/Non-His.	56	3,496	2,470	158,923	2,526	162,419		
Afr.-Amer.	4	257	206	12,025	210	12,283		
Total	*72*	*4,197*	*2,934*	*183,230*	*3,006*	*187,427*		
							2.12	0.11

[a]Category created in this analysis reflecting response options “Native American/Pacific Islander” and “Asian” combined.
* p < .05.

Table 1.4a: Frequency distributions of teacher educational background by gender (1 of 2)

	Gender							
	Male		Female		Combined			
	Sample	Pop.	Sample	Pop.	Sample	Pop.	*F*	Pr > *F*
High educ. rec.								
HS/assoc./bach.	21	1,388	781	53,147	802	54,534		
Some post-bach.	28	1,603	937	57,924	965	59,527		
Master's	14	696	867	50,170	881	50,866		
Ed. sp./prof. deg.	3	170	156	9,793	159	9,963		
Doctorate	0	0	4	469	4	469		
Total	*66*	*3,856*	*2,745*	*171,503*	*2,811*	*175,358*		
							.605	.636
# ECE courses								
0	7	537	223	15,464	230	16,001		
1	7	288	179	11,332	186	11,620		
2	12	839	293	17,882	305	18,722		
3	8	379	252	16,764	260	17,144		
4	9	440	228	13,585	237	14,025		
5	3	79	122	7,230	125	7,309		
6+	16	807	1,433	88,021	1,449	88,828		
Total	*62*	*3,370*	*2,730*	*170,278*	*2,792*	*173,648*		
							2.97	0.015*
# Elementary courses								
0	5	272	81	7,741	86	8,012		
1	2	61	52	3,455	54	3,516		
2	3	291	81	5,206	84	5,497		
3	3	127	100	6,096	103	6,224		
4	9	331	108	6,074	117	6,405		
5	2	144	94	6,738	96	6,882		
6+	43	2,396	2,225	135,611	2,268	138,007		
Total	*67*	*3,622*	*2,741*	*170,921*	*2,808*	*174,543*		
							1.645	0.143
# Reading method courses								
0	6	240	69	6,361	75	6,601		
1	10	558	441	26,485	451	27,044		
2	21	1,395	680	42,763	701	44,158		
3	13	555	524	32,209	537	32,764		
4	11	634	293	17,776	304	18,411		
5	2	93	146	8,303	148	8,395		
6+	5	164	597	36,641	602	36,804		
Total	*68*	*3,639*	*2,750*	*170,538*	*2,818*	*174,176*		
							2.44	0.034*

Note. Totals sum to different amounts on some measures due to variations in item non-response.
* $p < .05$

Table 1.4b: Frequency distributions of teacher educational background by gender (2 of 2)

	Gender							
	Male		Female		Combined			
	Sample	Pop.	Sample	Pop.	Sample	Pop.	*F*	Pr > *F*
Certif. type								
None	2	189	52	5,551	54	5,740		
Temporary	14	630	286	16,429	300	17,060		
Alternate	5	227	56	4,858	61	5,086		
Standard	15	1,005	595	39,188	610	40,193		
Highest	34	1,783	1,832	109,408	1,866	111,190		
Total	*70*	*3,835*	*2,821*	*175,434*	*2,891*	*179,269*		
							1.467	.214
EC certified								
Yes	17	1,260	1,559	94,357	1,576	95,617		
No	51	2,542	1,282	82,301	1,333	84,843		
Total	*68*	*3,801*	*2,841*	*176,659*	*2,909*	*180,460*		
							5.62	.018*
Elem. certified								
Yes	60	3,433	2,415	146,392	2,475	149,825		
No	8	368	414	2,9026	422	29,394		
Total	*68*	*3,801*	*2,829*	*175,418*	*2,897*	*179,219*		
							1.93	.166
Other certified								
Yes	21	937	773	48,290	794	49,227		
No	51	3,260	2,141	133,779	2,192	137,039		
Total	*72*	*4,197*	*2,914*	*182,069*	*2,986*	*186,266*		
							.545	0.461

Note. Totals sum to different amounts on some measures due to variations in item non-response.
* p < .05

Table 1.5: Simple mean comparison of total years kindergarten teaching experience by gender

	Mean	S.D.	95% Confidence Interval for Mean		Adjusted Wald Test	
			Lower	Upper	*F*	Pr > *F*
Years. K. exp.						
Male	4.54	.73	3.11	5.98		
Female	8.24	.164	7.92	8.56		
					25.87	0.00**

**p <.01.

Table 1.6: Frequency distributions of teacher job experience by gender

	Gender				Combined			
	Male		Female					
	Sample	Pop.	Sample	Pop.	Sample	Pop.	*F*	Pr > *F*
Class type								
AM only	22	1,376	439	28,699	461	30,075		
PM only	13	669	237	11,585	250	12,253		
AM & PM	7	423	471	29,564	478	29,987		
All day	30	1,729	1,779	112,502	1,809	114,231		
AM & all day	0	0	6	748	6	748		
PM & all day	0	0	2	132	2	132		
Total	*72*	*4,197*	*2,934*	*183,230*	*3,006*	*187,427*		
							3.04	0.015*
School type								
Public	66	3,376	2,578	145,925	2,644	149,301		
Private	6	822	356	37,305	362	38,127		
Total	*72*	*4,197*	*2,934*	*183,230*	*3,006*	*187,427*		
							0.01	0.92
Urbanicity								
Central city	47	2,396	1,516	88,419	1,563	90,815		
Urban fringe	11	393	886	53,324	897	53,717		
Small town	14	1,408	532	41,487	546	42,895		
Total	*72*	*4,197*	*2,934*	*183,230*	*3,006*	*187,427*		
							5.01	0.009**
Total enrollment[a]								
0 - 149	7	1,160	146	21,462	153	22,622		
150 - 299	7	410	326	31,128	333	31,538		
300 - 499	5	422	700	47,634	705	48,057		
500 - 749	26	1,232	951	50,633	977	51,864		
750 +	26	823	793	30,876	819	31,699		
Total	*71*	*4,047*	*2,916*	*181,732*	*2,987*	*185,779*		
							3.32	0.013*
% Minor. enrol.[a]								
<10%	12	1,142	698	52,115	710	53,256		
10% - <25%	6	552	473	32,636	479	33,188		
25% - <50%	7	314	516	32,407	523	32,721		
50% - <75%	10	347	366	20,170	376	20,518		
>75%	36	1,692	826	41,696	862	43,388		
Total	*71*	*4,047*	*2,879*	*179,024*	*2,950*	*183,071*		
							2.14	0.09

[a]Data associated with teachers in catalog but taken from administrator questionnaire responses, thus reflecting a different amount of missing data.
* p < .05. **p <.01.

Table 1.7: Classroom instructional practice simple mean comparisons of teachers by gender

	Mean	L.S.E.	95% Confidence Interval for Mean		Adjusted Wald Test	
			Lower	Upper	*F*	Pr > *F*
Read/writ. activities.						
Male	3.05	.13	2.79	3.32		
Female	2.92	.04	2.84	2.99		
					0.96	0.33
Phonics						
Male	4.4	.09	4.23	4.58		
Female	4.49	..03	4.44	4.54		
					0.83	0.36
Didactic						
Male	2.67	.22	2.23	3.1		
Female	2.52	.04	2.44	2.6		
					0.42	0.52
Comprehension						
Male	3.19	.3	2.59	3.78		
Female	3.63	.03	3.57	3.69		
					2.03	0.15
Student centered inst.						
Male	2.28	.11	2.06	2.5		
Female	2.23	.03	2.17	2.29		
					0.19	0.67
Read/writ. skills						
Male	2.26	.26	1.75	2.78		
Female	2.11	.5	2.01	2.2		
					0.33	0.57
Mixed ability group.						
Male	2.85	.27	2.32	3.39		
Female	2.92	.04	2.84	2.99		
					.05	.82

Table 1.8: Regression of significant educational background qualifications on age, race and gender

	Coef.	SE	*t*	P > |*t*|	95% CI Lower	95% CI Upper
Early childhood courses[a]						
Age	0.012	0.005	2.69	0.007*	0.003	0.021
Race	0.024	0.061	0.39	0.698	-0.096	0.143
Gender	-0.878	0.274	-3.21	0.001**	-1.416	-0.34
Reading method courses[a]						
Age	0.032	0.004	7.64	0.00**	0.024	0.041
Race	0.005	0.071	0.07	0.942	-0.135	0.145
Gender	-0.423	0.235	-1.8	0.072	-0.885	0.039
Early childhood cert.[b]						
Age	-0.001	0.005	-0.21	0.835	-0.011	0.009
Race	-0.133	0.058	-2.28	0.023*	-0.247	-0.018
Gender	0.821	0.363	2.26	0.024*	0.108	1.534

Note. Table contains regression results only for variables determined to have a simple significant association with gender in prior analysis.
[a]Conducted by ordered logistic regression (for ordered categorical variables). [b]Conducted by multinomial logistic regression (for unordered categorical variables).
* p < .05. **p <.01.

Table 1.9: Regression of significant job experience qualifications on age, race and gender

	Coef.	SE	*t*	P > |*t*|	95% CI Lower	Upper
Years K Experience						
Age	0.406	0.015	26.35	0.000**	0.376	0.437
Race	0.687	0.171	4.03	0.000**	0.352	1.023
Gender	-1.988	0.469	-4.24	0.000**	-2.91	-1.065
Class type: AM only[a]						
Age	0.01	0.006	1.65	0.099	-0.002	0.023
Race	0.033	0.112	0.29	0.771	-0.188	0.253
Gender	1.176	0.436	2.69	0.007**	0.318	2.034
Class type: PM only[a]						
Age	-0.01	0.01	-1.03	0.305	-0.03	0.009
Race	-0.21	0.105	-2.00	0.046*	-0.417	-0.004
Gender	1.25	0.427	2.92	0.004**	0.41	2.089
Class type: AM + All day[a]						
Age	0.023	0.05	0.47	0.641	-0.075	0.121
Race	16.49	0.678	24.29	0.000**	15.14	17.8
Gender	-28.64	0.593	-48.32	0.000**	-29.8	-27.47
Class type: PM + All day[a]						
Age	-0.08	0.084	-0.96	0.339	-0.245	0.085
Race	16.19	1.057	15.32	0.000**	14.11	18.27
Gender	-28.72	0.928	-30.93	0.000**	-30.55	-26.9
Urban fringe location[a]						
Age	-0.001	0.007	0.21	0.835	-0.015	0.012
Race	0.378	0.087	4.33	0.000**	0.206	0.549
Gender	-1.29	0.436	-2.95	0.003**	-2.14	-0.43
Total school enrollment[b]						
Age	-0.006	0.004	-1.27	0.205	-0.014	0.003
Race	-0.244	0.07	-3.49	0.001**	-0.382	-0.106
Gender	-0.211	0.4	-0.53	0.593	-0.988	-0.565

Note. Table contains regression results only for variables determined to have a simple significant association with gender in prior analysis.
[a]Conducted by multinomial logistic regression (for unordered categorical variables). [b]Conducted by ordered logistic regression (for ordered categorical variables).
* p < .05. **p <.01.

Table 2.1a: Variables and data sources for Study Part 2 (1 of 2)

Variable Name (Variable ID)	**ECLS-K Instrument Source**	**ECLS-K Variable ID**
Student Level Variables		
Student gender (STGENDER)[a]	ECLS-K created composite	GENDER
Fall IRT Reading Score (C1RSCALE)	Fall-kindergarten direct child assessment	C1RSCALE
Time between assessments (TIMELAPSE)[a]	Fall-kindergarten direct child assessment; spring-kindergarten direct child assessment	R1_KAGE, R2KAGE
Stud. Race: White (STRACEWH) [a]	ECLS-K created composite	RACE
Stud. Race: Afr. Am (STRACEAA) [a]	ECLS-K created composite	RACE
Stud. Race: Hisp. (STRACEHS) [a]	ECLS-K created composite	RACE
Stud. Race: Asian (STRACEAS) [a]	ECLS-K created composite	RACE
Stud. Race: Other (STRACEOT) [a]	ECLS-K created composite	RACE
Student family SES (WKSESL)	ECLS-K created composite	WKSESL
Resident father status (RESFATH)[a]	Fall-kindergarten parent interview	P1HDADTYP
Behavior/Emotional services status (BEHSER)[a]	Spring-kindergarten teacher questionnaire C	T2BEHPRB
Classroom Level Variables		
Teacher ID (NT2_ID) [a]	ECLS-K assigned identification variable	T2_ID
Teacher gender (TRGENDER)[a]	Fall-kindergarten teacher questionnaire B	B1TGEND
Class type (CLASSLENGTH2)[a]	Spring-kindergarten field management system	F2CLASS
Teacher age (continuous) (B1AGE)	Fall-kindergarten teacher questionnaire B	B1AGE
Total years of kindergarten teaching experience (B1YRSKIN)	Fall-kindergarten teacher questionnaire B	B1YRSKIN
Total number of early childhood courses taken (B1EARLY)	Fall-kindergarten teacher questionnaire B	B1EARLY
Certification in early childhood education (B1ERLYCT)	Fall-kindergarten teacher questionnaire B	B1ERLYCT
School Level Variables		
School ID (NS2_ID)[a]	ECLS-K assigned identification variable	S2_ID
Percent free lunch eligible (S2KFLNCH)	Spring-kindergarten school administrator questionnaire	S2KFLNCH
Percent minority enrollment (S2KMINOR)	Spring-kindergarten school administrator questionnaire	S2KMINOR
School urbanicity (KURBAN)	ECLS-K created composite	KURBAN
Dependent Variable		
Spring IRT Reading Score (C2RSCALE)	Spring-kindergarten direct child assessment	C2RSCALE
Predictor Variable		
Student gender x Teacher gender interaction (GENINTERACT)[a]	ECLS-K created composite, Fall-kindergarten teacher questionnaire B	GENDER, B1TGEND

Table 2.1b: Variables and data sources for Study Part 2 (2 of 2)

Survey Design Weights		
Base year, child-level full sample weight (BYCOMW0)	ECLS-K designed sample weight	BYCOMW0
Base year, child-level stratum nesting weight (BYCOMSTR)	ECLS-K designed sample weight	BYCOMSTR
Base year, child-level PSU nesting weight (BYCOMPSU)	ECLS-K designed sample weight	BYCOMPSU

[a]Recoded or created variable. See Appendix B for full detail of data preparation procedures.

