Examining the Relationship between Teacher Math-Gender Stereotypes and Students' Math Outcomes

Rachel Conlon

FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

EXAMINING THE RELATIONSHIP BETWEEN TEACHER MATH-GENDER STEREOTYPES AND STUDENTS' MATH OUTCOMES

By

RACHEL CONLON

A Thesis submitted to the
Department of Psychology
in partial fulfillment of the
requirements for the degree of
Master of Science

2020

Rachel Conlon defended this thesis on November 17, 2020.
The members of the supervisory committee were:

Colleen M. Ganley
Professor Directing Thesis

Sara Hart
Committee Member

E. Ashby Plant
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the thesis has been approved in accordance with university requirements.

To Mam and Dad;

For everything

ACKNOWLEDGMENTS

Firstly, I would like to thank my committee members for their assistance and expertise. A huge thank you to my wonderful advisor, Dr Colleen Ganley, for her endless patience and guidance over the past couple of years, and for caring about me as a person just as much as she cares about my work. To my fellow Ganley Lab graduate students, both past and present, I appreciate every bit of your support and reassurance. To my roommates, fellow cohort members, and partner, thank you for being there for me through every moment of the process, and for making these past few years all the more enjoyable because I got to spend them will you all. Lastly, and most importantly, thank you to my amazing parents: for raising me to be curious, for encouraging me to me brave, and for supporting me to move halfway across the world to pursue my dreams. This accomplishment is every bit as much yours as it is mine.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Common stereotypes contend that boys have more natural ability and interest in math than do girls, which may contribute to the underrepresentation of women in some Science, Technology, Engineering and Math fields. Research suggests that many teachers endorse these stereotypes to some extent. These stereotypes have the potential to be transferred to their students via their behavior, which may impact math self-concept and achievement, particularly in girls. Thus, we hypothesized that math-gender stereotyping in teachers would be associated with lower math achievement and self-concept in their female, but not male students, and that student math-gender stereotyping would mediate these relationships. In a sample of 2,387 kindergarten through 3rd grade students and their 223 teachers from 30 schools in the United States, we measured student math achievement and self-concept in Fall and Spring of the 2018/19 school year. We also measured math-gender stereotyped beliefs in both the teachers and the 3rd grade students from the sample. We found evidence of male-biased math-gender stereotypes in both teachers and male students, and female-biased stereotypes in female students. Using multi-level models, we found that there were no significant relations between teacher stereotyping and student math achievement or self-concept. We also found no relation between teacher stereotyping and student stereotyping in the 3rd grade students, and thus did not test the overall mediation model. However, when stereotypes about men and women were examined separately from those about girls and boys, we found that male-biased stereotyping about adults in teachers was associated with higher math achievement overall, and higher stereotyping and lower math self-concept in 3rd grade boys. We suggest that teacher math-gender stereotyping overall may not be a particularly effective avenue for intervention to improve math outcomes for female students, but that the influence of stereotypes about adults specifically merits exploration. Future research

with a developmental approach is required understand how these relations may change over time and what potential gender stereotyping may have as a candidate for intervention as children grow.

CHAPTER 1

INTRODUCTION

The underrepresentation of women in certain Science, Technology, Engineering and Mathematics (STEM) subfields has been a topic of major interest in recent years. Whilst great strides have been made towards gender equity in the workplace, STEM fields such as computer science and engineering continue to see large gender imbalances. For example, the US Census Bureau's American Community Survey has indicated that whilst women make up around 47% of the total US workforce, they comprise only 24% of the country's STEM workers (Noonan, 2017). Women's avoidance of STEM careers is problematic given the growing demand for workers in many STEM-focused industries (Xue & Larson, 2015). In addition, many of the most highly paid careers are concentrated in STEM, so women's underrepresentation in these fields may be contributing to the gender pay gap, and thus perpetuating gender inequality more widely (Beede et al., 2011).

Given research suggesting that self-concept and math achievement predict later STEM career aspirations (Sells, 1980; Seo et al., 2018), and that girls typically express lower levels of math self-concept and have marginally lower math achievement than boys (Good et al., 2012), examining the factors that influence girls' math achievement and self-concept could shed light on some of the contributors to these imbalances and provide pathways to address them.

Male-biased stereotypes about math abilities are prevalent, and may be an important contributor to girls' relatively poorer math self-concept and achievement. Understanding how math-gender stereotypes develop and how they contribute to students' math attitudes is key to devising ways to counteract them. Given that children are highly influenced by the views and behavior of the adults around them, teachers' own stereotyped views about girls' math abilities, and their subsequent influence on teacher behavior, may contribute to the absorption of these views by their students. This, in turn, may reduce girls' confidence in their math abilities, as well as negatively impacting their performance. This could result in a lower likelihood of pursuing STEM careers in the future.

The present investigation seeks to examine the relationship between teacher math-gender stereotyping and student math outcomes using a longitudinal design. Specifically, we test whether there is a relationship between math-gender stereotyping in teachers, measured in the beginning of a school year, and their students' math-gender stereotyping, math self-concept, and math achievement at the end of the year. We also examine if any such relationship could be partially explained by students' own stereotypes about gender and math ability, and whether any of these relationships are different for boys and girls

Math-Gender Stereotypes

Whilst there is some evidence for a gender gap in math performance in the US (e.g. Cimpian et al., 2016), a meta-analysis by Lindberg et al. (2010) found that these differences are effectively non-existent at the elementary school level, and data from the 2015 Trends in International Math and Science Study (TIMSS) suggests that, overall, boys and girls have similar levels of mathematical ability, with male advantages in math performance largely narrowing over time (Mullis et al., 2016). Despite this, school-aged girls have been found to express less interest in math than boys (Marsh et al., 2005) and are less likely to pursue academic and career paths in STEM fields (Beede et al., 2011). The prevalence of male-biased math-gender stereotypes has often been proposed as contributing to this disparity. These stereotypes contend that boys have more natural talent for mathematical studies and have greater natural interest in mathematical pursuits (Master & Meltzoff, 2016; Nosek et al., 2009). This can make girls feel as though math is not ‘for them,’ impacting performance and self-concept, and impairing math engagement (Bian et al., 2017).

Research suggests that children demonstrate awareness of math-gender stereotypes from a young age. Cvencek et al. (2011) found that, by the beginning of elementary school, children were already aware, at least implicitly, that math is considered a ‘male domain’. However, children who are aware of math-gender stereotypes do not necessarily endorse them, with research indicating that younger children tend not to explicitly endorse these stereotypes (Tomasetto et al., 2011). Instead, they display small amounts of math-gender stereotyping with ingroup bias - girls believe that girls are better at math and boys believe that boys are better (Heyman & Legare, 2004). Around age 9 or 10, these attitudes shift towards endorsement of more traditional math-gender stereotypes, with both sexes more likely to view boys as being better at math (Kurtz-Costes et al., 2008; Steffens et al., 2010). This stereotype endorsement becomes more prevalent by adolescence, with girls typically holding stronger implicit math-gender stereotyped beliefs than boys (Steffens et al., 2010). These trends indicate that the early to mid-elementary school period is a key time for math-gender stereotype development and, subsequently, a crucial period at which to examine this topic.

Many factors contribute to the development of stereotyped beliefs in children. Children absorb ideas about the world from the people around them, particularly those with whom they spend the most time: their peers, their families, and their teachers (Tiedemann, 2000). In their position as authority figures and facilitators of both formal and informal learning, adults have a strong potential to influence the children in their care through their own stereotyped views. A number of studies have found that adults’ stereotyped views can influence their expectations for children’s performance in various domains, which in turn can influence the views of the children themselves (e.g. Jacobs, 1991; Lummis & Stevenson, 1990; see Gunderson et al., 2012 for a review). Children spend a large portion of their day in the classroom, where teachers can influence not only an individual child, but also their peers, allowing beliefs and ideas to be both transmitted directly from the teacher and reinforced by their peers (Gunderson et al., 2012). Thus, given teachers' important role in the development of academic attitudes (Tiedemann, 2000), the math-gender stereotypes they hold have a strong potential to influence their students' own attitudes about gender and math

Teacher Math-Gender Stereotyping

A number of studies have shown that many teachers hold stereotypic beliefs about math and gender (see Li, 1999 for a review). In a study of first-grade teachers by Fennema et al. (1990), teachers were found to believe boys are more logical and independent at math, and they were more likely to overestimate boys' math performance and underestimate that of girls. Boys' success in math was typically attributed to natural talent, while girls' performance was attributed to both ability and effort (Fennema et al., 1990). Gunderson et al. (2012) proposed that these kinds of differential attributions may account for the steadfastness of math-gender stereotypes, despite evidence of largely equal math abilities across genders – math is seen to 'come more easily' to boys than to girls.

Research has demonstrated a positive relationship between gender stereotyping in teachers and their students, such that teachers who more strongly endorse math-gender stereotypes tend to have students in their classroom who more strongly endorse those stereotypes (Keller, 2001). Though the size of this effect was relatively small, the finding remained significant even when controlling for achievement, interest, and math self-concept. Fennema et al. (1990) propose that teachers who hold math-gender stereotypes transmit these through their instruction. A gender-stereotyped classroom environment and peer reinforcement have also been posited as contributing to this effect (Keller, 2001). Gender stereotyped beliefs might lead teachers to unconsciously interact differently with male and female students, which may have further consequences for children's math self-concept and achievement. For example, an observational study by Becker (1981) found that teachers interacted more with boys in the classroom, asking them more questions and providing more feedback following incorrect answers.

More recent research (e.g., Harrop & Swinson, 2011) has failed to replicate these findings, suggesting that overt differences in teacher-student interaction may have diminished over time. However, the possibility of more subtle differences in teaching practices remains. For example, a study in pre-schools by Simpson and Linder (2016) found that while there were few differences in the amount of mathematical talk directed at girls and boys by teachers, the context of the interactions tended to differ by gender, with mathematical talk directed at boys more often in academic contexts (e.g. in the math/science area of the classroom), and girls in more domestic contexts (e.g. whilst playing house). Thus, while teachers do not explicitly discourage girls' interest in STEM, they may still be unintentionally reinforcing the stereotype that STEM careers are unsuitable for them.

It has been shown that children particularly attend to and model the beliefs and behavior of same-gender adults (Gunderson et al., 2012). Given the vast majority of elementary school teachers are female (89% in 2015/16; National Center for Education Statistics, 2020), math-gender stereotypes may have a disproportionately large effect on girls compared to boys. This is supported by findings from Beilock and colleagues (2010) showing that female teachers' math anxiety was related to more math-gender stereotyping amongst female, but not male, students across an academic year. However, it is unclear if teachers' own gender stereotyping contributes to this relationship. Ultimately, a greater understanding of how teachers' gender stereotyping may influence their students is of importance given the influential role teachers play in a child's development.

Student Gender Stereotype Influences on Math Self-Concept and Performance

Awareness and endorsement of math-gender stereotypes have been associated with negative outcomes for girls' math performance in a number of studies. For example, a study of Singaporean students by Cvencek et al. (2015) found a significant relationship between implicit math-gender stereotyping and student math achievement. Beilock et al. (2010) measured gender stereotype endorsement by asking children to draw the subject of a story about a child who was good at math, or a child who was good at reading. Similarly, they found that girls who drew a picture of a girl for the reading story, and a boy for the math story had significantly poorer math achievement than those who did not. It must be noted that these findings do not speak to directionality, so it may be that lower math achievement leads girls to endorse these stereotypes, or that endorsing stereotypes leads to lower achievement. Interestingly, Beilock et al. (2010) found no difference in math achievement between boys who endorsed male-biased math-gender stereotypes and those who did not, which is contrary to what we might expect if achievement influenced stereotype endorsement. Thus, math-gender stereotypes seem connected to math achievement for girls, whereas there is a lack of evidence pointing to a similar relationship in boys. This suggests that the relation between math-gender stereotypes and girls' math achievement requires particular attention.

Gender stereotyping has also been associated with girls' lower math self-concept, or the extent to which one identifies with, and believes they have the capacity to do well at, math. Math self-concept is an important predictor for both math performance and STEM career ambitions (Seo et al., 2018). Low math self-concept can result in more negative attitudes towards math, potentially leading to math anxiety and math avoidance, which ultimately limit math ambitions (Marsh et al., 2005). Studies using both US and international samples have indicated that girls have lower math ability self-concept than do boys (e.g. Else-Quest et al., 2010; Ganley & Lubienski, 2016), and that gender stereotyping may play a role in this relationship (e.g. Cvencek et al., 2011). Consistent with this view, research with Italian children by Passolunghi et al. (2014) indicated that explicit, but not implicit, math-gender stereotyping was associated with higher math self-concept in boys, and lower math self-concept in girls.

In addition, Dickhäuser and Meyer (2006) found that girls were less likely than boys to attribute their math successes to high ability, and more likely to attribute failure to low ability than boys. This mirrors similar attitudes seen in teachers, as discussed above, which suggests that girls may be picking up on and mirroring their teachers' beliefs that they may have less 'natural talent' for math. Given findings by Barth et al. (2018) that ability beliefs are strongly related to STEM career choice, even among high-performing math students, it is important to examine the source of these often inaccurate ability beliefs to avoid talented women turning away from STEM careers.

Present Study

Whilst there is some evidence to support a link between teacher math-gender stereotyping and students' development of their own stereotyped beliefs, as well as between students' beliefs and their math achievement or self-concept, whether there is a relationship between teacher stereotyping and student achievement and self-concept remains unstudied and therefore unclear. Thus, the current study aims to address several research questions: Does teacher math-gender stereotyping measured in the Fall of a school year relate to student math self-concept or math achievement measured in the Spring (1a); does teacher math-gender stereotyping in the Fall relate to student math-gender stereotyping in Spring? (1b); and, if such associations exist, is the relationship between teacher math-gender stereotyping and student math self-concept or achievement mediated by student math-gender stereotyping (2). For each of these research questions, I will also examine if the results are different for boys and girls.

I hypothesize that when teachers express stronger male-biased math-gender stereotyped beliefs at the beginning of the school year, this will relate to lower math achievement and math self-concept for their female, but not male, students at the end of the school year, and that this relationship will be explained, at least partially, by the students' own math-gender stereotypes.

CHAPTER 2

METHODS

Participants

Data were drawn from the Research on Experiences, Attitudes, and Learning in Math (REALM) Project - a longitudinal study investigating the relationship between kindergarten through 3rd grade teacher math anxiety and student math attitudes and achievement. Participants in the study were 2,847 students and their 233 teachers from 30 public elementary schools across 4 school districts in one state in the United States. Three teachers were removed from the final analytic sample because they did not complete the teacher measures in the fall. Two of these teachers had not completed these measures because they became the classroom teacher after fall teacher surveys were completed, and one did not complete the measures for unknown reasons. Their 34 students were also removed from the sample. A further six teachers and their 41 students were excluded from the final sample because their classes did not complete the spring measures. Two of these teachers left their position during the school year and their replacement teachers did not agree to administer test materials to the students, three teachers chose not to return spring materials, and one teacher was on medical leave during the spring data collection period. One student was removed because they switched grade levels during the school year, and one teacher and their two students were removed as the teacher taught multiple grade levels. Finally, 366 students were removed because they completed neither the math self-concept nor math achievement measures in spring. Thus, in total, 10 teachers and 444 students were removed from the dataset, leaving a final analytic sample of 223 teachers and their 2,387 students. The 3rd grade subsample consisted of 58 teachers and their 642 students.

Teachers received financial compensation for their participation, and students were compensated with a pencil and bookmark. Both parental consent and student written assent were required for students' responses to be included in analyses. We obtained information about student race/ethnicity, gender, free/reduced lunch status, parent occupation, and parent education from the students' parents via a paper survey sent home with each student. Demographic information for teachers and students is available in Tables 1a and 1b.

Teachers had an average of 14.62 years teaching experience (*SD* = 8.83). The majority of teachers were female (93.7%), 2.2% identified as male and 4% did not provide their gender. Seventy-seven percent of teachers identified as white, 13.9% as Black, 0.4% as American Indian or Alaskan Native, 2.7% as multiracial and 0.4% as another race, with 5.4% not providing their race.

Table 1

Demographic Information for Teacher Analytic Sample

Variable	Group	n	%	Valid %
Grade	Kindergarten	53	23.8	23.8
	1st	56	25.1	25.1
	2nd	56	25.1	25.1
	3rd	58	26.0	26.0
Gender	Male	5	2.2	2.3
	Female	209	93.7	97.7
	Missing	9	4.0	-
Race	American Indian or Native Alaskan	1	0.4	0.5
	Black	31	13.9	14.7
	White	172	77.1	81.5
	Other	1	0.4	0.5
	Multirace	6	2.7	2.8
	Missing	12	5.4	-

	M	SD	Min	Max
Teaching Experience (years)	14.62	8.83	1	46

Table 2
Demographic Information for Student Analytic Sample

Variable	Group	n	%	Valid %
Grade	Kindergarten	523	21.9	21.9
	1st	595	24.9	24.9
	2nd	627	26.3	26.3
	3rd	642	26.9	26.9
Gender	Boys	1138	47.7	50.5
	Girls	1162	48.7	49.5
	Missing	87	3.6	-
Race	Asian	96	4.0	4.2
	Black	484	20.0	21.3
	Hispanic	279	11.7	12.3
	White	1112	46.6	48.9
	Other	11	0.5	0.5
	Multirace	290	12.1	12.8
	Missing	115	4.8	-
Lunch Status	Free or Reduced Lunch	997	41.8	47.2
	None	1116	46.8	52.8
	Missing	274	11.5	-
Parent Education 1	Less than 9th grade	18	0.8	0.9
	9th to 12th grade, no diploma	89	3.7	4.2
	High school graduate (including equivalency GED)	279	11.7	13.2
	Some college, no degree	463	19.4	12.9
	Associate's degree	258	10.8	12.2
	Bachelor's degree	538	22.5	25.5
	Graduate or professional degree	467	19.6	22.1
	Missing	275	11.5	-
Parent Education 2	Less than 9th grade	24	1.0	1.5
	9th to 12th grade, no diploma	76	3.2	4.9
	High school graduate (including equivalency GED)	283	11.9	18.1
	Some college, no degree	274	11.5	17.5
	Associate's degree	154	6.5	9.9
	Bachelor's degree	374	15.7	23.9
	Graduate or professional degree	378	15.8	24.2
	Missing	824	34.5	-

	M	SD	Max	Min
Age (in months)	94.37	14.06	67	142

Students came from a wide variety of backgrounds, with 41.8% qualifying for free or reduced-price lunch. In terms of racial identity, 46.6% of students identified as white, 20.3% as Black, 11.7% as Hispanic, 4% as Asian, 12.1% as multiracial, and 0.5% as another race, while 4.8 percent did not provide information on their race. Roughly half of students (48.7%) identified as female, 47.7% identified as male, and gender was unavailable for 3.6% of students. Students were on average 7.86 years old (SD=1.17 years) with ages ranging from 5.58 to 11.83 years.

Materials

Gender stereotyped beliefs and knowledge. The same measure of gender stereotyped beliefs and awareness was completed by teachers from all grades, as well as students from 3rd grade only. The eight items in the scale were adapted from Steele (2003), with some additional novel items included by the researchers. Items related to reading, and items about stereotype awareness were completed but were not analyzed for the current study. To assess math-gender stereotyped beliefs, participants responded to questions asking how good they believe most girls, boys, men, and women to be at math (e.g. *"How good are most boys at math?"*). Answers were in the form of a 4-point Likert scale *(0=Not very good, 1=a little bit good, 2=mostly good, 3=very good).* Scores for items about girls and women were subtracted from those for items about men and boys to create a math-gender stereotyping score. These scores were then combined, resulting in scores from negative six to positive 6, with values above zero indicating male-biased math gender stereotyping and values below zero indicating female-biased math-gender stereotyping.

Math achievement assessment. At Wave 1, students completed the Elementary Mathematics Student Assessment (EMSA; adapted from the Fall 2015 version of the EMSA; Schoen et al., 2018a) which is a math achievement measure featuring a mixture of multiple-choice and constructed-response questions in both numeric and word problem formats. The measure was designed to test math knowledge and assessed the following mathematical domains: 1) counting and basic number facts, 2) word problems, and 3) number relations, fractions, and computation. In fall, the kindergarten test form consisted of 16 items, the 1st grade test had 22 items, the 2nd grade test had 27 items, and the grade 3 test had 28 items. In spring, the kindergarten test form consisted of 17 items, the 1st grade test had 22 items, the 2nd grade test had 25 items, and the grade 3 test had 28 items. The Wave 2 assessment (adapted from the Spring 2016 version of the EMSA; Schoen et al., 2018b) featured some more challenging questions to account for what students would have learned throughout the school year, but some items overlapped with the Wave 1 assessment in a common-item design (Kolen & Brennan, 2004). Test scores were determined using a two-parameter logistic model based on item-response theory. The expected a posteriori (EAP) method was used to estimate the person ability in each grade level, and these ability estimates were mapped onto a single scale by the Stocking-Lord method for vertical equating (Kolen & Brennan, 2014) which allows for comparing scores between grades. Thus, a student's EMSA score for each wave was operationalized as a theta score, with a higher score indicating a higher level of math achievement and a lower score indicating a lower level of math achievement. The test proved to be reliable across all four grade levels (kindergarten α = .77, 1st grade α = .79, 2nd grade α = .82, 3rd grade α = .87).

Child math self-concept scale. At both timepoints, students completed a five-item math self-concept measure adapted from Fredricks & Eccles' (2002) Math Competence measure. Students were asked to respond to five statements surrounding their feelings about their math abilities (e.g., *"I am good at math," "I think I could learn how to solve hard math problems")* on a 4-point Likert scale *(No, Not really, Kind of, or Yes).* In this study, the scale was found to have reasonable reliability at both timepoints for the overall sample, (Wave 1 α = .73, Wave 2 α = .69) and the 3rd grade subsample (Wave 1 α = .69, Wave 2 α = .72).

Procedure

In fall 2018 (Wave 1), all teachers completed the gender-stereotyped beliefs and knowledge scale online as part of 90 minutes of surveys and assessments delivered via the Qualtrics platform. To assess math self-concept and achievement, all students completed the child math self-concept scale (as part of a 15-minute math attitudes questionnaire) and the 30-minute math achievement assessment. Additionally, the 3rd grade students completed the same gender-stereotyped beliefs and knowledge scale as the teachers. In spring 2019 (Wave 2), students repeated the previous assessments. At Waves 1 and 2, both students and teachers also completed a number of other measures and assessments which are beyond the scope of this project. At both timepoints, teachers administered all student measures in their regular classroom using materials provided by the researchers. Each math test item and math self-concept item was read aloud by the teacher to ensure students could comprehend the item, and students were instructed to read along with them in the provided survey booklet, and then respond to the question in the space provided. The gender stereotyping items were not read aloud to students, as we were concerned that having teachers reading these out loud to their students may impact students' responses.

Analysis Plan

Outlier management. We planned to specifically focus on outliers that were due to error (i.e., data entry error, child response error) in these data. To identify error outliers, we planned to run descriptive statistics on each variable to confirm that there were no impossible values included in these data. If there were data points that were extreme but reflected possible values and we had no reason to believe that they are incorrect (e.g., the mean for the item may be a rating of 2 on a 4-point scale but there were scores of 4 from some individuals), they were to be kept in the data set. Any data entry errors for individual data points that we believed were true inaccuracies were to be corrected for or coded as missing in our final data set and the reasons for data exclusion have been noted.