Table 2.2: Study Part 2 unadjusted and adjusted model frequency descriptives

	Unadjusted ECLS-K Sample		Analytic Full Sample		Analytic Sample (boys only)	
Characteristic	Unweighted sample	Population estimate	Unweighted sample	Population estimate	Unweighted sample	Population estimate
Gender						
Male	7,649	1,668,479 (51%)	7,499	1,643,557 (51%)	7,499	1,643,557 (100%)
Female	7,386	1,584,775 (49%)	7,237	1,560,525 (49%)	--	--
Total	*15,035*	*3,253,254 (100%)*	*14,736*	*3,204,083 (100%)*	*7,499*	*1,643,557 (100%)*
Race/Eth.						
White/Non-His.	9,225	2,040,555 (63%)	9,050	2,011,352 (63%)	4,660	1,042,717 (63%)
Afr.-Amer.	2,240	522,392 (16%)	2,197	515,770 (16%)	1,093	260,606 (16%)
Hispanic	N/A[a]	N/A[a]	1,937	448,279 (14%)	970	225,184 (14%)
Asian	N/A[a]	N/A[a]	715	76,809 (2%)	355	38,444 (2%)
NH/PI/MR	3,548	685,365 (21%)	837	151,873 (5%)	421	76,606 (5%)
Total	*15,013*	*3,248,312 (100%)*	*14,736*	*3,204,083 (100%)*	*7,499*	*1,643,557 (100%)*
Teacher gender						
Male	227	49,793 (2%)	218	48,545 (2%)	116	25,827 (2%)
Female	14,634	3,167,793 (98%)	14,518	3,155,539 (98%)	7,383	1,617,731 (98%)
Total	*14,861*	*3,217,586 (100%)*	*14,736*	*3,204,083 (100%)*	*7,499*	*1,643,557 (100%)*
Class type						
Half day	6,614	1,432,680 (44%)	6,513	1,415,955 (44%)	3,364	737,091 (45%)
All day	8,419	1,820,497 (56%)	8,221	1,788,051 (56%)	4,134	906,428 (55%)
Total	*15,033*	*3,253,177 (100%)*	*14,734*	*3,204,006 (100%)*	*7,498*	*1,643,519 (100%)*
School urbanicity						
Central city	6,857	1,503,424 (46%)	6,744	1,488,230 (46%)	3,391	754,914 (46%)
Urban fringe	4,879	1,013,385 (31%)	4,745	988,583 (31%)	2,436	513,044 (31%)
Small town	3,299	736,445 (23%)	3,247	727,271 (23%)	1,672	375,599 (23%)
Total	*15,035*	*3,253,254 (100%)*	*14,736*	*3,204,083 (100%)*	*7,499*	*1,643,557 (100%)*
% minority enroll.						
<10%	5,055	1,055,371 (33%)	5,105	1,063,830 (33%)	2,557	535,088 (33%)
10% - <25%	2,850	638,108 (20%)	2,891	645,560 (20%)	1,495	335,475 (20%)
25% - <50%	2,483	560,447 (18%)	2,496	563,000 (18%)	1,339	306,177 (18%)
50% - <75%	1,457	354,472 (11%)	1,473	356,473 (11%)	721	176,281 (11%)
>75%	2,770	575,638 (18%)	2,771	575,221 (18%)	1,387	290,535 (18%)
Total	*14,615*	*3,184,036 (100%)*	*14,736*	*3,204,083 (100%)*	*7,499*	*1,643,557 (100%)*

[a] – Estimates not available due to analysis recoded values

Table 2.3: Study Part 2 student sample mean descriptives

Characteristic	n	Mean	SE	95% CI	
				Lower	Upper
Spring IRT score[a]					
Unadj. ECLS-K sample	15,006	32.25	0.24	31.79	32.72
Analytic full sample	14,736	32.24	0.24	31.77	32.7
Analytic sample (boys)	7,499	31.27	0.24	30.8	31.74
Fall IRT score					
Unadj. ECLS-K sample	14,978	22.33	0.18	21.96	22.69
Analytic full sample	14,736	22.29	0.18	21.92	22.65
Analytic sample (boys)	7,499	21.64	0.19	21.26	22.02
Family SES level					
Unadj. ECLS-K sample	15,035	0.04	0.02	0.003	0.08
Analytic full sample	14,736	0.04	0.02	0.001	0.08
Analytic sample (boys)	7,499	0.04	0.02	-0.004	0.08
Time between assessments					
Unadj. ECLS-K sample	14,986	6.23	0.03	6.18	6.28
Analytic full sample	14,736	6.23	0.03	6.18	6.29
Analytic sample (boys)	7,499	6.26	0.03	6.2	6.31
School free lunch					
Unadj.ECLS-K sample[b]	9,835	30.11	--	--	--
Analytic full sample	14,736	30.96	1.14	28.72	33.2
Analytic sample (boys)	7,499	30.94	1.12	28.73	33.15

[a]Dependent variable. [b]Missing standard error because of stratum with single sampling unit.

Table 2.4: Test results for best model design: Two level vs. three level structure (full analytic sample)

	Two level model		Three level model	
Fixed Effect	Estimate	SE	Estimate	SE
$ß_0$	32.43	0.179	32.43	0.178
Covariance Parameter	Estimate	SE	Estimate	SE
Level 3 (Schools)	24.33	1.393	21.39	1.425
Level 2 (Teachers)	--	--	5.94	0.815
Level 1 (Students)	83.42	0.957	80.27	0.975
Variance Decomposition	% by level		% by level	
Level 3	22.58		19.89	
Level 2	--		5.52	
Level 1	77.42		74.59	
Model Information	$\bar{\chi}^2$ (df)	$P > \|\bar{\chi}^2\|$	χ^2 (df)	$P > \|\chi^2\|$
LR test vs. linear regression	2410.4 (1)	0.000**	2484.6 (2)	0.000**
Between models LRT	$\bar{\chi}^2$ (1)	$P > \|\bar{\chi}^2\|$		
H_0[a]	80.49	0.000**		

[a]H_0: The random effects associated with Level 2 (teachers nested within schools) can be omitted from the model.
**p <.01.

Table 2.5: Three-level model sample clustering

		Students per group		
Group variable	# of groups	Minimum	Average	Maximum
Schools	942	1	17.1	27
Teachers	2,852	1	5.7	27

Table 2.6: Model Results (full analytic sample)

Fixed Effect	Unconditional	Level 1 Model	Level 2 Model	Level 3 Model	Full Model
$ß_0$	32.43 (0.178)**	9.81 (0.434)**	7.95 (0.572)**	7.82 (0.533)**	8.16 (0.466)**
$ß_1$ (Fall score)		0.93 (0.006)**	0.93 (0.006)**	0.93 (0.006)**	0.93 (0.006)**
$ß_2$ (Time lapse)		0.31 (0.055)**	0.32 (0.055)**	0.32 (0.055)**	0.32 (0.055)**
$ß_3$ (Stud. gender)		-0.69 (0.089)**	-0.69 (0.09)**	-0.69 (0.09)**	-0.69 (0.09)**
$ß_4$ (SR:White)		0.14 (0.226)	-	-	-
$ß_5$ (SR:Af. Am.)		-0.87 (0.259)**	-1.15 (0.16)**	-1.23 (0.166)**	-1.16 (0.16)**
$ß_6$ (SR:Hisp.)		0.25 (0.252)	-	-	-
$ß_7$ (SR:Asian)		1.61 (0.294)**	1.49 (0.213)**	1.46 (0.215)**	1.5 (0.213)**
$ß_8$ (Family SES)		0.71 (0.072)**	0.72 (0.071)**	0.76 (0.074)**	0.72 (0.071)**
$ß_9$ (Class type)			1.13 (0.171)**	1.07 (0.172)**	1.13 (0.169)**
$ß_{10}$ (Teacher age)			0.005 (0.008)	-	-
$ß_{11}$ (Yrs teach. K)			0.01 (0.01)	-	-
$ß_{12}$ (ECE courses)			-0.04 (0.035)	-	-
$ß_{13}$ (ECE certified)			0.19 (0.152)	-	-
$ß_{14}$ (Tch. gender)			0.45 (0.505)	-	-
$ß_{15}$ (School loc.)				0.05 (0.12)	-
$ß_{16}$ (% free lunch)				0.002 (0.004)	-
$ß_{17}$ (% min. enr.)				0.1 (0.075)	-
$ß_{18}$ (Stud. gender x Tch. gender)					-0.12 (0.572)
Covariance Parameter	Estimate (SE)				
Schools (Level 3)	21.39 (1.425)	4.56 (0.376)	4.26 (0.361)	4.26 (0.36)	4.26 (0.36)
Teachers (Level 2)	5.94 (0.815)	1.98 (0.287)	1.92 (0.285)	1.92 (0.284)	1.93 (0.284)
Students (Level 1)	80.27 (0.975)	29.89 (0.363)	29.91 (0.363)	29.91 (0.363)	29.91 (0.363)
Variance Decomposition	% by level				
Level 3	19.89	12.51	11.82	11.81	11.81
Level 2	5.52	5.44	5.33	5.33	5.34
Level 1	74.59	82.05	82.85	82.86	82.85
LR Test vs. linear regression	χ^2 (df)				
	2484.6 (2)**	-	-	-	-
AIC	Estimate				
	118664.9	102360.9	102324.3	102313.7	102297.5

**p <.01.

Table 2.7: Model results (restricted analytic sample)

Fixed Effect	Unconditional	Level 1 Model	Level 2 Model	Level 3 Model	Full Model
ß$_0$	31.56 (0.194)**	8.12 (0.596)**	5.77 (0.747)**	6.42 (0.686)**	6.31 (0.613)**
ß$_1$ (Fall score)		0.94 (0.008)**	0.94 (0.008)**	0.94 (0.008)**	0.94 (0.008)**
ß$_2$ (Time lapse)		0.32 (0.074)**	0.34 (0.074)**	0.34 (0.074)**	0.34 (0.074)**
ß$_3$ (SR:White)		0.44 (0.314)	-	-	-
ß$_4$ (SR:Af. Am.)		-0.33 (0.359)	-	-	-
ß$_5$ (SR:Hisp.)		0.5 (0.353)	-	-	-
ß$_6$ (SR:Asian)		1.96 (0.416)**	1.7 (0.303)**	1.69 (0.306)**	1.7 (0.303)**
ß$_7$ (Family SES)		0.73 (0.102)**	0.77 (0.1)**	0.76 (0.106)**	0.78 (0.101)**
ß$_8$ (Behav. serv.)		0.03 (0.515)	-	-	-
ß$_9$ (Father status)		0.63 (0.176)**	0.83 (0.171)**	0.83 (0.173)**	0.83 (0.171)**
ß$_{10}$ (Class type)			1.18 (0.199)**	1.18 (0.2)**	1.17 (0.196)**
ß$_{11}$ (Teacher age)			0.02 (0.01)	-	-
ß$_{12}$ (Yrs teach. K)			0.0004 (0.013)	-	-
ß$_{13}$ (ECE courses)			-0.054 (0.045)	-	-
ß$_{14}$ (ECE certified)			0.25 (0.193)	-	-
ß$_{15}$ (Tch. gender)			0.16 (0.623)	-	-
ß$_{16}$ (School loc.)				-0.05 (0.139)	-
ß$_{17}$ (% free lunch)				-0.003 (0.005)	-
ß$_{18}$ (% min. enr.)				0.02 (0.089)	-
ß$_{19}$ (Father status x Tch. gender)					0.12 (0.69)
Covariance Parameter	Estimate (SE)				
Schools (Level 3)	21.32 (1.713)	4.77 (0.496)	4.49 (0.481)	4.48 (0.481)	4.46 (0.48)
Teachers(Level 2)	5.3 (1.318)	2.09 (0.47)	1.98 (0.467)	2.01 (0.467)	2.02 (0.467)
Students (Level 1)	82.04 (1.493)	29.87 (0.543)	29.97 (0.545)	29.97 (0.545)	29.96 (0.545)
Variance Decomposition	% by level				
Level 3	19.62	12.98	12.32	12.29	12.23
Level 2	4.88	5.69	5.43	5.52	5.54
Level 1	75.5	81.33	82.25	82.19	82.23
LR Test vs. linear regression	χ^2 (df)				
	903.6(2)**	-	-	-	-
AIC	Estimate				
	61038.09	52526	52516.69	52512.12	52493.13

* p < .05. **p <.01.

Figure 1: Decision matrix for determining significant unique associations with teacher gender.

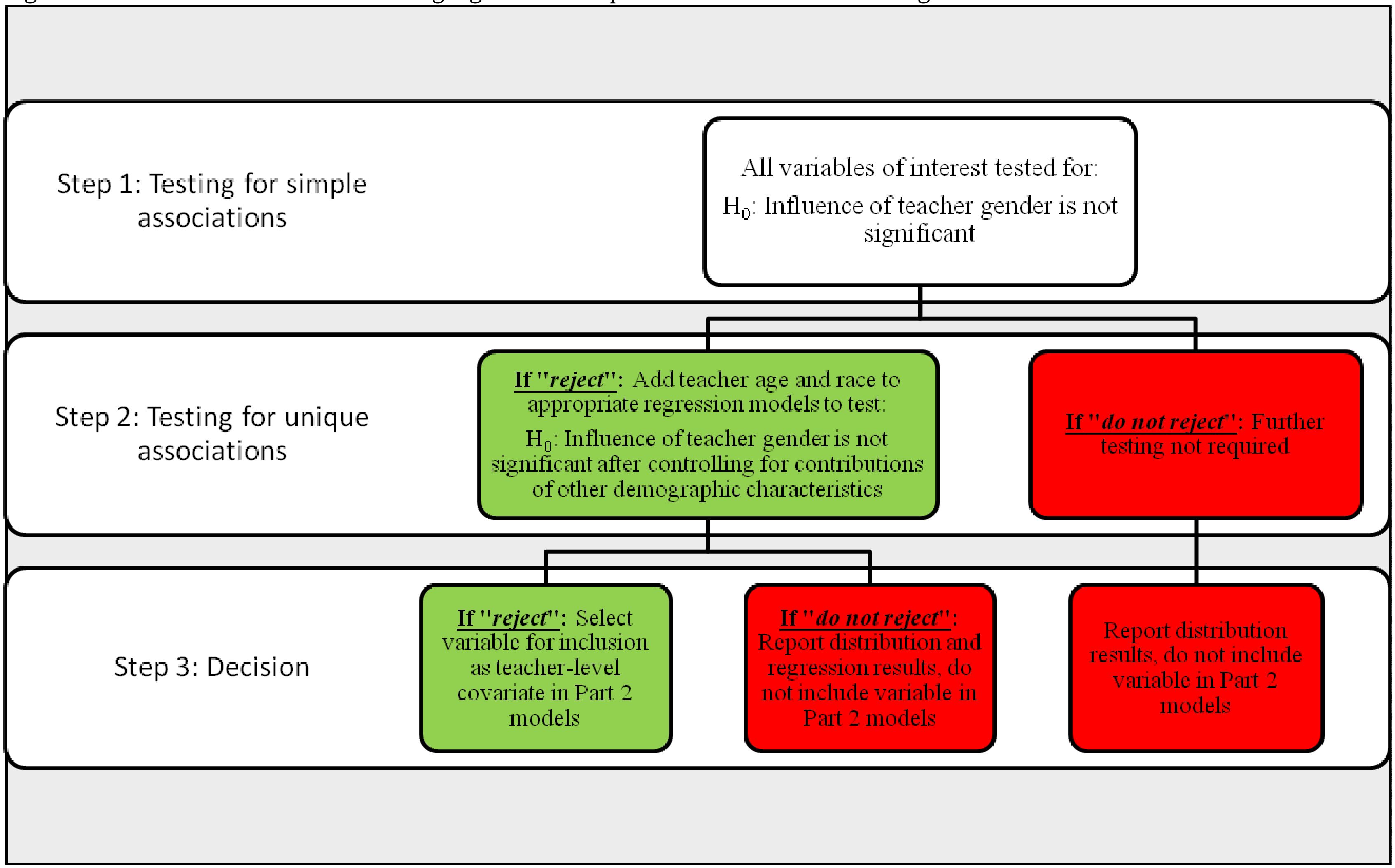


Figure 2: HLM path model (full model – full analytic sample)

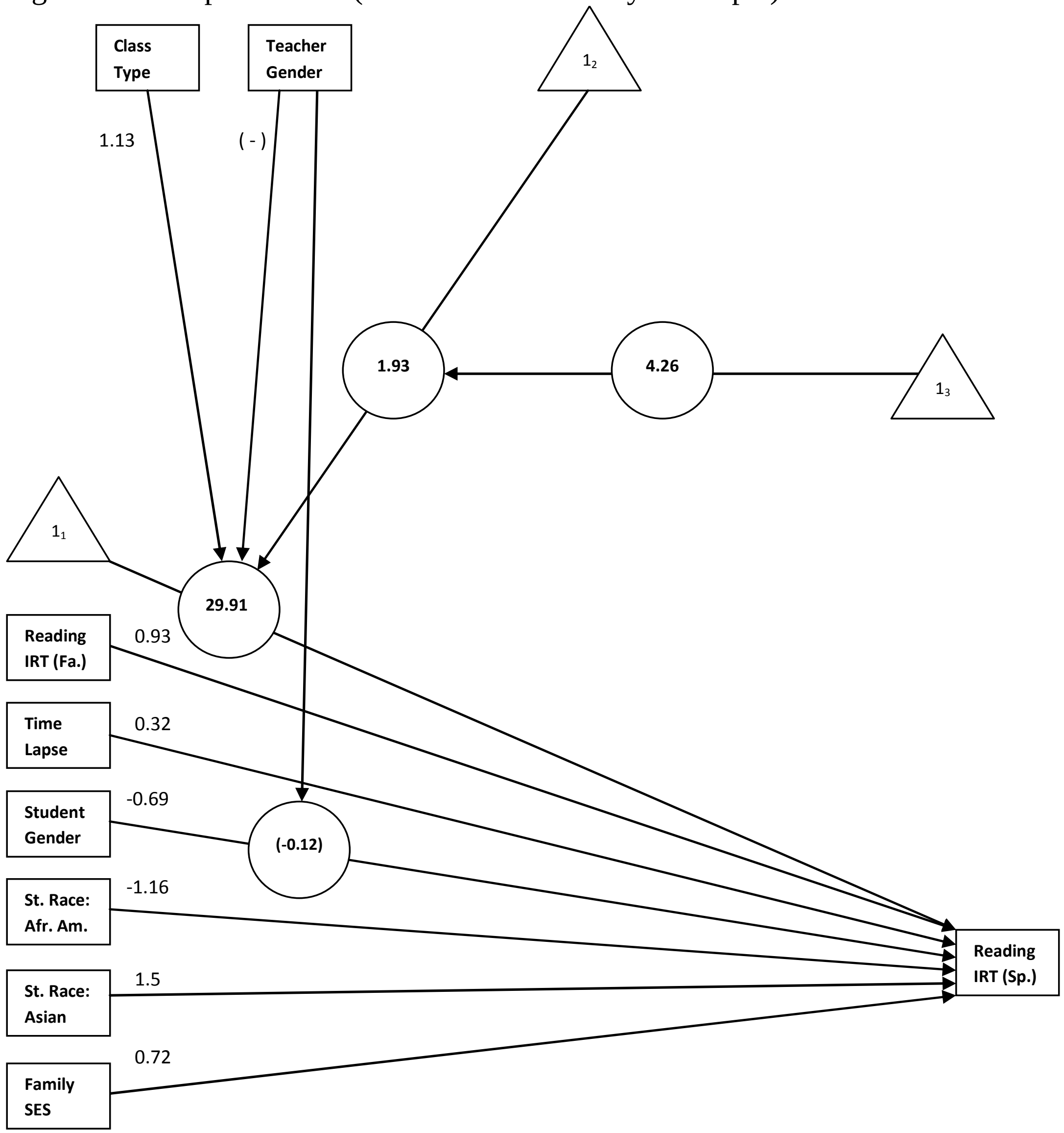


Figure 3: Actual reading IRT score distributions by student-teacher gender groupings (full analytic sample)

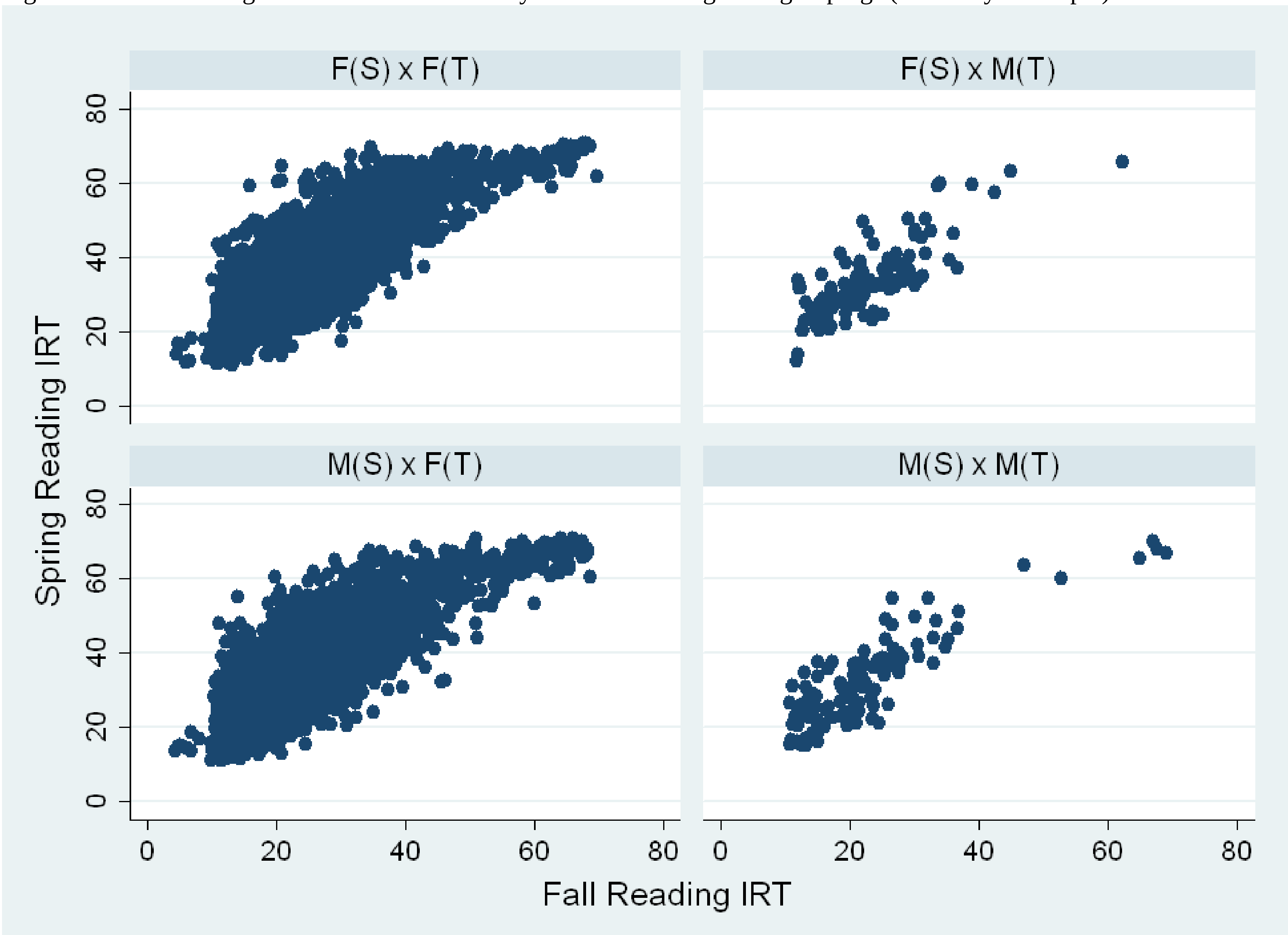


Figure 4: Full model predicted reading IRT score distributions by student-teacher gender groupings (full analytic sample)