Preliminary analyses. Due to social desirability effects in responses to stereotyping questions, we had concern about the potential that stereotyping scores would lack the variability required to perform the analyses below. As such, we planned to examine the distribution of stereotyping responses prior to conducting any further analyses. If scores lacked sufficient variability, teachers and/or students would be divided into three categories depending on whether their score leaned towards stereotyping boys to be better at math, girls to be better, or no gender differences and those categories would be utilized in subsequent analyses.

Research question 1a: Do teacher math-gender stereotyping scores at Wave 1 relate to student self-concept or achievement at Wave 2, and are these relationships different for boys and girls? We planned to examine the relationships between teacher math-gender stereotyping and student math self-concept or student math achievement using multi-level modeling with the entire sample (see Figure 1). For our main analyses we planned to use the combined gender stereotyping scores from the items about boys and girls and men and women, but also to conduct supplementary analyses with these modeled separately, based on evidence that adults (Powlishta, 2000) and children (Steele, 2003) often hold different stereotypes about boys and girls versus men and women. We planned to use Wave 1 teacher math-gender stereotyping (level 2) to predict student math self-concept and math achievement at Wave 2 (level 1). Wave 1 math self-concept and math achievement were to be included as covariates in the model to account for initial levels of these variables. Student age in months and SES (comprised of combined z-scores of student lunch status, parent level of education and parent occupational prestige) were also to be included as covariates. A significant path estimate between Wave 1 teacher math-gender stereotyping and Wave 2 student math self-concept and/or achievement would provide support for the hypothesis that the extent to which a teacher holds stereotypical math-gender stereotyped beliefs at Wave 1 relates to student achievement and/or math self-concept at Wave 2.

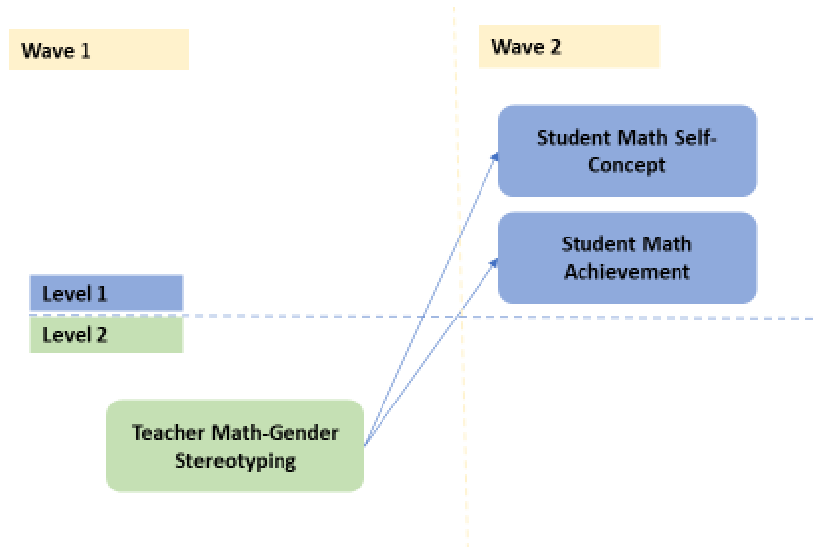

Figure 1
Proposed Model of Relationship Between Teacher Math-Gender Stereotyping and Student Achievement and Math Self-Concept
Note. Model includes student math-gender self-concept and achievement at Wave 1, grade, student age, and SES as covariates.

We intended to use a multiple group approach, as recommended by Kline (2005) to examine if these relationships differ by gender. First, both a free model (where intercepts and path are allowed to vary across groups) and constrained model (where intercepts and paths are fixed) would be estimated. If these models are significantly different, we would then proceed to explore which paths are different between the two groups. We would then constrain and release the various paths of interest (from teacher math-gender stereotyping to math self-concept and from teacher stereotyping to math achievement) to examine if there are group differences in these paths. We then planned to repeat the analyses with a subsample of only 3rd grade teachers (approx. n = 56) and their students to align with the subsequent research questions below, which utilize data from this subsample only, as student gender stereotyping data was not available for the other grades.

Research question 1b: Do teacher math-gender stereotyping scores at Wave 1 relate to student math-gender stereotype scores at Wave 2, and are these relationships different for boys and girls? Next, we proposed to examine the relationship between teacher math-gender stereotyping and student math-gender stereotyping. First, using data from the students and teachers from the 3rd grade cohort only (as math-gender stereotyping data was not collected from students in the younger grade levels), the above analyses would be repeated with Wave 2 student math-gender stereotyping as the outcome variable (see Figure 2).

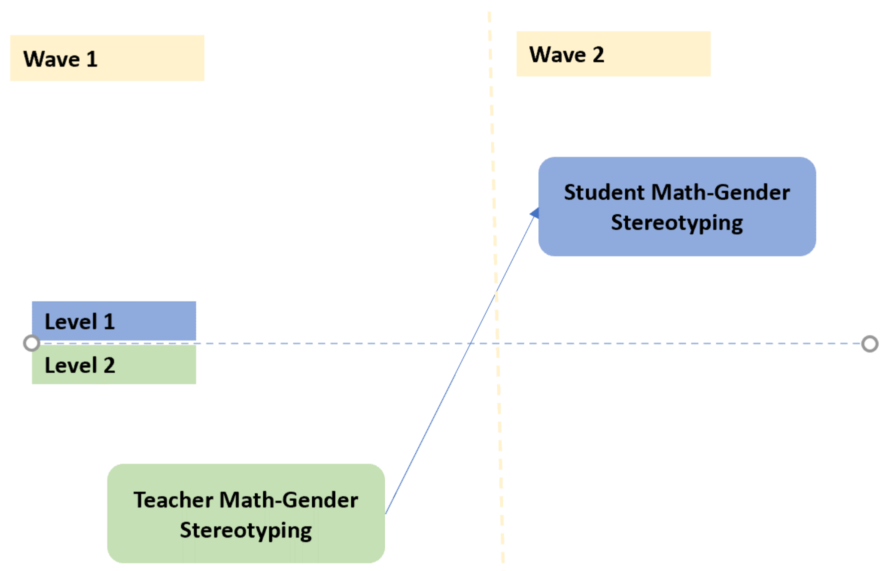

Figure 2

Proposed Model of Relationship Between Teacher and Student Math-Gender Stereotyping

Note. Model includes student age and SES as covariates.

A significant path estimate between Wave 1 teacher math-gender stereotyping and Wave 2 student math-gender stereotyping would provide support for the hypothesis that teachers can transmit their gender-stereotyped beliefs about math to their students. As for research question 1a, we again planned to use a multi-group approach to examine if these relationships differ by gender. We also again planned to test these models with gender-math stereotyping combined as our main model, and then separated into boys/girls and men/women for supplementary analyses.

Research question 2: Is the relationship between teacher math-gender stereotyping and student math self-concept or achievement mediated by student math-gender stereotyping, and are these relationships different for boys and girls? If significant relationships were found in the preceding analyses for the 3rd grade subsample, we would then combine the two models, to see if student math-gender stereotyping acts as a mediator of the relationship between teacher math-gender stereotyping and student math self-concept or student math achievement, and once again if any relationship is different for boys and girls.

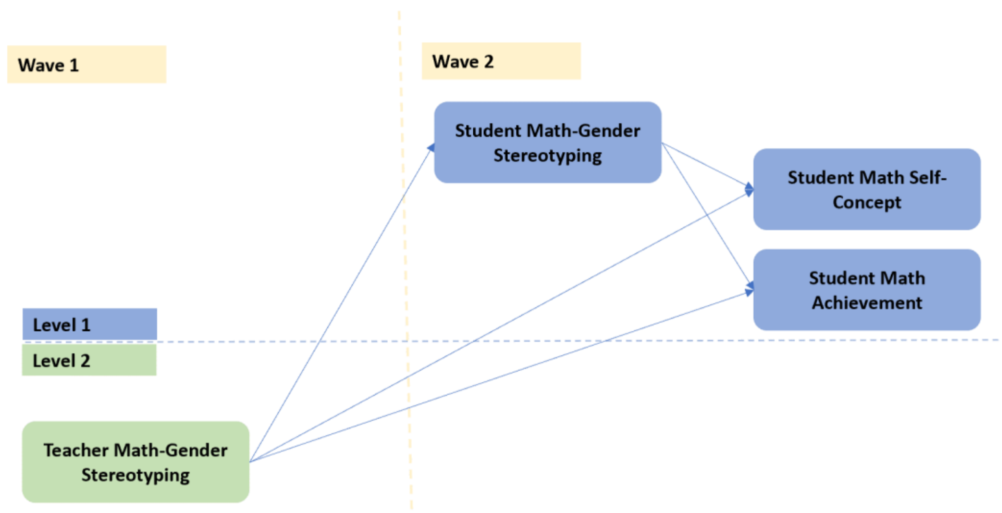

Figure 3
Proposed Model of Relationship Between Teacher Math-Gender Stereotyping and Student Math Achievement and Self-Concept Mediated by Student Math-Gender Stereotyping
Note. Model includes student math self-concept and achievement at Wave 1, student age, and SES as covariates.

Specifically, we planned to use the framework and models proposed by Preacher, Zyphur, & Zhang (2010) to investigate student math-gender stereotyping as a potential mediator of teacher math-gender stereotyping (see Figure 3). In these models, teacher math-gender stereotyping (level 2) is the independent variable, student math-gender stereotyping (level 1) is the potential mediator, and covariance-adjusted changes over time on the student math self-concept/achievement variables (level 1) are the outcome of interest. As above, a multigroup approach would be utilized to examine if there are gender differences in the various paths of interest.

CHAPTER 3

RESULTS AND DISCUSSION

Descriptive Statistics

Descriptive statistics for all variables of interest are presented in Table 3, separately for both the overall sample and the 3rd grade subsample. Kurtosis and skew for all variables were within the acceptable range as outlined by Kline (2005), suggesting that the variables were approximately normally distributed. Correlations between all variables in the model are available in Table 4.

On average, students in the overall sample showed relatively high math self-concept at both waves, falling around 3.3 on a scale of 1-4 (Wave 1 M = 3.32, SD = 0.66; Wave 2 M = 3.35, SD = 0.59). Independent samples t-tests showed that there were no significant gender differences in math self-concept at either Wave 1, $t(2031) = .399$, $p = .36$, or Wave 2, $t(2184) = -1.34$, $p = .44$.

Table 3
Descriptive Statistics

	Full Sample					3rd Grade Cohort			
Variable	M	SD	Min	Max		M	SD	Min	Max
Teacher gender stereotyping	0.26	0.70	-1	5		0.19	0.61	-1	3
Student gender stereotyping	-	-	-	-		-0.28	1.70	-6	6
Wave 1 math achievement	0.056	0.90	-2.84	2.41		0.07	0.92	-2.66	2.19
Wave 2 math achievement	0.006	0.91	2.87	1.93		0.02	0.93	-2.49	1.78
Wave 1 math self-concept	3.32	0.66	1	4		3.35	0.54	1.2	4
Wave 2 math self-concept	3.35	0.59	1	4		3.28	0.60	1	4

Table 4
Correlations Between Study Variables

	Teacher gender stereotyping	Student gender stereotyping	Wave 1 math anxiety	Wave 2 math anxiety	Wave 1 math self-concept	Wave 2 math self-concept	SES	Age	Grade
Teacher gender stereotyping	1	.002	.02	-.04	-.10*	-.10*	.11*	-.02	-
Student gender stereotyping	-	1	.003	.01	-01	.06	.002	-.01	-
Wave 1 math anxiety	-.05*	-	1	.77**	.27**	.35**	.09	-.05	-
Wave 2 math anxiety	.01	-	.66**	1	.27**	.42**	.12*	-.12**	-
Wave 1 Math Self-Concept	-.01	-	.18**	.17**	1	.45**	.06	.05	-
Wave 2 Math Self-concept	-.05*	-	.22**	.27**	.33**	1	.07	-.01	-
SES	-.01	-	.14**	.13**	.01	.07**	1	-.05	-
Age	-.02	-	.03	-.01	.01	-.06**	-.06*	1	-
Grade	-.04	-	.001	-.003	.02	-.08**	-.03	.93**	1

Note. Values below the diagonal refer to the entire sample. Values above the diagonal refer to the 3rd grade subsample only.
**p* < .05; ***p* < .01.