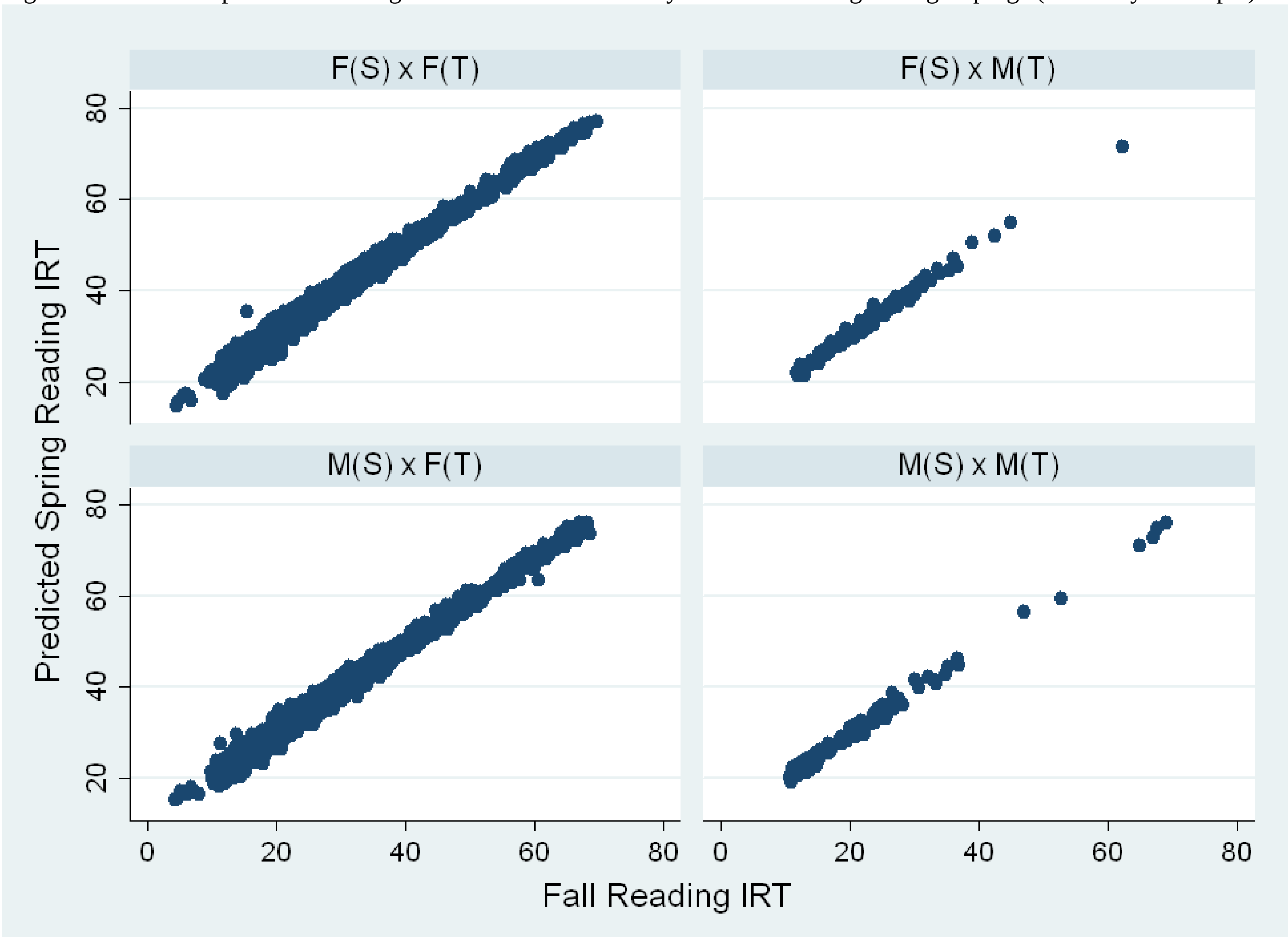


Figure 5: HLM path model (full model – restricted analytic sample)

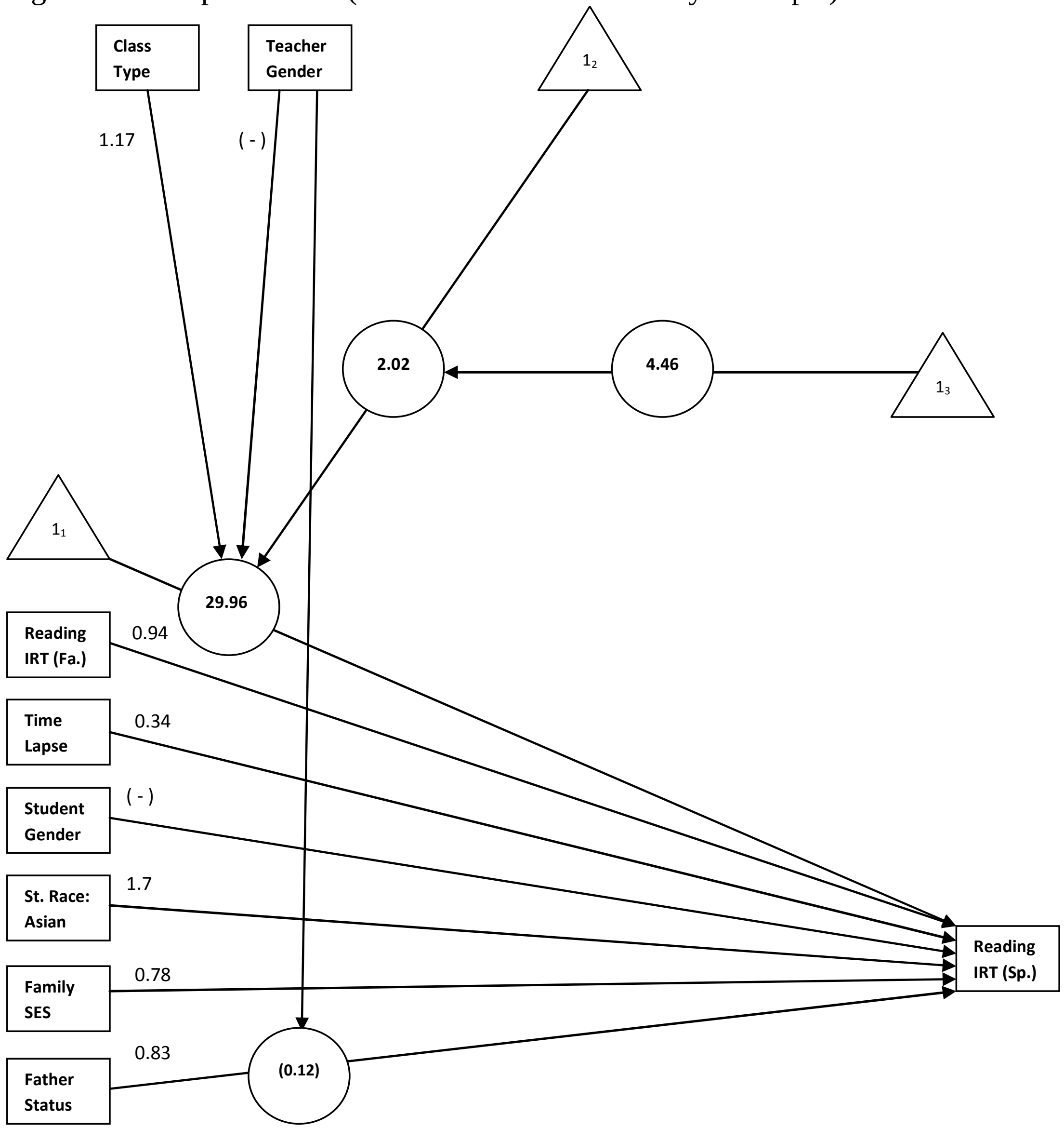


Figure 6: Actual reading IRT score distributions by student-teacher gender groupings (restricted analytic sample)

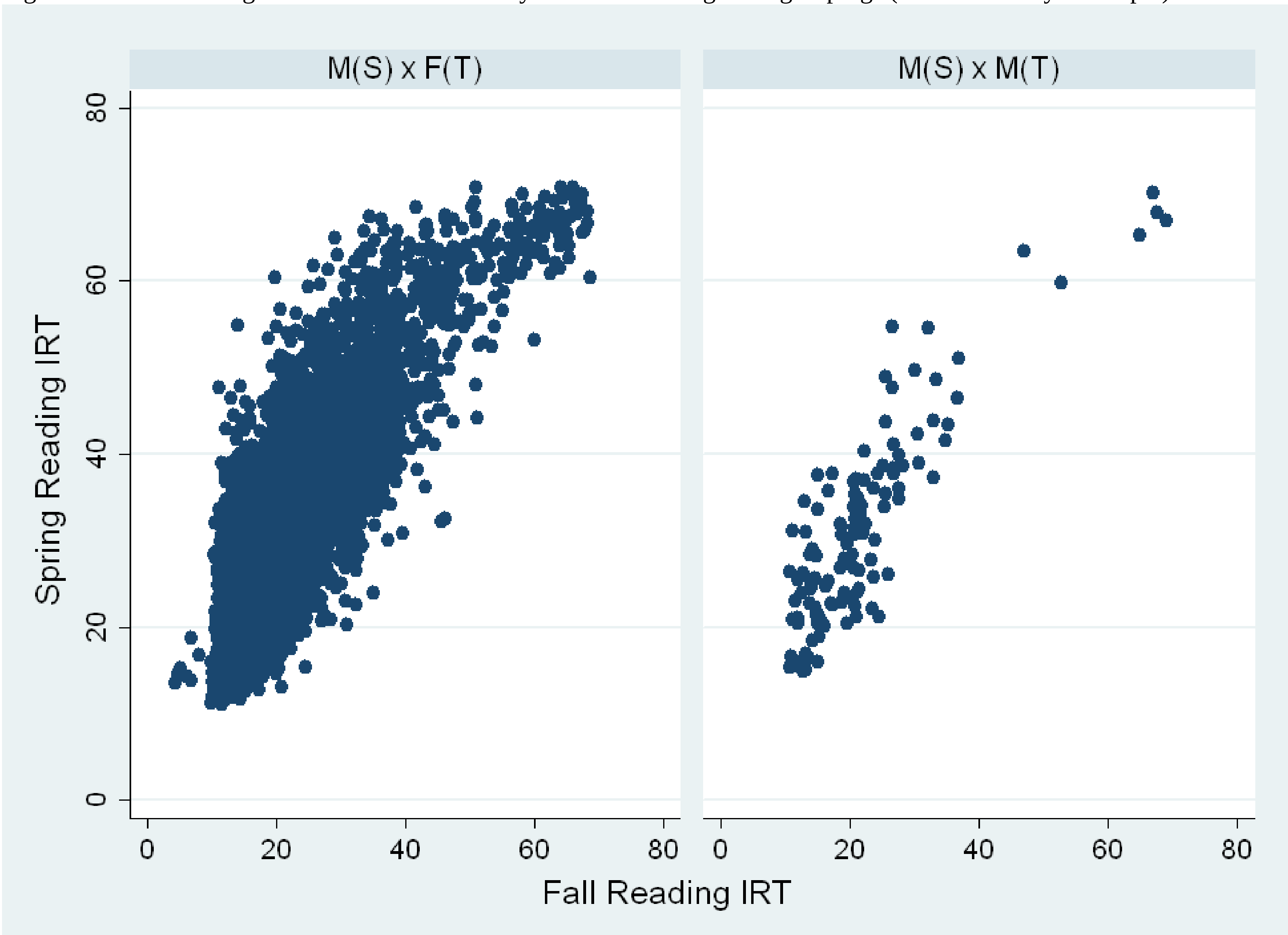
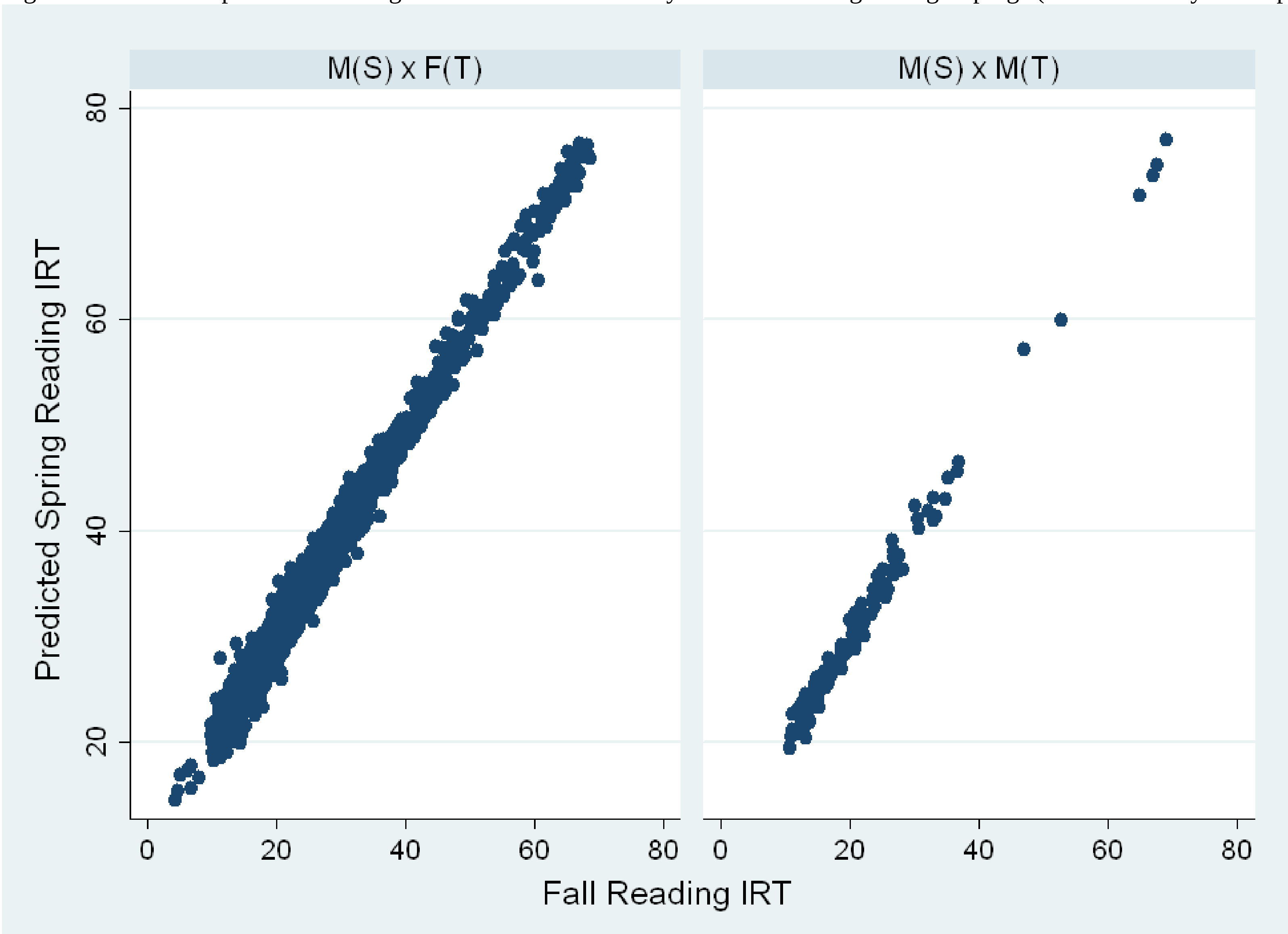


Figure 7: Full model predicted reading IRT score distributions by student-teacher gender groupings (restricted analytic sample)



Appendix A: Research Question 1 data preparation and analysis code (Stata)

```
************/ CLEANING/RECODING BEGINS HERE ************/

recode KURBAN S2KSCTYP S2KPUPRI S2KENRLS S2KMINOR KGCLASS B1AGE B1ENJOY
B1MKDIFF B1TGEND B1HISP B1RACE1 B1RACE2 B1RACE3 B1RACE4 B1RACE5 B1YRSKIN
B1HGHSTD B1EARLY B1ELEM B1MTHDRD B1TYPCER B1ELEMCT B1ERLYCT B1OTHCRT (-9 =
.)

recode A2LERNLT A2PRACLT A2NEWVOC A2DICTAT A2PHONIC A2RETELL A2READLD
A2BASAL A2SILENT A2WRKBK A2WRTWRD A2INVENT A2CHSBK A2COMPOS A2DOPROJ
A2PUBLSH A2SKITS A2JRNL A2MXDGRP A2PRTUTR A2CONVNT A2RCGNZE A2MATCH
A2WRTNME A2RHYMNG A2SYLLAB A2PREPOS A2MAINID A2PREDIC A2TEXTCU A2ORALID
A2DRCTNS A2PNCTUA A2COMPSE A2WRTSTO A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT (-9
= .)

recode A2LERNLT A2PRACLT A2NEWVOC A2DICTAT A2PHONIC A2RETELL A2READLD
A2BASAL A2SILENT A2WRKBK A2WRTWRD A2INVENT A2CHSBK A2COMPOS A2DOPROJ
A2PUBLSH A2SKITS A2JRNL A2MXDGRP A2PRTUTR (1 = 0)

recode A2LERNLT A2PRACLT A2NEWVOC A2DICTAT A2PHONIC A2RETELL A2READLD
A2BASAL A2SILENT A2WRKBK A2WRTWRD A2INVENT A2CHSBK A2COMPOS A2DOPROJ
A2PUBLSH A2SKITS A2JRNL A2MXDGRP A2PRTUTR (2 = 1)

recode A2LERNLT A2PRACLT A2NEWVOC A2DICTAT A2PHONIC A2RETELL A2READLD
A2BASAL A2SILENT A2WRKBK A2WRTWRD A2INVENT A2CHSBK A2COMPOS A2DOPROJ
A2PUBLSH A2SKITS A2JRNL A2MXDGRP A2PRTUTR (3 = 2)

recode A2LERNLT A2PRACLT A2NEWVOC A2DICTAT A2PHONIC A2RETELL A2READLD
A2BASAL A2SILENT A2WRKBK A2WRTWRD A2INVENT A2CHSBK A2COMPOS A2DOPROJ
A2PUBLSH A2SKITS A2JRNL A2MXDGRP A2PRTUTR (4 = 3)

recode A2LERNLT A2PRACLT A2NEWVOC A2DICTAT A2PHONIC A2RETELL A2READLD
A2BASAL A2SILENT A2WRKBK A2WRTWRD A2INVENT A2CHSBK A2COMPOS A2DOPROJ
A2PUBLSH A2SKITS A2JRNL A2MXDGRP A2PRTUTR (5 = 4)

recode A2LERNLT A2PRACLT A2NEWVOC A2DICTAT A2PHONIC A2RETELL A2READLD
A2BASAL A2SILENT A2WRKBK A2WRTWRD A2INVENT A2CHSBK A2COMPOS A2DOPROJ
A2PUBLSH A2SKITS A2JRNL A2MXDGRP A2PRTUTR (6 = 5)

recode A2CONVNT A2RCGNZE A2MATCH A2WRTNME A2RHYMNG A2SYLLAB A2PREPOS
A2MAINID A2PREDIC A2TEXTCU A2ORALID A2DRCTNS A2PNCTUA A2COMPSE A2WRTSTO
A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT (1 = 0)

recode A2CONVNT A2RCGNZE A2MATCH A2WRTNME A2RHYMNG A2SYLLAB A2PREPOS
A2MAINID A2PREDIC A2TEXTCU A2ORALID A2DRCTNS A2PNCTUA A2COMPSE A2WRTSTO
A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT (2 = 0)

recode A2CONVNT A2RCGNZE A2MATCH A2WRTNME A2RHYMNG A2SYLLAB A2PREPOS
A2MAINID A2PREDIC A2TEXTCU A2ORALID A2DRCTNS A2PNCTUA A2COMPSE A2WRTSTO
A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT (3 = 1)

recode A2CONVNT A2RCGNZE A2MATCH A2WRTNME A2RHYMNG A2SYLLAB A2PREPOS
A2MAINID A2PREDIC A2TEXTCU A2ORALID A2DRCTNS A2PNCTUA A2COMPSE A2WRTSTO
A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT (4 = 2)
```

```
recode A2CONVNT A2RCGNZE A2MATCH A2WRTNME A2RHYMNG A2SYLLAB A2PREPOS
A2MAINID A2PREDIC A2TEXTCU A2ORALID A2DRCTNS A2PNCTUA A2COMPSE A2WRTSTO
A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT (5 = 3)

recode A2CONVNT A2RCGNZE A2MATCH A2WRTNME A2RHYMNG A2SYLLAB A2PREPOS
A2MAINID A2PREDIC A2TEXTCU A2ORALID A2DRCTNS A2PNCTUA A2COMPSE A2WRTSTO
A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT (6 = 4)

recode A2CONVNT A2RCGNZE A2MATCH A2WRTNME A2RHYMNG A2SYLLAB A2PREPOS
A2MAINID A2PREDIC A2TEXTCU A2ORALID A2DRCTNS A2PNCTUA A2COMPSE A2WRTSTO
A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT (7 = 5)

generate RACE = 1
replace RACE = 0 if B1RACE1 == -9
replace RACE = 0 if B1RACE2 == -9
replace RACE = 0 if B1RACE3 == -9
replace RACE = 0 if B1RACE4 == -9
replace RACE = 0 if B1RACE5 == -9
replace RACE = 0 if B1RACE1 == .
replace RACE = 0 if B1RACE2 == .
replace RACE = 0 if B1RACE3 == .
replace RACE = 0 if B1RACE4 == .
replace RACE = 0 if B1RACE5 == .
replace RACE = 1 if B1RACE1 == 1
replace RACE = 1 if B1RACE2 == 1
replace RACE = 2 if B1RACE3 == 1
replace RACE = 3 if B1RACE5 == 1

generate GENDER = 2
replace GENDER = 0 if B1TGEND == 2
replace GENDER = 1 if B1TGEND == 1
replace GENDER = . if B1TGEND == -9
replace GENDER = . if B1TGEND == .

generate AGE = 8
replace AGE = 0 if B1AGE == -9
replace AGE = 0 if B1AGE == .
replace AGE = 1 if B1AGE == 24 | B1AGE == 25 | B1AGE == 26 | B1AGE == 27 |
B1AGE == 28
replace AGE = 2 if B1AGE == 29 | B1AGE == 30 | B1AGE == 31 | B1AGE == 32 |
B1AGE == 33
replace AGE = 3 if B1AGE == 34 | B1AGE == 35 | B1AGE == 36 | B1AGE == 37 |
B1AGE == 38
replace AGE = 4 if B1AGE == 39 | B1AGE == 40 | B1AGE == 41 | B1AGE == 42 |
B1AGE == 43
replace AGE = 5 if B1AGE == 44 | B1AGE == 45 | B1AGE == 46 | B1AGE == 47 |
B1AGE == 48
replace AGE = 6 if B1AGE == 49 | B1AGE == 50 | B1AGE == 51 | B1AGE == 52 |
B1AGE == 53
replace AGE = 7 if B1AGE == 54 | B1AGE == 55 | B1AGE == 56 | B1AGE == 57 |
B1AGE == 58

egen ReadWriteActMean=rowmean(A2SILENT A2INVENT A2CHSBK A2COMPOS A2JRNL
A2DICTAT A2DOPROJ A2PUBLSH)

egen PhonicsMean=rowmean(A2LERNLT A2PRACLT A2NEWVOC A2PHONIC A2CONVNT
A2RCGNZE A2MATCH A2WRTNME A2RHYMNG)
```

```
egen DidacticMean=rowmean(A2BASAL A2WRKBK A2WRTWRD A2READLD)
egen ComprehensionMean=rowmean(A2PREPOS A2MAINID A2PREDIC A2TEXTCU
A2ORALID A2DRCTNS)

egen StudCentInstMean=rowmean(A2RETELL A2SKITS)

egen ReadWriteSkillMean=rowmean(A2SYLLAB A2PNCTUA A2COMPSE A2WRTSTO
A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT)

egen MixedAbilGrpMean=rowmean(A2MXDGRP A2PRTUTR)

svyset [pweight = B1TW0], strata (B1TTWSTR) psu (B1TTWPSU)


*********/ SIMPLE COMPARISON ESTIMATES AND TESTS BEGIN HERE *********/

svy: mean B1AGE, over (GENDER)
test [B1AGE]0 = [B1AGE]1
svy: mean B1YRSKIN, over (GENDER)
test [B1YRSKIN]0 = [B1YRSKIN]1
svy linearized : tabulate GENDER AGE, count obs format(%9.0f)
svy linearized : tabulate GENDER RACE, count obs format(%9.0f)
svy linearized : tabulate GENDER KGCLASS, count obs format(%9.0f)
svy linearized : tabulate GENDER B1TYPCER, count obs format(%9.0f)
svy linearized : tabulate GENDER B1ERLYCT, count obs format(%9.0f)
svy linearized : tabulate GENDER B1ELEMCT, count obs format(%9.0f)
svy linearized : tabulate GENDER B1OTHCRT, count obs format(%9.0f)
svy linearized : tabulate GENDER B1HGHSTD, count obs format(%9.0f)
svy linearized : tabulate GENDER B1EARLY, count obs format(%9.0f)
svy linearized : tabulate GENDER B1ELEM, count obs format(%9.0f)
svy linearized : tabulate GENDER B1MTHDRD, count obs format(%9.0f)
svy linearized : tabulate GENDER KURBAN, count obs format(%9.0f)
svy linearized : tabulate GENDER S2KPUPRI, count obs format(%9.0f)
svy linearized : tabulate GENDER S2KENRLS, count obs format(%9.0f)
svy linearized : tabulate GENDER S2KMINOR, count obs format(%9.0f)
svy: mean ReadWriteActMean, over(GENDER)
test [ReadWriteActMean]0 = [ReadWriteActMean]1
svy: mean PhonicsMean, over(GENDER)
test [PhonicsMean]0 = [PhonicsMean]1
svy: mean DidacticMean, over(GENDER)
test [DidacticMean]0 = [DidacticMean]1
svy: mean ComprehensionMean, over(GENDER)
test [ComprehensionMean]0 = [ComprehensionMean]1
svy: mean StudCentInstMean, over(GENDER)
test [StudCentInstMean]0 = [StudCentInstMean]1
svy: mean ReadWriteSkillMean, over(GENDER)
test [ReadWriteSkillMean]0 = [ReadWriteSkillMean]1
svy: mean MixedAbilGrpMean, over(GENDER)
test [MixedAbilGrpMean]0 = [MixedAbilGrpMean]1


*********/ SPECIFIED REGRESSION ESTIMATES BEGIN HERE *********/

svy: regress B1AGE RACE GENDER

svy: mlogit KURBAN B1AGE RACE GENDER
```

```
svy: mlogit KGCLASS B1AGE RACE GENDER
svy: ologit B1EARLY B1AGE RACE GENDER
svy: ologit B1MTHDRD B1AGE RACE GENDER
svy: mlogit B1ERLYCT B1AGE RACE GENDER
```