Results were fairly similar when the 3rd grade cohort was examined separately. Again, math self-concept was relatively high at both waves (Wave 1 $M = 3.35$, $SD = 0.53$; Wave 2 $M = 3.28$, $SD = 0.60$) and there were no significant gender differences in math self-concept at either wave (*p*s > .67).

For the overall sample, the mean theta score for math achievement was 0.056 in Wave 1 ($SD = 0.90$), and 0.0056 at Wave 2 ($SD = 0.91$). At Wave 1, boys ($M = 0.11$, $SD = 0.87$) performed significantly better than did girls ($M = 0.01$, $SD = 0.92$; $t(1922) = -2.38$, $p = .02$). There were no significant gender differences in achievement at Wave 2 ($p = .51$). In the 3rd grade cohort, the mean score for achievement was 0.07 at Wave 1 ($SD = 0.92$), and 0.02 at Wave 2 ($SD = 0.93$). There were no significant gender differences in 3rd grade math achievement scores at either wave (*p*s > .10).

As detailed in the analysis plan, before analyzing the math-gender stereotyping scores, we examined these scores for both students and teachers to determine if there was sufficient variability to include these scores in analyses in their current form. The full student sample and 3rd grade subsample both showed math-gender stereotyping scores across the range of possible values, from -6 to 6, with negative values indicating female-biased stereotyping and positive values indicating male-biased stereotyping. In the teacher sample, math-gender stereotyping scores ranged from -1 to 5 in the overall sample, and -1 to 3 in the 3rd grade subsample. Based on these ranges, and the relatively normal distribution of the scores, it was determined that there was sufficient variability in scores to perform subsequent analyses with these variables as-is.

In the 3rd grade student subsample, the mean math-gender stereotyping score was small ($M = -0.28$, $SD = 1.70$). A one-sample t-test was conducted to examine if this mean was significantly different from zero, which would provide evidence for the existence of math-gender stereotyping in the sample. Results suggested the presence of math-gender stereotyping that favors females ($t(572) = -3.98$, $p < .001$, $d = .17$). However, the small effect size suggests that this difference may not be meaningful, and that, overall, what stereotyping does exist is minimal. When we repeated this analysis separately by gender, we found evidence of stereotyping scores significantly different from zero in both girls, $t(283) = -10.88$, $p < .001$, $d = .65$, and boys, $t(265) = 3.97$, $p < .001$, $d = .24$. Girls' stereotyping appeared to favor females ($M = -0.95$, $SD = 1.47$) while boys' stereotyping appeared to favor males ($M = 0.41$, $SD = 1.70$). The larger effect sizes in this case indicate that the opposite direction of these effects may have masked their magnitude in the overall sample. An independent samples t-test indicated that the difference in stereotyping scores between boys and girls was significant, $t(549) = -10.09$, $p <.001$, $d =.86$. Overall, this suggests that math-gender stereotyping is different in boys and girls, with students displaying small but significant amounts of math-gender stereotyping that biases their own gender.

In the full sample, teachers showed evidence of male-biased math-gender stereotyping ($M = 0.26$, $SD = 0.70$). A one-sample t-test indicated that mean math-gender stereotyping teachers is significantly different from zero, $t(222) = 5.46$, $p < .001$, $d = .37$. Results were similar when the 3rd-grade cohort of teachers were examined separately. They also showed a small male bias on average that was significantly different from zero ($M = 0.19$, $SD = 0.61$, $t(57) = 2.39$, $p = .02$, $d = .31$).

In the 3rd grade students, an independent samples t-test suggested that endorsement of stereotypes about girls and boys ($M = -0.26$, SD = 1.12) was significantly different from endorsement of stereotypes about women and men ($M = -0.02$, SD = 0.85, $t(573) = 5.38$, $p < .001$, $d = 0.23$), although the scores did show a significant, moderate-to-strong correlation, $r = .47$, $p < .001$. In the full sample of teachers, an independent samples t-test suggested that endorsement of stereotypes about girls and boys (*M* = 0.06, *SD* = 0.39) was significantly different from endorsement of stereotypes about women and men (*M* = 0.19, *SD* = 0.50, $t(222) = 3.53$, $p = .001$, $d = .24$) and the two types of stereotypes showed a small-to-moderate correlation, $r = .24$, $p < .001$. Results were similar for the 3rd grade subsample of teachers, $t(57) = 2.18$, $p = .03$, $d = .29$ with the two types of stereotypes showing a moderate correlation, $r = .36$, $p = .006$. As a result, the main analyses will be conducted using the combined adult and child item scores, but supplementary analyses will be conducted using them separately.

To address the three main research questions (1a, 1b, and 2), multi-level models were constructed using Mplus version 8.4. Data was considered to be missing at random (MAR), so full imputation maximum likelihood (FIML) was used to handle missing data, as it has been shown to produce unbiased parameter estimates and standard errors when data is MAR (Enders & Bandalos, 2001).

Research Question 1a: Do Teacher Math-Gender Stereotyping Scores at Wave 1 Relate to Student Self-Concept or Achievement at Wave 2, and are these Relationships Different for Boys and Girls?

Overall sample. Our first research question was to examine whether teacher gender stereotyping in fall predicts student math self-concept or math achievement in spring. Whilst the initial analysis plan envisioned examining both outcomes in a single model, due to the computational complexities of examining the difference between paths in boys and girls (discussed in further detail below), we made the decision to conduct two separate, multi-level models, with students (level 1) nested in classrooms (level 2) - one with teacher math-gender stereotyping predicting student math achievement, and another with teacher stereotyping predicting student math self-concept. In both models, the covariates were the same: age, SES, and students' Wave 1 scores in the particular outcome of interest. In the analysis plan, we had planned to include dummy-coded grade variables as covariates, to account for the difference in math tests between the grades. When we ran models with grade included, it was a significant predictor for math self-concept, but not math achievement. However, the addition of grade in the models caused some issues with model fit. Therefore, we re-ran the models without grade and the pattern of results was the same. In the current paper, we report the results of the models without grade as a covariate.

Our first model predicted Wave 2 math achievement from fall teacher math-gender stereotyping. Wave 1 math achievement, SES, and age in months were included as covariates at the student level. Model fit statistics for all math achievement models are available in Table 5. We used the following model fit indices to evaluate and compare each model: chi-square (χ^2) *p*-values greater than .05, root mean square error of approximation (RMSEA) values less than .08, and comparative fit index (CFI) values greater than or equal to .90 (Kline, 2015). The chi-square significance test for this model was significant, suggesting poor model fit, $\chi^2(12) = 98.56$, $p < .0001$.

Table 5

Goodness-Of-Fit Indicators for Multi-Level Models Predicting Wave 2 Math Achievement from Teacher Math-Gender Stereotyping

Sample	χ^2	*df*	*p*	CFI	TLI	RMSEA	*n*
K-3rd grade	98.56	12	<.0001	0.91	0.88	.05	2387
K-3rd (boys)	66.39	12	<.0001	0.88	0.84	.06	1138
K-3rd (girls)	71.62	12	<.0001	0.89	0.86	.07	1162
3rd grade only	43.91	12	<.0001	0.92	0.90	.06	642
3rd grade (boys)	31.79	12	.002	0.91	0.88	.07	300
3rd grade (girls)	40.33	12	.0001	0.87	0.83	.09	319

Note. Estimator used was standard maximum likelihood (ML).

However, it should be noted that, as chi-square values are related to sample size, in large samples such as this, the chi-square value will almost always be significant, irrespective of model fit (Bentler & Bonett, 1980). As a result, additional fit statistics were also examined. The RMSEA of .05 and CFI of 0.91 suggested an acceptable fit, although the TLI of 0.88 was slightly low.

When intraclass correlations (ICCs) were examined, classroom membership was shown to account for 22.2% of the variance in math achievement. Results of the model did not show a significant path between teacher math-gender stereotyping and student math achievement, *B* = 0.03, *SE* = 0.4, *p* = .51. Wave 1 math achievement, age, and SES were significant predictors of Wave 2 math achievement at the student level, such that students with higher prior math achievement, older students and students from higher SES backgrounds tended to have higher math achievement at Wave 2 (see Table 6 for path coefficients for all models involving the overall sample).

We fit an almost-identical model to predict math self-concept, with Wave 1 and Wave 2 math self-concept replacing Wave 1 and Wave 2 math achievement. Model fit indices for all models involving math self-concept are available in Table 7. The model fit for this model was acceptable, $\chi^2(12) = 26.41$, *p* = .01, RMSEA = .02, CFI = .94, TLI = 0.92.

Here, the ICCs showed that classroom membership accounted for about 4.6% of the variance in math self-concept. Results did not indicate a significant relation between teacher math-gender stereotyping and student math self-concept, *B* = -0.04, *SE* = 0.02, *p* = .10. Wave 1 math self-concept and SES were significant predictors of Wave 2 math self-concept at the within level, although age was not.

Table 6

Multilevel Models with Teacher Math-Gender Stereotyping Predicting Student Math Achievement and Math Self-Concept for the Overall Sample

Predictor	All students			Boys only			Girls only		
	B	S.E.	p	B	S.E.	p	B	S.E.	p
Wave 2 math achievement									
Teacher math-gender stereotyping	0.03	0.04	.51	0.03	0.05	.56	0.04	0.05	.34
Wave 1 math achievement	0.66***	0.02	<.001	0.66***	0.03	<.001	0.68***	0.03	<.001
Age	-.0.01**	0.003	.002	-0.01**	0.01	.004	-0.01	0.004	.16
SES	0.03*	0.01	.01	0.04*	0.02	.03	0.02	0.02	.19
Wave 2 math self-concept									
Teacher math-gender stereotyping	-0.04	0.02	.10	-0.03	0.03	.45	-0.03	0.03	.36
Wave 1 math self-concept	0.30***	0.02	<.001	0.26***	0.03	<.001	0.33***	0.04	<.001
Age	0.003	0.003	.26	-0.001	0.004	.75	0.01	0.004	.18
SES	0.03*	0.01	.02	0.02	0.02	.24	0.03*	0.01	.02

Note. Wave 1 math achievement, Wave 1 math self-concept, age, and SES were within level predictor variables; Teacher math-gender stereotyping was a between level variable.

*$p < .05$; ** $p < .01$; *** $p < .001$

Table 7
Goodness-Of-Fit Indicators for Multi-Level Models Predicting Wave 2 Math Self-Concept from Teacher Math-Gender Stereotyping

Sample	χ^2	*df*	*p*	CFI	TLI	RMSEA	*n*
K-3rd grade	26.41	12	.01	0.94	0.92	.02	2387
K-3rd (boys)	8.18	12	.77	1	1	0	1138
K-3rd (girls)	53.86	12	<.0001	0.79	0.72	.06	1162
3rd grade only	22.46	12	.03	0.93	0.90	.04	642
3rd grade (boys)	19.50	12	.08	0.84	0.80	.05	300
3rd grade (girls)	31.98	12	.001	0.80	0.73	.07	319

Note. Estimator used was standard maximum likelihood (ML).