Appendix B: Research Question 2 data preparation and analysis code (Stata)

```
************/ CLEANING/RECODING BEGINS HERE ************/

recode F1CLASS F2CLASS KURBAN GENDER RACE R1_KAGE R2_KAGE FKCHGTCH
FKCHGSCH A1CLASS A2CLASS C1SCTOT C1RSCALE C2RSCALE P1HDAD P1DADTYP
P1FIRKDG P2HDAD P2DADTYP WKSESQ5 WKPOVRTY B1AGE S2KMINOR S2KFLNCH S2KPUPRI
A2LERNLT A2PRACLT A2NEWVOC A2DICTAT A2PHONIC A2RETELL A2READLD A2BASAL
A2SILENT A2WRKBK A2WRTWRD A2INVENT A2CHSBK A2COMPOS A2DOPROJ A2PUBLSH
A2SKITS A2JRNL A2MXDGRP A2PRTUTR A2CONVNT A2RCGNZE A2MATCH A2WRTNME
A2RHYMNG A2SYLLAB A2PREPOS A2MAINID A2PREDIC A2TEXTCU A2ORALID A2DRCTNS
A2PNCTUA A2COMPSE A2WRTSTO A2SPELL A2VOCAB A2ALPBTZ A2RDFLNT B1TGEND
B1YRSKIN B1EARLY B1ERLYCT T2BEHPRB (-9 = .)

generate TRGENDER = 3
replace TRGENDER = 1 if B1TGEND == 1
replace TRGENDER = 0 if B1TGEND == 2
replace TRGENDER = . if B1TGEND == -9
replace TRGENDER = . if B1TGEND == .

generate STGENDER = 3
replace STGENDER = 1 if GENDER == 1
replace STGENDER = 0 if GENDER == 2
replace STGENDER = . if GENDER == -9
replace STGENDER = . if GENDER == .

generate STRACE = 9
replace STRACE = . if RACE == -9
replace STRACE = . if RACE == -1
replace STRACE = . if RACE == .
replace STRACE = 1 if RACE == 1
replace STRACE = 2 if RACE == 2
replace STRACE = 3 if RACE == 3
replace STRACE = 3 if RACE == 4
replace STRACE = 4 if RACE == 5
replace STRACE = 5 if RACE == 6
replace STRACE = 5 if RACE == 7
replace STRACE = 5 if RACE == 8

generate RESFATH = 3
replace RESFATH = 1 if P2DADTYP == 1
replace RESFATH = 1 if P2DADTYP == 2
replace RESFATH = 0 if P2DADTYP == 3
replace RESFATH = . if P2DADTYP == .
replace RESFATH = . if P2DADTYP == -9

generate BEHSER = 3
replace BEHSER = 1 if T2BEHPRB == 1
replace BEHSER = 0 if T2BEHPRB == 2
replace BEHSER = . if T2BEHPRB == .
replace BEHSER = . if T2BEHPRB == -7
replace BEHSER = . if T2BEHPRB == -8
replace BEHSER = . if T2BEHPRB == -9

generate CLASSLENGTH2 = 3
replace CLASSLENGTH2 = 1 if F2CLASS == 1
```

```
replace CLASSLENGTH2 = 1 if F2CLASS == 2
replace CLASSLENGTH2 = 2 if F2CLASS == 3
replace CLASSLENGTH2 = . if F2CLASS == -9
replace CLASSLENGTH2 = . if F2CLASS == .

destring S1_ID, generate (NS1_ID)
destring S2_ID, generate (NS2_ID)
destring CHILDID, ignore("C") generate (NCHILDID)
destring T1_ID, ignore("TO") generate (NT1_ID)
destring T2_ID, ignore("TO") generate (NT2_ID)

generate TIMELAPSE = R2_KAGE - R1_KAGE

drop if NT2_ID == .
drop if NS2_ID == .
drop if FKCHGTCH == .
drop if FKCHGTCH == -9
drop if FKCHGTCH == 1
drop if FKCHGSCH == .
drop if FKCHGSCH == -9
drop if FKCHGSCH == 1
drop if C1RSCALE == -1
drop if C2RSCALE == -1

svyset [pweight = BYCOMW0], strata (BYCOMSTR) psu (BYCOMPSU)

************/ PRE-IMPUTATION SAMPLE ESTIMATES BEGIN HERE ************/

svy linearized : tabulate STGENDER, cell count obs format(%11.3f)
svy linearized : tabulate TRGENDER, cell count obs format(%11.3f)
svy linearized : tabulate STRACE, cell count obs format(%11.3f)
svy linearized : tabulate KURBAN, cell count obs format(%11.3f)
svy linearized : tabulate S2KMINOR, cell count obs format(%11.3f)
svy linearized : tabulate CLASSLENGTH2, cell count obs format(%11.3f)
svy linearized : mean S2KFLNCH
svy linearized : mean C1RSCALE
svy linearized : mean C2RSCALE
svy linearized : mean WKSESL
svy linearized : mean B1AGE
svy linearized : mean TIMELAPSE

************/ IMPUTATION PROCESS BEGINS HERE ************/

misschk C1RSCALE C2RSCALE WKSESL STGENDER STRACE TIMELAPSE RESFATH BEHSER
B1AGE B1YRSKIN B1EARLY B1ERLYCT TRGENDER KURBAN S2KMINOR S2KFLNCH

ice C1RSCALE C2RSCALE WKSESL STGENDER STRACE TIMELAPSE RESFATH BEHSER
B1AGE B1YRSKIN B1EARLY B1ERLYCT TRGENDER KURBAN S2KMINOR S2KFLNCH,
saving(imputed18) m(3)

use imputed18, clear

tab _mj
```

```
sum NT2_ID NS2_ID C1RSCALE C2RSCALE WKSESL STGENDER STRACE TIMELAPSE
RESFATH BEHSER B1AGE B1YRSKIN B1EARLY B1ERLYCT TRGENDER KURBAN S2KMINOR
S2KFLNCH

drop if _mj==0

sum NT2_ID NS2_ID C1RSCALE C2RSCALE WKSESL STGENDER STRACE TIMELAPSE
RESFATH BEHSER B1AGE B1YRSKIN B1EARLY B1ERLYCT TRGENDER KURBAN S2KMINOR
S2KFLNCH

collapse (mean) NS2_ID NT2_ID F1CLASS F2CLASS R1_KAGE R2_KAGE FKCHGTCH
FKCHGSCH A1CLASS A2CLASS BYCOMW0 BYCOMSTR BYCOMPSU C1RSCALE C2RSCALE
READGAIN WKSESL STGENDER STRACE TIMELAPSE RESFATH BEHSER B1AGE B1YRSKIN
B1EARLY B1ERLYCT TRGENDER CLASSLENGTH2 KURBAN S2KMINOR S2KFLNCH S2KPUPRI
WKPOVRTY P1HDAD P1DADTYP P1FIRKDG P2HDAD P2DADTYP WKSESQ5 C1SCTOT,
by(NCHILDID)


************/ POST-IMPUTATION SAMPLE ESTIMATES BEGIN HERE ************/

sum NT2_ID NS2_ID C1RSCALE C2RSCALE WKSESL STGENDER STRACE TIMELAPSE
RESFATH BEHSER B1AGE B1YRSKIN B1EARLY B1ERLYCT TRGENDER CLASSLENGTH2
KURBAN S2KMINOR S2KFLNCH
svy linearized : tabulate STGENDER, cell count obs format(%11.3f)
svy linearized : tabulate TRGENDER, cell count obs format(%11.3f)
svy linearized : tabulate STRACE, cell count obs format(%11.3f)
svy linearized : tabulate KURBAN, cell count obs format(%11.3f)
svy linearized : tabulate S2KMINOR, cell count obs format(%11.3f)
svy linearized : tabulate CLASSLENGTH2, cell count obs format(%11.3f)
svy linearized : mean S2KFLNCH
svy linearized : mean C1RSCALE
svy linearized : mean C2RSCALE
svy linearized : mean WKSESL
svy linearized : mean B1AGE
svy linearized : mean TIMELAPSE

****/RECODING 'STRACE' FOR USE & INTERPRETATION****/

generate STRACEWH = 9
replace STRACEWH = . if STRACE == -9
replace STRACEWH = . if STRACE == -1
replace STRACEWH = . if STRACE == .
replace STRACEWH = 1 if STRACE == 1
replace STRACEWH = 0 if STRACE == 2
replace STRACEWH = 0 if STRACE == 3
replace STRACEWH = 0 if STRACE == 4
replace STRACEWH = 0 if STRACE == 5

generate STRACEAA = 9
replace STRACEAA = . if STRACE == -9
replace STRACEAA = . if STRACE == -1
replace STRACEAA = . if STRACE == .
replace STRACEAA = 0 if STRACE == 1
replace STRACEAA = 1 if STRACE == 2
replace STRACEAA = 0 if STRACE == 3
replace STRACEAA = 0 if STRACE == 4
replace STRACEAA = 0 if STRACE == 5
```

```
generate STRACEHS = 9
replace STRACEHS = . if STRACE == -9
replace STRACEHS = . if STRACE == -1
replace STRACEHS = . if STRACE == .
replace STRACEHS = 0 if STRACE == 1
replace STRACEHS = 0 if STRACE == 2
replace STRACEHS = 1 if STRACE == 3
replace STRACEHS = 0 if STRACE == 4
replace STRACEHS = 0 if STRACE == 5

generate STRACEAS = 9
replace STRACEAS = . if STRACE == -9
replace STRACEAS = . if STRACE == -1
replace STRACEAS = . if STRACE == .
replace STRACEAS = 0 if STRACE == 1
replace STRACEAS = 0 if STRACE == 2
replace STRACEAS = 0 if STRACE == 3
replace STRACEAS = 1 if STRACE == 4
replace STRACEAS = 0 if STRACE == 5

generate STRACEOT = 9
replace STRACEOT = . if STRACE == -9
replace STRACEOT = . if STRACE == -1
replace STRACEOT = . if STRACE == .
replace STRACEOT = 0 if STRACE == 1
replace STRACEOT = 0 if STRACE == 2
replace STRACEOT = 0 if STRACE == 3
replace STRACEOT = 0 if STRACE == 4
replace STRACEOT = 1 if STRACE == 5

svy linearized : tabulate STRACEWH, cell count obs format(%11.3f)
svy linearized : tabulate STRACEAA, cell count obs format(%11.3f)
svy linearized : tabulate STRACEHS, cell count obs format(%11.3f)
svy linearized : tabulate STRACEAS, cell count obs format(%11.3f)
svy linearized : tabulate STRACEOT, cell count obs format(%11.3f)

************/ MODEL STRUCTURE TESTS BEGIN HERE ************/

xi: xtmixed C2RSCALE || NS2_ID: || NT2_ID:, variance
estimates store model2_18Un3
estat group
estat ic
xtmrho

xi: xtmixed C2RSCALE || NS2_ID:, variance
estimates store model2_18Un2
estat group
estat ic
xtmrho

lrtest model2_18Un2 model2_18Un3

************/ THREE-LEVEL MODEL BUILDING BEGINS HERE ************/
```

```
*****/ BUILDING UNCONDITIONAL MODEL *****/
xi: xtmixed C2RSCALE || NS2_ID: || NT2_ID:, variance
estimates store model2_18UnB
estat group
estat ic
xtmrho

*****/ BUILDING LEVEL 1 MODEL *****/
xi: xtmixed C2RSCALE C1RSCALE TIMELAPSE STGENDER STRACEWH STRACEAA
STRACEHS STRACEAS WKSESL || NS2_ID: || NT2_ID:, variance
estimates store model2_18_StlevelB
estat group
estat ic
xtmrho


*****/ BUILDING LEVEL 2 MODEL *****/
xi: xtmixed C2RSCALE C1RSCALE TIMELAPSE STGENDER STRACEAA STRACEAS WKSESL
CLASSLENGTH2 B1AGE B1YRSKIN B1EARLY B1ERLYCT TRGENDER || NS2_ID: ||
NT2_ID:, variance
estimates store model2_18_TchlevelB
estat group
estat ic
xtmrho

*****/ BUILDING LEVEL 3 MODEL *****/
xi: xtmixed C2RSCALE C1RSCALE TIMELAPSE STGENDER STRACEAA STRACEAS WKSESL
CLASSLENGTH2 KURBAN S2KFLNCH S2KMINOR || NS2_ID: || NT2_ID:, variance
estimates store model2_18_SclevelB
estat group
estat ic
xtmrho

*****/ BUILDING FULL MODEL *****/
generate GENINTERACT = STGENDER*TRGENDER

xi: xtmixed C2RSCALE C1RSCALE TIMELAPSE STGENDER STRACEAA STRACEAS WKSESL
CLASSLENGTH2 GENINTERACT || NS2_ID: || NT2_ID:, variance
estimates store model2_18_FullB
estat group
estat ic
xtmrho

******/CREATING COMPARISONS, SCATTERPLOTS ******/

egen TMATCH = group(STGENDER TRGENDER), label

svy linearized : mean C1RSCALE, over (TMATCH)
svy linearized : mean C2RSCALE, over (TMATCH)

gen PredictedSpIRT = 8.16 + .93*C1RSCALE + .32*TIMELAPSE + -.69*STGENDER +
-1.16*STRACEAA + 1.5*STRACEAS + .72*WKSESL + 1.13*CLASSLENGTH2 + -
.12*GENINTERACT
label variable PredictedSpIRT "Predicted Spring Reading IRT"

twoway (scatter C2RSCALE C1RSCALE), by(TMATCH, note(""))
twoway (scatter PredictedSpIRT C1RSCALE), by(TMATCH, note(""))
```

```
***********/ BOYS-ONLY SAMPLE EXTENSION MODELS BEGIN HERE ***********/

drop if STGENDER == 0

sum NT2_ID NS2_ID C1RSCALE C2RSCALE WKSESL STGENDER STRACE TIMELAPSE
RESFATH BEHSER B1AGE B1YRSKIN B1EARLY B1ERLYCT TRGENDER CLASSLENGTH2
KURBAN S2KMINOR S2KFLNCH
svy linearized : tabulate STGENDER, cell count obs format(%11.3f)
svy linearized : tabulate TRGENDER, cell count obs format(%11.3f)
svy linearized : tabulate STRACEWH, cell count obs format(%11.3f)
svy linearized : tabulate STRACEAA, cell count obs format(%11.3f)
svy linearized : tabulate STRACEHS, cell count obs format(%11.3f)
svy linearized : tabulate STRACEAS, cell count obs format(%11.3f)
svy linearized : tabulate STRACEOT, cell count obs format(%11.3f)
svy linearized : tabulate KURBAN, cell count obs format(%11.3f)
svy linearized : tabulate S2KMINOR, cell count obs format(%11.3f)
svy linearized : tabulate CLASSLENGTH2, cell count obs format(%11.3f)
svy linearized : mean S2KFLNCH
svy linearized : mean C1RSCALE
svy linearized : mean C2RSCALE
svy linearized : mean WKSESL
svy linearized : mean B1AGE
svy linearized : mean TIMELAPSE

*****/ BUILDING UNCONDITIONAL MODEL *****/
xi: xtmixed C2RSCALE || NS2_ID: || NT2_ID:, variance
estimates store model2_18bUn
estat group
estat ic
xtmrho

*****/ BUILDING LEVEL 1 MODEL *****/
xi: xtmixed C2RSCALE C1RSCALE TIMELAPSE STRACEWH STRACEAA STRACEHS
STRACEAS WKSESL RESFATH BEHSER || NS2_ID: || NT2_ID:, variance
estimates store model2_18b_Stlevel
estat group
estat ic
xtmrho

*****/ BUILDING LEVEL 2 MODEL *****/
xi: xtmixed C2RSCALE C1RSCALE TIMELAPSE STRACEAS WKSESL RESFATH
CLASSLENGTH2 B1AGE B1YRSKIN B1EARLY B1ERLYCT TRGENDER || NS2_ID: ||
NT2_ID:, variance
estimates store model2_18b_Tchlevel
estat group
estat ic
xtmrho

*****/ BUILDING LEVEL 3 MODEL *****/
xi: xtmixed C2RSCALE C1RSCALE TIMELAPSE STRACEAS WKSESL RESFATH
CLASSLENGTH2 KURBAN S2KFLNCH S2KMINOR || NS2_ID: || NT2_ID:, variance
estimates store model2_18b_Sclevel
estat group
estat ic
xtmrho
```

```
*****/ BUILDING FULL MODEL *****/

generate TCHFATH = RESFATH*TRGENDER

xi: xtmixed C2RSCALE C1RSCALE TIMELAPSE STRACEAS WKSESL RESFATH
CLASSLENGTH2 TCHFATH || NS2_ID: || NT2_ID:, variance
estimates store model2_18b_Full
estat group
estat ic
xtmrho

******/CREATING COMPARISONS, SCATTERPLOTS ******/

svy linearized : mean C1RSCALE, over (TMATCH)
svy linearized : mean C2RSCALE, over (TMATCH)

gen PredictedSpIRTBoys = 6.31 + .94*C1RSCALE + .34*TIMELAPSE +
1.7*STRACEAS + .78*WKSESL + .83*RESFATH + 1.17*CLASSLENGTH2 + .12*TCHFATH
label variable PredictedSpIRTBoys "Predicted Spring Reading IRT"

twoway (scatter C2RSCALE C1RSCALE), by(TMATCH, note(""))
twoway (scatter PredictedSpIRTBoys C1RSCALE), by(TMATCH, note(""))
```

Appendix C: Factor Loadings for Reading and Language Arts Instructional Activities and Skills Items, Full Sample (n=2323): 1999

Item number	Variable name	Reading and writing activities	Phonics	Didactic instruction	Compre-hension	Student-centered instruction	Reading and writing skills	Mixed-ability grouping
29f	A2SYLLAB	0.09	0.02	0.11	-0.20	-0.21	**-0.61**	0.01
29m	A2PNCTUA	0.22	0.14	0.04	-0.22	0.18	**-0.68**	0.02
29n	A2COMPSE	0.40	0.05	-0.03	-0.18	0.05	**-0.72**	-0.01
29o	A2WRTSTO	0.25	-0.03	0.00	-0.14	-0.23	**-0.66**	0.03
29p	A2SPELL	0.13	0.07	0.17	-0.03	-0.04	**-0.71**	0.08
29q	A2VOCAB	-0.03	0.17	0.09	-0.20	-0.03	**-0.52**	0.33
29r	A2ALPBTZ	-0.01	0.06	0.20	-0.08	-0.20	**-0.53**	0.17
29s	A2RDFLNT	0.13	0.01	0.45	-0.14	-0.03	**-0.56**	0.03
28a	A2LERNLT	0.05	**0.76**	0.05	-0.10	-0.05	0.04	0.00
28b	A2PRACLT	0.09	**0.69**	0.22	-0.04	-0.13	-0.09	0.03
28c	A2NEWVOC	0.11	**0.39**	0.08	-0.32	-0.13	-0.11	0.26
28e	A2PHONIC	0.11	**0.73**	0.15	-0.09	0.02	-0.07	0.11
29a	A2CONVNT	0.07	**0.40**	-0.12	-0.36	0.01	-0.14	0.12
29b	A2RCGNZE	0.02	**0.76**	-0.11	-0.08	0.02	-0.02	0.04
29c	A2MATCH	0.09	**0.80**	0.00	-0.12	0.01	-0.09	0.05
29d	A2WRTNME	0.08	**0.54**	-0.04	-0.14	-0.19	-0.08	0.01
29e	A2RHYMNG	0.11	**0.40**	0.08	-0.32	-0.32	-0.18	0.04
28j	A2BASAL	0.02	-0.00	**0.76**	0.01	-0.06	-0.14	-0.02
28l	A2WRKBK	-0.29	0.23	**0.56**	-0.07	0.05	-0.04	0.08
28m	A2WRTWRD	0.06	0.14	**0.45**	0.03	-0.07	-0.35	0.17
29g	A2PREPOS	-0.11	0.16	0.04	**-0.44**	-0.37	-0.34	0.05
29h	A2MAINID	0.10	0.04	0.08	**-0.68**	-0.30	-0.22	0.01
29i	A2PREDIC	0.26	0.13	-0.04	**-0.72**	-0.14	-0.12	0.10
29j	A2TEXTCU	0.18	0.05	0.10	**-0.69**	-0.07	-0.27	0.10
29k	A2ORALID	0.17	0.24	-0.01	**-0.65**	0.17	-0.10	0.20
29l	A2DRCTNS	0.10	0.30	0.01	**-0.59**	0.18	-0.10	0.11
28i	A2READLD	*0.38	0.22	**0.41**	-0.27	0.02	-0.24	0.03
28k	A2SILENT	**0.43**	0.03	0.34	-0.22	0.25	-0.13	0.18
28n	A2INVENT	**0.74**	0.17	-0.03	-0.18	-0.01	-0.21	0.09
28o	A2CHSBK	**0.59**	0.19	0.15	-0.32	0.19	-0.09	0.16
28p	A2COMPOS	**0.68**	0.05	0.04	-0.07	-0.18	-0.30	0.14
28t	A2JRNL	**0.71**	0.05	-0.09	-0.14	-0.10	-0.21	0.00
28d	A2DICTAT	**0.46**	0.22	-0.01	-0.15	*-0.45	-0.11	0.03
28h	A2RETELL	0.30	0.23	0.18	-0.39	**-0.47**	-0.03	-0.01
28q	A2DOPROJ	**0.42**	-0.01	0.05	-0.32	*-0.21	-0.04	0.25
28r	A2PUBLSH	**0.53**	-0.01	-0.01	-0.02	*-0.38	-0.19	0.22
28s	A2SKITS	0.28	0.04	0.00	-0.15	**-0.50**	-0.09	0.22
28v	A2MXDGRP	0.18	0.09	-0.05	-0.19	0.02	-0.03	**0.73**
28w	A2PRTUTR	0.16	0.10	0.14	-0.11	-0.13	-0.16	**0.67**

NOTE: Each item's highest loading appears in boldface type.
* indicates item assigned to a factor other than the one on which it had its highest loading.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99, Base Year Restricted-Use data files.

REFERENCES

Allan, J. (1994). *Anomaly as exemplar: The meanings of role-modeling for men elementary teachers* Iowa: EDRS. (ERIC Document Reproduction Service No. ED378190).

Altermatt, E., Jovanovic & Perry (1998). Bias or responsivity? Sex and achievement-level effects on teachers' classroom questioning practices. *Journal of educational psychology, 90*(3), 516.

American Association of University Women. (2001, February). *Beyond the "gender wars": A conversation about girls, boys and education*. Washington, DC: AAUW Educational Foundation.

Arnold, D. & Doctoroff, G. (2003). The early education of socioeconomically disadvantaged children. *Annual Review of Psychology, 54*(1), 517-545.

Barbarin, O. (2002). Ready or not! African American males in kindergarten. In B. Bowman (Ed.), *Love to read: Preparing African American children for reading success*. (pp. 1-15) Washington, DC: National Black Child Development Institute.

Barkley, R. A., Ullman, D. G., Otto, L., & Brecht, J. M. (1977). The effects of sex typing and sex appropriateness of modeled behavior on children's imitation. *Child Development, 48*, 721-725.

Barnett. R.C., & Rivers, C. (2007). Gender myths and the education of boys. *Independent school, 66*(2), 92-103.

Bennett, R. E. (1993). Influence of behavior perceptions and gender on teachers. *Education, 85*(2), 347.

Beswick, J.F., Willms, J.D. & Sloat, E.A. (2005). A comparative study of teacher ratings of emergent literacy skills and student performance on a standardized measure. *Journal of School Psychology, 126*(1), 116-137.

Biddulph, S. (2004). *Manhood*. Vermilion.

Birch, S.H. & Ladd, G.W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology, 35*(1), 61-79.

Blount, J. (2000). Spinsters, bachelors, and other gender transgressors in school employment, 1850-1990. *Review of Educational Research, 70*(1), 83-101.

Brookhart, S. M., & Loadman, W. E. (1996). Characteristics of male elementary teachers in the U.S.A., at teacher education program entry and exit. *Teaching and Teacher Education, 12*(2), 197-210.

Brophy, J. (1985). Interactions of male and female students with male and female teachers. In Wilkinson, L.C., Marrett, C.B (Eds.), *Gender influences in classroom interaction*. (pp. 115-142) Orlando, FL: Academic Press, Inc.

Brophy, J. E., & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Education Psychology, 61*(5), 365-374.

Brutsaert, H. & Bracke, P. (1994). Gender context of the elementary school: Sex differences in affective outcomes. *Educational Studies, 20*(1), 3-11.

Buchanan, T., Burts, D. C., Bidner, J., White, V. F., & Charlesworth, R. (1998). Predictors of the developmental appropriateness of the beliefs and practices of first, second, and third grade teachers. *Early Childhood Research Quarterly, 13*(3), 459-83.

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review, 106*, 676-713.

Carrington, B., Francis, B., Hutchings, M., Skelton, C., Read, B., & Hall, I. (2007). Does the gender of the teacher really matter? seven- to eight-year-olds' accounts of their interactions with their teachers. *Educational Studies, 33*(4), 397-413.

Carter, S.B. (1989). Incentives and rewards to teaching. In D. Warren (Ed.), *American teachers : Histories of a profession at work*. (pp. 49-62) New York: Macmillan.

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the early childhood longitudinal study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology, 98*(3), 489-507.

Christian, K., Morrison, F. J., & Bryant, F. B. (1998). Predicting kindergarten academic skills: Interactions among child care, maternal education, and family literacy environments. *Early Childhood Research Quarterly, 13*(3), 501-521.

Cognard-Black, A.J. (2004). Will they stay, or will they go?: Sex-atypical work among token men who teach. *The Sociological Quarterly, 45*(1), 113-139.

Condron, D. (2007). Stratification and educational sorting: Explaining ascriptive inequalities in early childhood reading group placement. *Social Problems (Berkeley, Calif.), 54*(1), 139-160.

Costello, B. (2008). Leveraging gender differences to boost test scores. *The Education Digest, 73*(9), 32-35.

Deaux, K., & Major, B. (1987). Putting gender into context: An interactive model of gender-related behavior. *Psychological Review, 94*, 369.

DeCorse, C. J. B., & Vogtle, S. P. (1997). In a complex voice: The contradictions of male elementary teachers' career choice and professional identity. *Journal of Teacher Education, 48*(1), 37-46.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity or gender matter? *The American Economic Review, 95*(2), 158.

Dee, T. S. (2006a). Teachers and the gender gaps in student achievement. *The journal of human resources, 42*(3), 528-554.

Dee, T. S. (2006b). The why chromosome: How a teacher's gender affects boys and girls. *Education Next*, *6*(4), 68-75.

Dubow, E. F., & Ippolito, M.F. (1994). Effects of poverty and quality of the home environment on changes in the academic and behavioral adjustment of elementary school-age children. *Journal of Clinical Child Psychology, 23*(4), 401-412.

Duffy, J. Warren, K. & Walsh, M. (2002). Classroom Interactions: Gender of Teacher, Gender of Student, and Classroom Subject *Sex Roles*, 45, (9-10)

Dusek, J. B., & Joseph, G. (1983). The bases of teacher expectancies: A meta-analysis. *Journal of Educational Psychology, 75*(3), 327-46.

Dutro, E. (2001). "But that's a girls' book!" exploring gender boundaries in children's reading practices. *The Reading Teacher, 55*(4), 376-384.

Eccles, J.S. & Blumenfeld, P. (1985). Classroom experiences and student gender: Are there differences and do they matter?. In Wilkinson, L.C., Marrett, C.B (Eds.), *Gender influences in classroom interaction*. (pp. 79-113) Orlando, FL: Academic Press, Inc.

Enoch, J. (2008). A woman's place is in the school: Rhetorics of gendered space in nineteenth century America. *College English, 70*(3), 275-292.

Entwisle, D.R. & Alexander, K.L. (1992). Summer Setback: Race, Poverty, School Composition, and Mathematics Achievement in the First Two Years of School. *American Sociological Review, 57*(1), 72-84.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *Sociology of Education, 80*(2), 114-138.