As our hypothesis predicted that path estimates would be different for boys and girls, it is important to examine these findings separately by gender. In a combined sample, any potential effects that differ by gender could be washed out if the effects for girls and boys run in different directions, so we next examined the same models separately for boys and girls. The fit for the models predicting math achievement was acceptable in terms of RMSEA, although not CFI or TLI, for both boys, $\chi^2(12) = 66.39$, $p = <.0001$, RMSEA = .06, CFI = .88, TLI = .84, and girls, $\chi^2(12) = 71.62$, $p = <.0001$, RMSEA = .07, CFI = .89, TLI = .86. Classroom membership accounted for 22% of the variance in math achievement for boys, and 25% for girls. Results were consistent with the combined model, with no significant path estimates between teacher math-gender stereotyping and student math achievement for either boys, $B = 0.03$, $SE = 0.05$, $p = .56$, or girls, $B = 0.04$, $SE = 0.05$, $p = .34$. For math self-concept, the model fit well in boys, $\chi^2(12) = 8.18$, $p = .77$, RMSEA = 0, CFI = .1, TLI = 1, but less so in girls, $\chi^2(12) = 53.86$, $p = <.0001$, RMSEA = .06, CFI = .79, TLI = .72. Classroom membership accounted for 2.9% of the variance in self-concept in boys, and 9.2% in girls. Again, consistent with the combined model, there were no significant path estimates between teacher math-gender stereotyping and student math self-concept for either boys, *B* = -0.03, *SE* = 0.03, *p* = .45, or girls, *B* = -0.03, *SE* = 0.03, *p* = .36.

Initially, we had planned to compare the path estimates from teacher stereotyping to student math achievement and self-concept by gender using a multiple-group approach. However, the traditional multi-group approach is not suitable for multi-level models where the grouping variable (in this case, gender) is on the within (student) level, because it cannot account for the fact that boys and girls within the same classroom can have different means for variables at the within level (Asparouhov & Muthen, 2012). Thus, we need a model that can allow these cluster-specific gender effects which go beyond the overall population level gender effects. We used the approach specified by Asparouhov and Muthen (2012), wherein we create a latent class variable that gender is a perfect indicator for, rather than using gender as grouping variable. We then established a two-level mixture model, and introduce two correlated latent variables which each represent the random effect of one gender at the between level. This allows us to specify that certain paths can be different for the two groups.

We used a Wald test of parameter constraints to examine whether the path estimates for the relationships between teacher math-gender stereotyping and math achievement or math self-concept were different for girls and boys. The Wald tests revealed that there were no significant differences between girls and boys in the path estimates for the relationship between teacher math-gender stereotyping and either math achievement, $\chi^2(1) = 0.08$, $p = .78$, or math self-concept, $\chi^2(1) = 0.17$, $p = .68$.

When examined separately, results of models including stereotypes about adults or stereotypes about children separately were largely consistent with the above. One exception was the relationship between teacher math-gender stereotyping about adults and student achievement.

Both overall, and for girls and boys separately, teachers' male-biased math-gender stereotyping about adults was positively associated with student math achievement (*p*s <.03). Goodness-of-fit and path estimate statistics for these supplemental models are available in appendices A and B.

3rd grade subsample. Addressing research question 1b (below) required us to use student math-gender stereotyping data, which was only available for 3rd grade students, so we repeated the above analyses using just the 3rd grade cohort of students and teachers (*n* = 642 students, 58 teachers; see Table 8).

The results from model 1a, in combination with those from model 1b below, are required to establish if we have sufficient support for the mediation model proposed in research question 2. As such, it is important to ensure that we are combining results from the same sample in order to make any conclusions about the appropriateness of examining research question 2. The models for the 3rd grade cohort were the same as for the overall sample. Results of these models were similar to that of the model in the full sample.

The ICCs suggest that classroom membership accounts for only around 12.8% of the variance in math achievement for this subsample, which is lower than in the overall sample. It accounts for 4% of the variance in math self-concept, which is similar to the ICC value in the overall sample. Like for the overall sample, for the 3rd grade subsample there was no significant path estimate between teacher math-gender stereotyping and student math achievement, B = -.06, SE = .08, p =.47, or math self-concept, B = -.07, SE = .06, p = .25.

Table 8

Multilevel Models with Teacher Math-Gender Stereotyping Predicting Student Math Achievement and Math Self-Concept for the 3rd Grade Subsample

Predictor	All students			Boys only			Girls only		
	B	S.E.	p	B	S.E.	p	B	S.E.	p
Wave 2 math achievement									
Teacher math-gender stereotyping	-0.06	0.08	.47	-0.14	0.12	.23	-0.03	0.08	.71
Wave 1 math achievement	0.81***	0.03	<.001	0.84***	0.05	<.001	0.83***	0.05	<.001
Age	-0.01*	0.01	.02	-0.01	0.01	.41	-0.02*	0.01	.02
SES	0.05*	0.02	.02	0.05	0.03	.15	0.03	0.04	.38
Wave 2 math self-concept									
Teacher math-gender stereotyping	-0.07	0.06	.25	-0.13	0.07	.07	-0.02	0.10	0.87
Wave 1 math self-concept	0.48***	0.05	<.001	0.41***	0.08	<.001	0.53***	0.06	<.001
Age	-0.002	0.004	.69	-0.01	0.01	.25	0.001	0.01	.93
SES	0.03	0.02	.23	0.03	0.04	.47	0.02	.03	.50

Note. Wave 1 math achievement, Wave 1 math self-concept, age, and SES were within level predictor variables; Teacher math-gender stereotyping was a between level variable.

$*p < .05$; $** p < .01$; $*** p < .001$

Wave 1 math achievement predicted achievement at Wave 2, and self-concept showed the same pattern. Age and SES were significant predictors of math achievement, but not math self-concept. When examined separately, neither boys nor girls showed significant path estimates for teacher stereotyping predicting either math achievement (*p*s > .23) or math self-concept (*p*s > .07).

As we did for the full sample, we also conducted two-level mixture models in the 3rd grade subsample to calculate the path estimate between teacher math-gender stereotyping and math achievement or self-concept separately by gender, and compared them using a Wald test. Results were consistent with the overall sample – there was no significant difference between girls and boys in the path estimates between teacher math-gender stereotyping and math achievement, $\chi^2(1) = 0.001$, $p = .97$. or math self-concept, $\chi^2(1) = 0.13$, $p = .72$.

Once again, results of separate models for stereotypes about adults and children in 3rd grade were largely consistent with the above. However, in boys, teacher math-gender stereotyping about adults was negatively associated with math self-concept, such that as stereotyping increased, boys' math self-concept decreased. However, the *p*-value of this finding was .046, which is very close to the $p < .05$ threshold for determining significance, so it should be interpreted with caution. Goodness-of-fit and path estimate statistics for these supplemental models are available in Appendix A and B.

Research Question 1b: Do Teacher Math-Gender Stereotyping Scores at Wave 1 Relate to Student Math-Gender Stereotype Scores at Wave 2, and are these Relationships Different for Boys and Girls?

For our next research question, we constructed a multilevel model to examine the relationship between teacher math-gender stereotyping at Wave 1 and student math-gender stereotyping at Wave 2 in the 3rd grade students and their teachers. As in the models above, age and SES were included as covariates. Unlike the models above, Wave 1 gender stereotyping was not included as a covariate because we were more interested in absolute values of stereotyping rather than how it changed over time. The chi-square significance test suggested good model fit, χ^2 (6) = 9.00, p = .17, and the RMSEA of .03 was also indicative of a well-fitting model. However, the CLI and TFI were <.001. This may be reflecting the fact that the model we specified is close to the baseline model (i.e. the structural paths are so small that it is essentially equivalent to a model where they are not related at all). Since the baseline model has a significant chi-square statistic (χ^2 (9) = 9.36, p = .41) it is then unsurprising that our specified model does too.

Table 9

Goodness-Of-Fit Indicators for Multi-Level Models Predicting Wave 2 Student Math-Gender Stereotyping from Teacher Math-Gender Stereotyping

Sample	χ^2	*df*	*p*	CFI	TLI	RMSEA	*n*
3rd grade only	9.00	6	.17	<.001	<.001	.03	642
3rd grade (boys)	10.64	6	.10	<.001	<.001	.05	300
3rd grade (girls)	30.42	6	<.001	<.001	<.001	.11	319

Note. Estimator used was standard maximum likelihood (ML).

Goodness-of-fit indicators for all models predicting student math-gender stereotyping are available in Table 9. According to the ICCs, classroom membership accounted for 1.8% of the variance in student math-gender stereotyping. There was no significant path between teacher and student math-gender stereotyping, B = -0.004, SE = 0.01, p = .98. Age and SES were also not significant predictors of student math-gender stereotyping at the within level (ps > .85). Results for all models predicting student math-gender stereotyping are available in Table 10.

Table 10

Multilevel Models with Teacher Math-Gender Stereotyping Predicting Student Math-Gender Stereotyping for the 3rd Grade Subsample

Predictor	All students			Boys only			Girls only		
	B	S.E.	p	B	S.E.	p	B	S.E.	p
Teacher math-gender stereotyping	-0.004	0.13	.98	0.02	0.16	.91	0.10	0.16	.55
Age	-0.002	0.01	.89	-0.03	0.02	.08	0.01	0.02	.66
SES	0.01	0.05	.85	-0.01	0.10	.96	0.02	0.08	.83

Note. *A*ge and SES were within level predictor variables; Teacher math-gender stereotyping was a between level variable.

Next, we ran the same model separately for male and female students. Similar to the overall model, the model for males had an acceptable chi square statistic and RMSEA, $\chi^2(6) = 10.64$, $p = 0.10$, RMSEA = .05, but again CLI and TFI were <.001, suggesting that the variables in the model are very weakly correlated. ICCs suggested that classroom membership explained 1.6% of the variance in student math-gender stereotyping for boys, and 7% for girls.

The path estimate between teacher and student gender stereotyping was also non-significant in this model, *B* = 0.017, SE = 0.16, *p* = .91, and neither age nor SES were significant predictors, either (*p*s > .08). For female students, the model showed poor fit for all fit indices, $\chi^2(6) = 30.42$, $p < .0001$, RMSEA = .11, CFI < .001, TLI < .001. Once again, the path estimate from teacher to student math-gender stereotyping was non-significant, *B* = 0.10, *SE* = 0.16, *p* = .55, and neither age nor gender was a significant predictor, either (*p*s > .66).

Finally, consistent with the approach for research question 1a, we ran a multi-level mixture model with gender as a latent class to compare the path between teacher and student math-gender stereotyping in males and females. Results of a Wald test of parameter constraints used to compare the male and female path estimates was not significant, $\chi^2(1) = 0.25$, $p = .62$. This suggests that within the context of the model, the path estimate between teacher and student math-gender stereotyping is not different for boys and girls.

When we repeated these models using stereotyping scores for stereotyping about adults and stereotyping about children separately, results were largely consistent with the above except that, in boys, when math-gender stereotyping was measured with the adult items only, there was a significant relationship between teacher and student math-gender stereotyping, $B = 0.21$, $SE = 0.08$, $p = .01$, such that as teachers' male-biased stereotyping about men and women in math increased, so did similar stereotyping in their male, but not female students. Age was also a significant predictor of student math-gender stereotyping in this model, $B = 0.01$, $SE = 0.01$, $p = .01$, such that, in boys, as age increased, math-gender stereotyping about men and women was stronger. Goodness-of-fit and path estimate statistics for these supplemental models are available in Appendix C.

Research Question 2: Is the relationship between teacher math-gender stereotyping and student math self-concept or achievement mediated by student math-gender stereotyping, and are these relationships different for boys and girls?

Research question 2 was to examine if student gender stereotyping serves as a mediator of the relationship between teacher math-gender stereotyping and student math achievement or self-concept. Given the lack of significant path estimates between teacher math-gender stereotyping and student math achievement or math self-concept, and between teacher math-gender stereotyping and student math-gender stereotyping, we do not have sufficient evidence to support testing the mediation model outlined in research question 2b (Preacher et al., 2010). Therefore, as per our analysis plan, the analyses associated with research question 2 were not conducted.

Discussion

Understanding the factors that influence math achievement and math attitudes has been a topic of major interest to those searching for avenues to increase the participation of women in the STEM industry. Male-biased math-gender stereotypes have been proposed as one factor that may negatively impact girls' math self-concept and achievement, which in turn could reduce math interest and engagement, removing them from the 'pipeline' that leads to STEM-intensive careers. Given their role as facilitators of learning, we proposed that teachers may be uniquely placed to influence the development of students' academic attitudes through their own attitudes and beliefs. Thus, in the current study, we tested if teacher math-gender stereotyping at the beginning of a school year was associated with students' math achievement and self-concept at the end of that year, and if student math-gender stereotyping served as a mediator of any such relationships. Due to the male-biased content of traditional math-gender stereotypes, we also chose to examine if these relations were different for boys and girls. We hypothesized that higher levels of teacher math-gender stereotyping in the fall would be associated with lower math achievement and self-concept in the spring for girls but not for boys, and that these relationships would be mediated by students' own math-gender stereotypes.

Descriptive Research Findings

Whilst, in the overall sample, boys outperformed girls in math achievement at Wave 1, this gap had closed by Wave 2, and there were no gender differences in 3rd grade math achievement at any stage. These findings run contrary to those of Cimpian et al. (2016) who found evidence of a gender gap in math achievement in kindergarten through 2nd grade students from two cohorts of the ECLS-K, but they are consistent with results from Lindberg et al. 's (2010) metanalysis on the topic, which suggested that there are no meaningful differences in math achievement at the elementary school level. These results may be reflective of the narrowing of male mathematical performance advantages over time noted by Mullis et al. (2016) in their examination of data from the 2015 TIMSS cohort. Unlike in previous studies (e.g. Else-Quest et al., 2010; Ganley & Lubienski, 2016), which have consistently shown a gender gap in math self-concept favoring males, the female students in our sample did not have significantly lower math self-concept than did the male students. For girls and boys, math self-concept was relatively high on average in both the overall sample and the 3rd grade subsample.

Our results indicated that the 3rd grade students in the sample endorsed math-gender stereotypes, although the content of stereotypes was different for girls and boys, with girls endorsing female-biased stereotypes and boys endorsing male-biased stereotypes. This is inconsistent with the findings of Tomasetto et al. (2011) in Italian Kindergarten through 2nd grade students, where participants did not explicitly endorse math-gender stereotypes. However, their particular measure required participants to make a direct comparison between a male and female character in terms of which, if any, was better at math. The separate items used in this study for stereotypes about males and females may have provided for more nuanced answers.

Based on the findings of Kurtz-Costes et al., (2008), which suggested a shift towards more traditional male-biased stereotypes in both genders around age 9 or 10, we might have expected to see male-biased stereotypes in both girls and boys in this sample. Instead, the students displayed the kind of ingroup bias found by Heyman & Legare (2004) in their study of kindergarten and 1st grade students, with girls on average endorsing female-biased stereotypes and boys endorsing male-biased stereotypes, for both items about men and women, and items about boys and girls. This may suggest that math-gender stereotyping in children is changing over time, or that girls do not develop male-biased math-gender stereotypes until an older age than previously thought.

It should be noted that, whilst the stereotyping scores are significantly different from zero for both boys and girls, they are quite small overall. Given the explicit nature of the stereotyping measure, it is possible that the results reflect a social desirability bias, with students being reluctant to admit to endorsement of gender stereotypes. An implicit measure, such as that used by Cvencek et al. (2015), might provide different results in this regard.

There was also some evidence of male-biased math-gender stereotyping in teachers, consistent with Li’s (1999) review of teachers’ math-gender beliefs, which indicated that teachers did show stereotyped, male-biased beliefs about math and gender across a range of studies. Once again, although mean stereotyping scores were significantly different from zero, they were quite low. This suggests the possibility that, as was suggested for students, social desirability bias may have made teachers reluctant to share their true beliefs, thus masking the true extent of any gender stereotyping. This is supported by Nürnberger et al.'s (2016) study of preservice teachers which showed that explicit and implicit gender-ability stereotypes were not highly correlated, so what teachers say they believe and what they actually believe may be quite different. Given one major proposed mechanism of action for the influence of teachers’ stereotypes on their students is through their influence on teacher behavior (e.g. Becker, 1981; Simpson & Linder, 2016) it is important to note that stereotypes may still play a role in the way teachers act, and thus influence their students, without them being explicitly expressed.

Is Teacher Math-Gender Stereotyping Associated with Student Math Achievement or Self-Concept?

Our models to address research question 1a – whether teacher math-gender stereotyping is associated student math achievement or math self-concept – did not find a significant association between teacher math-gender stereotyping and either student math achievement or student math self-concept. The models showed very weak relations for both outcomes in the overall sample, as well as in the 3rd grade subsample.

Given our hypothesis that these relationships would be significant and negative for girls, and non-significant for boys, we considered the possibility that this could cause associations to be obscured in the models if we examined models with boys and girls together. Thus, we next examined the relations in girls and boys separately. As expected, when we repeated the models divided by gender, there was no relation for male students between teacher math-gender stereotyping and math achievement, or between teach stereotyping and math self-concept. However, contrary to our hypothesis, the relations were also nonsignificant for girls. Whilst there is a dearth of literature relating teacher stereotyping directly to student math outcomes, we had expected, based on findings that male-biased math-gender stereotyping in girls is associated with lower math achievement (Beilock et al., 2010), and self-concept (Passolunghi et al., 2014), that the girls in this sample might show lower math achievement as their teachers showed higher male-biased math-gender stereotyping. However, this was not the case, and girls seemed relatively unaffected by their teachers' stereotyping.

Interestingly, when these models were repeated using teacher's math-gender stereotypes about men and women only as a predictor, we found a significant relation with math achievement in the overall sample, and for girls and boys separately, such that stronger male-biased stereotyping in teachers was associated with higher math achievement in their male and female students. This finding was unexpected based on the results from the initial models, and the results were opposite to our hypothesis. We had not predicted that teacher stereotyping would benefit boys, and we expected teachers' male-biased stereotypes to be associated with worse, not better, performance in girls (e.g. Eccles et al., 1990). This suggests that, consistent with Powlishta (2000) and Steele (2003), there is something different about teachers' stereotypes about adults versus those about children in terms of how they influence their students' outcomes, although the exact mechanism is unclear. It is possible that because stereotypes about adults relate more to how children think they should be in the future rather than how they are now, they inspire both genders to work hard at math – boys to live up to the stereotype and girls to prove the stereotype wrong - thus improving math performance. Further research is required to understand the nature of this unusual finding, as there is a dearth of work exploring how different subjects of math-gender stereotyping might influence teachers' behaviors differently.

We also found that teacher math-gender stereotypes about adults negatively predicted boys' math self-concept in the 3rd grade subsample, such that teachers holding stronger male-biased stereotypes about math and gender was associated with lower math self-concept in their male 3rd grade students. This, too, was an unexpected finding, as math-gender stereotyping is generally associated with higher math self-concept in boys (Passolunghi et al., 2014), again reinforcing the idea that there may be something different about teachers' stereotypes about adults specifically. This is especially interesting given there was no such negative relation for girls, who we might have expected to be negatively affected. However, this finding only just reached statistical significance, and given the relation was not significantly different from the non-significant relation for 3rd grade girls it may be a spurious result and should be interpreted with caution. Further research is required to ascertain if this finding is robust. It may be that, as boys get older, they feel pressure to live up to the stereotypes about their potential adult abilities which causes them to reflect negatively on their current abilities, whereas girls are not subjected to similar pressures, but further research is required to tease apart this unusual finding.

To examine the final part of this research question – whether the associations between teacher math-gender stereotyping and student math achievement, or between teacher math-gender stereotyping and student math self-concept, were significantly different for male and female students - we compared the relations in girls and boys, and found that, contrary to our hypothesis, they were not significantly different from one another, even for the relation between teacher stereotyping and self-concept in 3rd grade students, where the association was significant for boys but not for girls. This suggests that the association between teacher stereotypes and student math achievement and self-concept is similar for boys and girls.

Is Teacher Math-Gender Stereotyping Associated with Student Math-Gender Stereotyping?

Our models used to address research question 1b – whether teacher math-gender stereotyping predicts student math-gender stereotyping – did not support the hypothesis that stereotyping in teachers would be associated with stereotyping in students, with one unexpected exception. Contrary to the findings of Keller (2001), which showed that increased math-gender stereotyping in teachers was associated with increased stereotyping in their students, we did not find a significant relation between teacher and student math-gender stereotyping in our 3rd grade subsample, which suggests that teachers' math-gender stereotyping is not related to their student's math gender stereotyping at the end of the school year. This finding was consistent when girls and boys were examined separately, and the associations for girls and boys were not significantly different from one another. It may be that students' endorsement of math-gender stereotypes develops over a longer period of time, meaning that the duration of a school year is simply too short to see any influence on the formation of these ideas. However, when we examined stereotypes about adults only, we found that teacher gender stereotyping predicted boys' math-gender stereotyping, but not girls, such that as teachers' male-biased stereotyping about adults increases, boys' male-biased stereotyping about adults also increases. This was an interesting finding which does not necessarily fit with our expected results, as we had expected that our mostly-female teacher sample would influence the stereotyping of girls, but not boys, due to children's tendency to emulate the views of same-gender adults (Gunderson et al., 2012). This finding once again suggests that there is something different about stereotypes about adults for 3rd grade boys, and not girls, but the exact mechanisms are unclear. A developmental approach is required to better understand these effects. It is also likely that other social influencers, such as parents or peers, could play a larger role in the development of these gendered attitudes (Tiedemann, 2000).

In light of the results from these first two models, no further models were required to address our 3rd research question, of whether student math-gender stereotyping mediates a relationship between teacher math-gender stereotyping and student math self-concept and achievement. Given the lack of significant relations in the models above, we can conclude that there is no significant relationship between teacher math-gender stereotyping and student math self-concept, achievement, or gender stereotyping. As a result, it is not possible to say that teacher gender stereotyping predicts student math achievement and self-concept through students' gender stereotyping, and establishing if such a model would be different in boys and girls is unnecessary.

Limitations and Future Directions

A number of limitations should be taken into account when interpreting these results. The current study follows students across a single school year, and, in addition, the gender stereotyping data was available for the 3rd grade cohort only. Male-biased math-gender stereotypes are thought to be more prevalent in older children, so it is possible that the current sample did not include the age groups where we would be most likely to see relations. Having stereotyping data from younger students in the sample could have provided cross-sectional data on how gender stereotyping develops over time. In addition, longitudinal data collected over multiple school years with the same group of students could shed more light on how gender stereotypes change as students grow.

In this study, we utilized explicit measures to explore math-gender stereotyping, but it is possible that these could have masked the true extent of stereotyping in both teachers and students. This, in turn, could impact the relations between the major variables in this study, as attitudes do not need to be explicitly expressed to impact behavior. Future research should incorporate implicit stereotyping measures to examine if the attitudes being portrayed are an accurate reflection of how participants truly think about math and gender, in order to gain a fuller understanding of the influences of stereotyping on students' outcomes.

The inconsistent findings when stereotypes about men and women were considered separately from those about boys and girls merit further exploration. It may be that children relate differently to stereotypes about fellow children differently than to stereotypes about adults, which has the potential to impact how such stereotypes affect them. Future research should further explore how children relate math-gender stereotypes to themselves, depending on the subject of the stereotype. How different stereotype subjects influence teacher behavior should also be explored. Teachers may be conscious of avoiding behavior that reinforces stereotypes about children, but less vigilant in avoiding those about adults. The unexpected finding that male-biased math-gender stereotypes about adults is actually associated with reduced math self-concept in boys is particularly concerning and merits further investigation to understand what underlies it.

Conclusion

Taken together, these findings suggest that 1) explicit math-gender stereotyping is evident, in both teachers and students at the 3rd grade level, albeit at low levels, with students' stereotypes showing own-gender bias 2) math-gender stereotyping in teachers is generally not associated with their early elementary school students' math achievement or math self-concept over a school year, even when boys and girls are examined separately, although 3) teachers' male-biased stereotypes about adults specifically positively influence overall performance in boys and girls, but negatively influence self-concept in 3rd grade boys, 4) teachers' math-gender stereotyping is not associated with their students' math-gender stereotyping, except for boys when we examine stereotypes about adults specifically and therefore 5) there is no support for the role of student math-gender stereotyping as a mediator in the relationships between teacher math-gender stereotyping and either student math achievement or student math self- concept.