Fabes, R., Martin, C., Hanish, L., Madden-Derdich, D., & Anders, M. (2003). Early school competence: The roles of sex-segregated play and effortful control. *Developmental Psychology*, *39*(5), 848-858. Retrieved August 10, 2009, doi:10.1037/0012-1649.39.5.848

Fagot, B. I. (1981). Male and female teachers: Do they treat boys and girls differently? *Sex Roles: A Journal of Research*, *7*(3), 263-71.

Fennema, E. & Peterson, P.L. (1985). Autonomous learning behavior: A possible explanation of gender-related differences in mathematics. In Wilkinson, L.C., Marrett, C.B (Eds.), *Gender influences in classroom interaction*. (pp. 17-36) Orlando, FL: Academic Press, Inc.

Fennema, E., Peterson, P.L., Carter, T.P., & Lubinski, C.A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics, 21*(1), 55-69.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Mehta, P., & Schatschneider, C. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*(1), 37-55.

Foorman, B. R., Schatschneider, C., & Eakin, M. N. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology, 31*(1), 1-29.

Freidus, H. (1992, April). *Men in a women's world: A study of male second career teachers in elementary schools*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Gamble, R.J. & Wilkins, J. (1997). Beyond tradition: Where are the men in elementary education?: *Contemporary Education, 68,* 187-193.

Garraghy, D.A. (2006). Three third-grade teachers’ gender-related beliefs and behavior. *The Elementary School Journal*, *102*(1), 81-94.

Gilliam, W.S. (2005). *Prekindergarteners left behind*: *Expulsion rates in state prekindergarten systems*. New Haven, CT: Yale University Child Study Center.

Gold, D. & Reis, M. (1982). Male teacher effects on young children: A theoretical and empirical consideration. *Sex Roles: A Journal of Research, 8*(5), 493-513.

Good, T.L., Sikes, J.N. & Brophy, J.E. (1973). Effects of teacher sex and student sex on classroom interaction. *Journal of Education Psychology, 65*(1), 74-87.

Grant, L. (1985). Race-gender status, classroom interactions, and children's socialization in elementary school. In Wilkinson, L.C., Marrett, C.B (Eds.), *Gender influences in classroom interaction*. (pp. 57-75) Orlando, FL: Academic Press, Inc.

Gross, N. & Trask, A.E. (1976). The sex factor and the management of schools. New York: John Wiley& Sons.

Guarino, C.M., Hamilton, L. S., Lockwood, J. R., Rathbun, A. H., RAND Corp., A.,VA., Education Statistics Services Inst., Washington,DC., et al. (2006). *Teacher qualifications, instructional practices, and reading and mathematics gains of kindergartners. research and development report. NCES 2006-031* National Center for Education Statistics.

Halpern, D.F. (1997). Sex differences in intelligence: Implications for education. *American Psychologist*, *52*(1), 1091-1102.

Hamilton, L. S., & Guarino, C. M. (2006). *Measuring the practices, philosophies, and characteristics of kindergarten teachers. WR-199-EDU* National Center for Education Statistics.

Hamre, B. & Pianta, R. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development, 72*(2), 625-638.

Harris M. J., Rosenthal, R. & Snodgrass, S.E., (1986). The effects of teacher expectations, gender, and behavior on pupil academic performance and self-concept. *The Journal of Educational Research, 79*(3), 173.

Hedges, L.V. & Nowell, A. (1999). Changes in the black-white gap in achievement test scores. *Sociology of Education, 72*(2), 111.

Hyde, J.S., Lindberg, S.M., Linn M.C., Ellis, A.B. & Williams C.C. (2008). Gender similarities characterize math performance. *Science (New York, N.Y.), 321*(5888), 494-495.

Hyde, J. S. (1990). Meta-analysis and the psychology of gender differences. *Signs: Journal of Women in Culture and Society, 16*, 55.

Jacobs, J. Lanza, S. Osgood, D. Eccles, J. & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve *Child Development*, Vol. 73, No. 2 (Mar. - Apr., 2002), pp. 509-527

Jacobson, L. (2002, March 3). Longitudinal study finds gender and race gaps among first graders. *Education Week, 21*(27), 7.

Johnson, S.P. (2008). *The status of male teachers in public education today*. Education Policy Brief 6(4) Indiana University: Center for Evaluation and Education Policy.

Kainz, K., & Vernon-Feagans, L. (2007). The ecology of early reading development for children in poverty. *Elementary School Journal, 107*(5), 407-427.

Kimmel, M. (2006). A war against boys?. *Dissent, 53*(4), 65-70.

King, J. R. (1998). *Uncommon caring :Learning from men who teach young children*. New York: Teachers College Press.

King, K., & Gurian, M. (2006). Teaching to the minds of boys. *Educational Leadership, 64*(1), 56-61.

Lemkau, J. P. (1984). Men in female-dominated professions: Distinguishing personality and background features. *Journal of Vocational Behavior, 24*(1), 110-122.

Lummis, M. & Stevenson, H.W. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology, 26*(2), 254-263.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Maine Boys Network. (2008, October). *The gender divide in academic engagement. Perspectives from Maine boys and young men*. Maine: Colby-Bates-Bowdoin Collaborative Faculty Development Grant from the Andrew W. Mellon Foundation.

Mancus, D.S. (1992). Influence of male teachers on elementary school children's stereotyping of teacher competence. *Sex Roles: A Journal of Research, 26*, 109-28.

Marsh, H. W. (1998) Longitudinal structural equation models of academic self-Concept and achievement: Gender differences in the development of math and english constructs. *American Educational Research Journal*, *35*(4), 705-738.

Martin, C. L., & Halverson, C. F.,Jr. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development*, *52*(4), 1119-34.

McCracken, J.H. (1973). Sex typing of reading by boys attending all male classes. *Developmental Psychology, 8*(1), 148.

McCullen, M.B. & Kazat A. (2002). Education matters in the nurturing of the beliefs of preschool caregivers and teachers. *Early childhood research and practice* 4(2) Retrieved online at: http://ecrp.uiuc.edu/v4n1/saluja.html

McIntyre, L. (1988). Teacher gender: A predictor of special education referral?. *Journal of Learning Disabilities, 21*(6), 382-383.

Meece, J.L., Parsons, J.E., Kaczala, C.M., & Goff, S.B. (1982). Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin, 91*, 324-348.

National Institute of Child Health and Human Development and Early Child Care Research Network (2002). The relation of global first grade classroom environment to structural classroom features, teacher, and student behaviors. *The Elementary School Journal, 102*(5), 367.

O'Connor, E., & McCartney, K. (2007). Examining teacher-child relationships and achievement as part of an ecological model of development. *American Educational Research Journal*, *44*(2), 340-369.

Palardy, G.J., & Rumberger, R.W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation & Policy Analysis, 30*(2), 111-140.

Perkins, L.M. (1989). The history of blacks within teaching: Growth and decline within the profession. In D. Warren (Ed.), *American teachers : Histories of a profession at work.* (pp. 344-369) New York: Macmillan.

Perlmann, J. & Margo, R.A. (2001). *Women's work?: American schoolteachers, 1650-1920*. Chicago: University of Chicago Press.

Peterson, P. L., & Fennema, E. (1985). Effective teaching, student engagement in classroom activities, and sex-related differences in learning mathematics. *American Educational Research Journal*, *22*(3), 309-35.

Pianta, R.C. LaParo, K.M. Payne, C. Cox, M.J. & Bradley, R. (2002). The relation of kindergarten classroom environment to teacher, family, and school characteristics and child outcomes. *The Elementary School Journal, 102*(3), 225.

Raudenbush, S.W. & Bryk, A.S.. (2002). *Hierarchical linear models: Applications for data analysis methods (Second edition).* Thousand Oaks, CA: Sage.

Ready, D. D., LoGerfo, L. F., & Burkam, D. T. (2005). Explaining girls' advantage in kindergarten literacy learning: Do classroom behaviors make a difference? *The Elementary School Journal*, *106*(1), 21-38.

Rosenthal, R., and Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review, 3*(1), 16-20.

Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In N. Eisenberg, W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 3, social, emotional, and personality development (6th ed.).* (pp. 858-932) John Wiley & Sons Inc.

Rury, J.L. (1989). Who became teachers?: The social characteristics of teachers in American history. In D. Warren (Ed.), *American teachers : Histories of a profession at work*. (pp. 9-48) New York: Macmillan.

Sabbe, E. & Aelterman, A. (2007). Gender in teaching: a literature review. *Teachers and teaching, theory and practice, 13*(5), 521.

Sadker, M., Sadker, D. & Klein, S. (1991). The issue of gender in elementary and secondary education. *Review of Research in Education, 17*(2), 269.

Sargent, P. (2001). *Real men or real teachers?: Contradictions in the lives of men elementary school teachers* (1st ed.). Harriman, Tenn.: Men's Studies Press.

Scelfo, J. (2007, September 17). Come back Mr. Chips: Stereotyping, low pay, lack of role models. Why the number of men teaching in schools is at a 40-year low. *Newsweek, 150*(12), 44.

Schweigardt, W., Worrell, F., & Hale, R. (2001). Gender Differences in the Motivation for and Selection of Courses in a Summer Program for Academically Talented Students. *Gifted Child Quarterly 45***,** 283-293

Sedlak, M., Schlossman, S., & Rand Corp., S. M.,CA. (1986). *Who will teach? historical perspectives on the changing appeal of teaching as a profession*

Sexton, P. C. (1969). *The feminized male; classrooms, white collars, & the decline of manliness*. New York: Random House.

Shen, J., Wegenke, G. L., & Cooley, V. E. (2003). Has the public teaching force become more diversified? National and longitudinal perspectives on gender, race, and ethnicity. *Educational Horizons, 81*(3), 112-18.

Smith, E. Kirchner, H.L. Taylor, M. Hoffman, H. & Lemke, J.H. (1998). Voice problems among teachers: Differences by gender and teaching characteristics. *Journal of Voice*, 12(3), 328-334.

Snyder, T. D., Dillow, S. A., Hoffman, C. M., & National Center for Education Statistics (ED), Washington,DC. (2008). *Digest of education statistics, 2007. NCES 2008-022* National Center for Education Statistics.

Sommers, C. H. (2000). *The war against boys : How misguided feminism is harming our young men*. New York: Simon & Schuster.

Stake, J.E. & Katz, J.F. (1982). Teacher-pupil relationships in the elementary school classroom: Teacher-gender and pupil-gender differences. *American Educational Research Journal, 19*(3), 465-471.

Stinnett, T. A., Bull, K. S., Koonce, D. A., & Aldridge, J. O. (1999). Effects of diagnostic label, race, gender, educational placement, and definitional information on prognostic outlook for children with behavior problems. *Psychology in the Schools, 36*(1), 51-59.

Strizek, G.A., Pittsonberger, J.L., Riordan, K.E., Lyter, D.M. & Orlofsky, G.F. (2006). *Characteristics of schools, districts, teachers, and school libraries in the United States: 2003-2004 Schools and staffing survey* (NCES 2006-313) U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Sugg, R. S. (1978). *Motherteacher :The feminization of american education*. Charlottesville: University Press of Virginia.

Tomada, G., & Schneider, B. H. (1997). Relational aggression, gender, and peer acceptance: Invariance across culture, stability over time, and concordance among informants. *Developmental Psychology*, *33*(4), 601-09.

U.S. Department of Education, National Center for Education Statistics. (2001a). *Trends in educational equity of girls and women.* (NCES 2000-030R) by Yupin Bae, Susan Choy, Claire Geddes, Jennifer Sable and Thomas Snyder. Washington,DC.

U.S. Department of Education, National Center for Education Statistics (2001b). *User's manual for the ECLS-K base year public-use data files and electronic codebook* (NCES 2001-029 revised). Washington, D.C.: NCES

U.S. Department of Education, National Center for Education Statistics. (2004). *Kindergarten teachers: Public and private school teachers of the kindergarten class of 1998-99. NCES 2004-060* by Elvira Germino-Hausken, Jill Walston, and Amy H. Rathbun. Project officer: Elvira Germino-Hausken. Washington,DC.

Vernon-Feagans, L. (1996). *Children's talk in communities and classrooms*. Cambridge, Mass.: Blackwell Publishers.

Warren, D.R. (1989). Teachers, reformers and historians. In D. Warren (Ed.), *American teachers : Histories of a profession at work*. (pp. 1-6) New York: Macmillan.

Wayne, A.J. & Youngs, P.Y. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*(1), 89-122.

Weaver-Hightower, M. (2003). The "boy turn" in research on gender and education. *Review of Educational Research, 73*(4), 471-498.

Wehmeyer, M. L., & Schwartz, M. (2001). Disproportionate representation of males in special education services: Biology, behavior, or bias? *Education and Treatment of Children, 24*(1), 28-45.

West, B. T., Welch, K. B., & Gałecki, A. T. (2007). *Linear mixed models : A practical guide using statistical software.* Boca Raton: Chapman & Hall/CRC.

Williams, C. L. (1992). The glass escalator: Hidden advantages for men in “female” professions. *Social Problems, 39*(3), 253-267.

Williams, C. L. (1995). *Still a man's world: Men who do "women's" work*. Berkeley, CA: University of California Press.

Wright, S.P. Horn, S.P. & Sanders,W.L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57-67.

Xue, Y., & Meisels, S. J. (2004). Early literacy instruction and learning in kindergarten: Evidence from the early childhood longitudinal study--kindergarten class of 1998-1999. *American Educational Research Journal, 41*(1), 191-229.