Overall, these results suggest that risk of math-gender stereotyping and its potential impacts on students' outcomes may not be a topic of major concern for early elementary school students and their teachers. These results suggest that reducing teachers' math-gender stereotyping may not be the most useful candidate for intervention if we want to improve math performance or math self-concept in female students. It is possible that decreasing male-biased math-gender stereotyping could actually positively impact boys' self-concept in older elementary school, but further research is needed to establish this. Given the already low levels of math-gender stereotyping in teachers and students, there is no evidence to suggest that reducing these even further would provide any appreciable benefit for girls' math performance or self-concept in students this age group, and exploring child-centered approaches such as interventions aimed at increasing confidence or providing additional math support may provide greater benefits.

These findings are important because they provide us with more information about one potential avenue for interventions to improve math achievement and self-concept in girls. Given the importance of improving these outcomes, and the time and expense involved in the development of interventions, it is important to determine which factors are potential targets for producing meaningful change, allowing us to focus our efforts on those which can have the most profound impact and ultimately improve the representation for women in STEM. There is little research exploring how teacher gender stereotypes relate to student outcomes, so the current study provides a key contribution to the foundational literature surrounding these important questions. In this case, whilst math-gender stereotyping overall did not emerge as a strong candidate to target in terms of increasing girls' math self-concept and achievement at this age, math-gender stereotypes about adults did emerge as a candidate worthy of future research in this area. In addition, the results provide a strong rationale for further, longitudinal research to understand how these relations may change over time, which is currently lacking in the evidence base. Ultimately, this study serves as a jumping-off point for future developmental research exploring how stereotyping in teachers might influence their students' outcomes over time, which can inform our efforts in improving girls' math outcomes and, hopefully, increase the participation of women in STEM fields.

APPENDIX A

RESULTS FROM SUPPLEMENTARY MODELS FOR RESEARCH QUESTION 1A WITH MATH-GENDER STEREOTYPING ITEMS ABOUT ADULTS

Table 11

Goodness-Of-Fit Indicators for Multi-Level Models Predicting Wave 2 Math Achievement from Teacher Math-Gender Stereotyping About Adults

Sample	χ^2	*df*	*p*	CFI	TLI	RMSEA	*n*
K-3rd grade	101.97	12	<.0001	0.91	0.88	.06	2387
K-3rd (boys)	70.78	12	<.0001	0.87	0.83	.07	1138
K-3rd (girls)	68.38	12	<.0001	0.90	0.86	.06	1162
3rd grade only	48.81	12	<.0001	0.92	0.89	.07	642
3rd grade (boys)	32.11	12	.001	0.91	0.88	.08	300
3rd grade (girls)	45.26	12	<.0001	0.86	0.81	.09	319

Note. Estimator used was standard maximum likelihood (ML).

Table 12

Goodness-Of-Fit Indicators for Multi-Level Models Predicting Wave 2 Math Self-Concept from Teacher Math-Gender Stereotyping About Adults

Sample	χ^2	*df*	*p*	CFI	TLI	RMSEA	*n*
K-3rd grade	24.53	12	.02	0.95	0.93	.02	2387
K-3rd (boys)	8.15	12	.77	1	1	0	1138
K-3rd (girls)	49.39	12	<.0001	0.80	0.74	.05	1162
3rd grade only	20.44	12	.06	0.94	0.92	.03	642
3rd grade (boys)	19.95	12	.07	0.84	0.78	.05	300
3rd grade (girls)	27.18	12	.01	.85	.80	.06	319

Note. Estimator used was standard maximum likelihood (ML).

Table 13

Multilevel Models with Teacher Math-Gender Stereotyping about Adults Predicting Student Math Achievement and Math Self-Concept for the Overall Sample

Predictor	All students			Boys only			Girls only		
	B	S.E.	p	B	S.E.	p	B	S.E.	p
Wave 2 math achievement									
Teacher math-gender stereotyping	0.13*	0.06	.02	0.15*	0.07	.03	0.14*	0.06	.03
Wave 1 math achievement	0.66***	0.02	<.001	0.66***	0.03	<.001	0.68***	0.03	<.001
Age	-0.01**	0.003	.002	-0.01**	0.01	.004	-0.01	0.004	.16
SES	0.03*	0.01	.01*	0.04*	0.02	.02	0.02	0.02	.19
Wave 2 math self-concept									
Teacher math-gender stereotyping	-0.04	0.03	.29	-0.02	0.05	.79	-0.02	0.04	.56
Wave 1 math self-concept	0.30***	0.02	<.001	0.26***	0.03	<.001	0.33***	0.03	<.001
Age	0.003	0.003	.26	-0.001	0.004	.75	0.01	0.004	.18
SES	0.03*	0.01	.02	0.02	0.02	.24	0.03*	0.01	.02

Note. Wave 1 math achievement, Wave 1 math self-concept, age, and SES were within level predictor variables; Teacher math-gender stereotyping was a between level variable.

*$p < .05$; ** $p < .01$; *** $p < .001$

Table 14

Multilevel Models with Teacher Math-Gender Stereotyping About Adults Predicting Student Math Achievement and Math Self-Concept for the 3rd Grade Subsample

Predictor	All students			Boys only			Girls only		
	B	S.E.	p	B	S.E.	p	B	S.E.	p
Wave 2 math achievement									
Teacher math-gender stereotyping	-0.03	0.10	.77	-0.004	0.13	.98	-0.02	0.10	.85
Wave 1 math achievement	0.81***	0.03	<.001	0.84***	0.05	<.001	0.83***	0.05	<.001
Age	-0.01*	0.02	.02	-0.01	0.01	.41	-0.02*	0.01	.02
SES	0.05*	0.02	.03	0.04	0.03	.21	0.03	0.04	.39
Wave 2 math self-concept									
Teacher math-gender stereotyping	-0.08	0.06	.18	-0.16*	0.08	.046	-0.004	0.10	.97
Wave 1 math self-concept	0.49***	0.05	<.001	0.41***	0.08	<.001	0.53***	0.06	<.001
Age	-.002	0.004	.69	-0.01	0.01	.25	0.001	0.01	.93
SES	0.03	0.02	.24	0.03	0.04	.48	0.02	0.03	.50

Note. Wave 1 math achievement, Wave 1 math self-concept, age, and SES were within level predictor variables; Teacher math-gender stereotyping was a between level variable.

*$p < .05$; ** $p < .01$; *** $p < .001$

APPENDIX B

RESULTS FROM SUPPLEMENTARY MODELS FOR RESEARCH QUESTION 1A WITH MATH-GENDER STEREOTYPING ITEMS ABOUT CHILDREN

Table 15

Goodness-Of-Fit Indicators for Multi-Level Models Predicting Wave 2 Math Achievement from Teacher Math-Gender Stereotyping About Children

Sample	χ^2	*df*	*p*	CFI	TLI	RMSEA	*n*
K-3rd grade	95.76	12	<.0001	0.91	0.88	.05	2387
K-3rd (boys)	63.52	12	<.0001	0.88	0.84	.06	1138
K-3rd (girls)	70.88	12	<.0001	0.89	0.85	.07	1162
3rd grade only	40.04	12	.0001	0.93	0.91	.06	642
3rd grade (boys)	30.62	12	.002	0.92	0.89	.07	300
3rd grade (girls)	32.12	12	.001	0.90	0.86	.07	319

Note. Estimator used was standard maximum likelihood (ML).

Table 16

Goodness-Of-Fit Indicators for Multi-Level Models Predicting Wave 2 Math Self-Concept from Teacher Math-Gender Stereotyping About Children

Sample	χ^2	*df*	*p*	CFI	TLI	RMSEA	*n*
K-3rd grade	27.17	12	.01	0.94	0.92	.02	2387
K-3rd (boys)	8.63	12	.73	1	1	0	1138
K-3rd (girls)	53.79	12	<.0001	0.79	0.72	.06	1162
3rd grade only	14.73	12	.26	0.98	0.97	.02	642
3rd grade (boys)	14.16	12	.29	0.94	0.93	.02	300
3rd grade (girls)	24.15	12	.02	0.86	0.82	.06	319

Note. Estimator used was standard maximum likelihood (ML).

Table 17

Multilevel Models with Teacher Math-Gender Stereotyping About Children Predicting Student Math Achievement and Math Self-Concept for the Overall Sample

Predictor	All students			Boys only			Girls only		
	B	S.E.	p	B	S.E.	p	B	S.E.	p
Wave 2 math achievement									
Teacher math-gender stereotyping	-0.12	0.09	.19	-0.14	0.10	.17	-0.10	0.11	.35
Wave 1 math achievement	0.66***	0.02	<.001	0.66***	0.03	<.001	0.68***	0.03	<.001
Age	-0.01**	0.003	.002	-0.01**	0.01	.004	-0.01	0.004	.17
SES	0.03*	0.01	.01	0.04*	0.02	.02	0.03	0.02	.18
Wave 2 math self-concept									
Teacher math-gender stereotyping	-0.07	0.04	.09	-0.06	0.06	.29	-0.05	0.05	.30
Wave 1 math self-concept	0.30***	0.02	<.001	0.26***	0.03	<.001	0.33***	0.04	<.001
Age	0.003	0.003	.27	-0.001	0.004	.75	0.01	0.004	.18
SES	0.03*	0.01	.02	0.02	0.02	.23	0.03*	0.01	.02

Note. Wave 1 math achievement, Wave 1 math self-concept, age, and SES were within level predictor variables; Teacher math-gender stereotyping was a between level variable.

p* < .05; ** *p* < .01; * *p* < .001

Table 18

Multilevel Models with Teacher Math-Gender Stereotyping About Children Predicting Student Math Achievement and Math Self-Concept for the 3rd Grade Subsample

Predictor	All students			Boys only			Girls only		
	B	S.E.	p	B	S.E.	p	B	S.E.	p
Wave 2 math achievement									
Teacher math-gender stereotyping	-0.15	.16	0.34	-0.31	0.17	0.07	-0.08	0.18	.65
Wave 1 math achievement	0.81***	0.03	<.001	0.84***	0.05	<.001	0.83	0.05	<.001
Age	-0.01*	0.01	.02	-0.01	0.01	.44	-0.02*	0.01	.02
SES	0.05*	0.02	.02	0.05	0.03	.14	0.03	0.04	.39
Wave 2 math self-concept									
Teacher math-gender stereotyping	-0.09	0.13	.50	-0.13	0.16	.43	-0.06	0.22	.81
Wave 1 math self-concept	0.49***	0.05	<.001	0.41***	0.08	<.001	0.53***	0.06	<.001
Age	-0.002	0.004	.70	-0.01	0.01	.26	0.001	0.01	.93
SES	0.03	0.02	.25	0.02	0.04	.52	0.02	0.03	.50

Note. Wave 1 math achievement, Wave 1 math self-concept, age, and SES were within level predictor variables; Teacher math-gender stereotyping was a between level variable.

*p < .05; ** p < .01; *** p < .001

APPENDIX C

RESULTS FROM SUPPLEMENTARY MODELS FOR RESEARCH QUESTION 1B WITH MATH-GENDER STEREOTYPING ITEMS ABOUT ADULTS AND CHILDREN

Table 19

Goodness-Of-Fit Indicators for Multi-Level Models Predicting Wave 2 Student Math-Gender Stereotyping About Adults from Teacher Math-Gender Stereotyping About Adults

Sample	χ^2	*df*	*p*	CFI	TLI	RMSEA	*n*
3rd grade only	10.45	6	.11	<.001	<.001	.03	642
3rd grade (boys)	10.88	6	.09	.44	.16	.05	300
3rd grade (girls)	20.26	6	.003	<.001	<.001	.09	319

Note. Estimator used was standard maximum likelihood (ML).

Table 20

Goodness-Of-Fit Indicators for Multi-Level Models Predicting Wave 2 Student Math-Gender Stereotyping About Children from Teacher Math-Gender Stereotyping About Children

Sample	χ^2	*df*	*p*	CFI	TLI	RMSEA	*n*
3rd grade only	4.64	6	.59	1	1	0	642
3rd grade (boys)	5.8	6	.44	1	1	0	300
3rd grade (girls)	17.63	6	.01	<.001	<.001	.08	319

Note. Estimator used was standard maximum likelihood (ML).

Table 21

Multilevel Models with Teacher Math-Gender Stereotyping About Adults Predicting Student Math-Gender Stereotyping About Adults for the 3rd Grade Subsample

predictor	all students			boys only			girls only		
	B	S.E.	p	B	S.E.	p	B	S.E.	p
Teacher math-gender stereotyping	0.02	0.11	.87	0.21*	0.08	.01	-0.03	0.13	.83
Age	-0.01	0.01	.06	-0.02*	0.01	.01	-0.01	0.01	.54
SES	-0.01	0.02	.57	-0.02	0.05	.68	-0.01	0.04	.72

*Note. A*ge and SES were within level predictor variables; Teacher math-gender stereotyping was a between level variable.
p* < .05; ** *p* < .01; * *p* < .001

Table 22

Multilevel Models with Teacher Math-Gender Stereotyping About Children Predicting Student Math-Gender Stereotyping about Children for the 3rd Grade Subsample

predictor	all students			boys		only	girls only		
	B	S.E.	p	B	S.E.	p	B	S.E.	p
Teacher math-gender stereotyping	-0.07	.09	.41	-0.16	0.12	.17	-0.01	0.19	.98
Age	0.01	0.01	.45	-0.01	0.01	.47	0.01	0.01	.35
SES	0.02	0.04	.55	0.02	0.07	.83	0.04	0.05	.45

*Note. A*ge and SES were within level predictor variables; Teacher math-gender stereotyping was a between level variable.

APPENDIX D

IRB EXEMPTION

FLORIDA STATE UNIVERSITY
OFFICE *of the* VICE PRESIDENT *for* RESEARCH

NOT HUMAN RESEARCH

December 10, 2019

Rachel Conlon
[redacted]

Dear Rachel Conlon:

On 12/9/2019, the IRB staff reviewed the following submission:

Title of Study:	Examining the Relationship Between Teacher Math-Gender Stereotypes and Students' Math Outcomes
Investigator:	Rachel Conlon
Submission ID:	STUDY00000565
Study ID:	STUDY00000565
Funding:	None
IND, IDE, or HDE:	None
Documents Reviewed:	• Conlon_Proposal Document_10.11.19, Category: IRB Protocol;

The IRB staff determined that the proposed activity is not research involving human subjects as defined by DHHS and/or FDA regulations.

IRB review and approval by this organization is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these activities are research involving human subjects in which the organization is engaged, please submit a new request to the IRB for a determination. You can create a modification by clicking **Create Modification / CR** within the study.

Sincerely,

Human Subjects Research Office
humansubjects@fsu.edu

Page 1 of 1

REFERENCES

Asparouhov, T., & Muthen, B. (2012). *Multiple Group Multilevel Analysis: Mplus Web Notes: No. 16*.

Barth, J. M., Kim, H., Eno, C. A., & Guadagno, R. E. (2018). Matching abilities to careers for others and self: Do gender stereotypes matter to students in advanced math and science classes? *Sex Roles*, *79*(1), 83–97. https://doi.org/10.1007/s11199-017-0857-5

Becker, J. R. (1981). Differential treatment of females and males in mathematics classes. *Journal for Research in Mathematics Education*, *12*(1), 40–53. https://doi.org/10.2307/748657

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., & Doms, M. E. (2011). *Women in STEM: A gender gap to innovation* (No. 04–11). US Department of Commerce. http://www.ssrn.com/abstract=1964782

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *PNAS*, *107*(5), 1860–1863. https://doi.org/10.1073/pnas.0910967107

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Bian, L., Leslie, S.-J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389–391. https://doi.org/10.1126/science.aah6524

Cimpian, J. R., Lubienski, S. T., Timmer, J. D., Makowski, M. B., & Miller, E. K. (2016). Have gender gaps in math closed? Achievement, teacher perceptions, and learning behaviors across two ECLS-K cohorts. *AERA Open*, *2*(4), 1–19. https://doi.org/10.1177/2332858416673617

Cvencek, D., Kapur, M., & Meltzoff, A. N. (2015). Math achievement, stereotypes, and math self-concepts among elementary-school students in Singapore. *Learning and Instruction*, *39*, 1–10. https://doi.org/10.1016/j.learninstruc.2015.04.002

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Development*, *82*(3), 766–779. https://doi.org/10.1111/j.1467-8624.2010.01529.x

Dickhäuser, O., & Meyer, W.-U. (2006). Gender differences in young children's math ability attributions. *Psychology Science*, *48*(1), 3–16.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103–127. https://doi.org/10.1037/a0018053

Enders, C. K., & Bandalos, D. L. (2001). The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, *21*(1), 55–69. https://doi.org/10.1007/BF00311015

Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology*, *38*(4), 519–533. https://doi.org/10.1037/0012-1649.38.4.519

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences*, *47*, 182–193. https://doi.org/10.1016/j.lindif.2016.01.002

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700–717. https://doi.org/10.1037/a0026659

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*, 153–166. https://doi.org/10.1007/s11199-011-9996-2

Harrop, A., & Swinson, J. (2011). Comparison of teacher talk directed to boys and girls and its relationship to their behaviour in secondary and primary schools. *Educational Studies*, *37*(1), 115–125. https://doi.org/10.1080/03055691003729260

Heyman, G. D., & Legare, C. H. (2004). Children's beliefs about gender differences in the academic and social domains. *Sex Roles*, *50*(3), 227–239. https://doi.org/10.1023/B:SERS.0000015554.12336.30

Jacobs, J. E. (1991). Influence of gender stereotypes on parent and child mathematics attitudes. *Journal of Educational Psychology*, *83*(4), 518–527. https://doi.org/10.1037/0022-0663.83.4.518

Keller, C. (2001). Effect of teachers' stereotyping on students' stereotyping of mathematics as a male domain. *The Journal of Social Psychology*, *141*(2), 165–173. https://doi.org/10.1080/00224540109600544

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. The Gilford Press.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking—Methods and practices*. Springer-Verlag. https://www.springer.com/us/book/9781475743104

Kurtz-Costes, B., Rowley, S. J., Harris-Britt, A., & Woods, T. A. (2008). Gender stereotypes about mathematics and science and self-perceptions of ability in late childhood and early adolescence. *Merrill-Palmer Quarterly*, *54*(3), 386–409. http://dx.doi.org/10.1353/mpq.0.0001

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research*, *41*(1), 63–76. https://doi.org/10.1080/0013188990410106

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, *136*(6), 1123–1135. https://doi.org/10.1037/a0021276

Lummis, M., & Stevenson, H. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology*, *26*(2), 254–263. https://doi.org/10.1037/0012-1649.26.2.254

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. https://doi.org/10.1111/j.1467-8624.2005.00853.x

Master, A., & Meltzoff, A. N. (2016). Building bridges between psychological science and education: Cultural stereotypes, STEM, and equity. *PROSPECTS*, *46*(2), 215–234. https://doi.org/10.1007/s11125-017-9391-z

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *Student Achievement – TIMSS 2015 and TIMSS Advanced 2015 International Results*.

http://timssandpirls.bc.edu/timss2015/international-results/timss-2015/mathematics/student-achievement/

National Center for Education Statistics. (2020, May). *The condition of education—Preprimary, elementary, and secondary education—Teachers and staff—Characteristics of public school teachers—Indicator April (2018)*. National Center for Education Statistics. https://nces.ed.gov/programs/coe/indicator_clr.asp

Noonan, R. (2017). *Women in STEM: 2017 Update* (No. 06–17). US Department of Commerce. https://www.census.gov/programs-surveys/acs/.

Nosek, B. A., Smith, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., … Greenwald, A. G. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *PNAS*, *106*(26), 10593–10597. http://dx.doi.org/10.1073/pnas.0809921106

Nürnberger, M., Nerb, J., Schmitz, F., Keller, J., & Sütterlin, S. (2016). Implicit Gender Stereotypes and Essentialist Beliefs Predict Preservice Teachers' Tracking Recommendations. *The Journal of Experimental Education*, *84*(1), 152–174. https://doi.org/10.1080/00220973.2015.1027807

Passolunghi, M. C., Rueda Ferreira, T. I., & Tomasetto, C. (2014). Math-gender stereotypes and math-related beliefs in childhood and early adolescence. *Learning and Individual Differences*, *34*, 70–76. https://doi.org/10.1016/j.lindif.2014.05.005

Powlishta, K. K. (2000). The Effect of Target Age on the Activation of Gender Stereotypes. *Sex Roles*, *42*(3), 271–282. https://doi.org/10.1023/A:1007095307977

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*(3), 209–233. https://doi.org/10.1037/a0020141

Schoen, R. C., Anderson, D., & Bauduin, C. (2018). *Elementary Mathematics Student Assessment: Measuring grade 3, 4, and 5 Students' Performance in Number (Whole Numbers and Fractions), Operations, and Algebraic Thinking in Spring 2016.* (No. 2018–23). Florida State University.

Schoen, R. C., Anderson, D., Riddell, C. M., & Bauduin, C. (2018). *Elementary Mathematics Student Assessment: Measuring the Performance of Grade 3, 4, and 5 Students in*

Number (Whole Numbers and Fractions), Operations, and Algebraic Thinking in Fall 2015. (No. 2018–24). Florida State University.

Sells, L. W. (1980). Mathematics: The invisible filter. *Engineering Education*, *70*(4), 340–341.

Seo, E., Shen, Y., & Alfaro, E. C. (2018). Adolescents' beliefs about math ability and their relations to STEM career attainment: Joint consideration of race/ethnicity and gender. *Journal of Youth and Adolescence*, *48*(2), 1–20. https://doi.org/10.1007/s10964-018-0911-9

Simpson, A., & Linder, S. M. (2016). The indirect effect of children's gender on early childhood educators' mathematical talk. *Teaching and Teacher Education*, *54*, 44–53. https://doi.org/10.1016/j.tate.2015.11.011

Steele, J. (2003). Children's gender stereotypes about math: The role of stereotype stratification. *Journal of Applied Social Psychology*, *33*(12), 2587–2606. https://doi.org/10.1111/j.1559-1816.2003.tb02782.x

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology*, *102*(4), 947–963. https://doi.org/10.1037/a0019920

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, *92*(1), 144–151. https://doi.org/10.1037/0022-0663.92.1.144

Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology*, *47*(4), 943. https://doi.org/10.1037/a0024047

Xue, Y., & Larson, R. (2015). *STEM crisis or STEM surplus? Yes and yes: monthly labor review: U.S. Bureau of Labor Statistics*. Bureau of Labor and Statistics. http://www.bls.gov/opub/mlr/2015/article/stem-crisis-or-stem-surplus-yes-and-yes.htm#_edn32#_edn32

BIOGRAPHICAL SKETCH

Rachel Conlon

Florida State University

Education

PhD Developmental Psychology **Fall 2018-Present**

Florida State University, Tallahassee, FL

Advisor: Dr. Colleen Ganley

Expected Graduation Date: June 2023

BA Psychology **Fall 2011-Spring 2015**

Trinity College Dublin, Ireland

Publications

Ganley, C. M., **Conlon, R. A,** McGraw, A. L., Barroso, C., & Geer, E. A. (in press). The effectiveness of brief anxiety interventions on math test performance. *Journal of Numerical Cognition.*

Conlon, R.A., Hicks, A., Barroso, C. & Ganley, C.M. (in press). The effect of the timing of math anxiety measurement on math outcomes**.** *Learning and Individual Differences.*

In Preparation

Conlon, R. A., Barroso, C., & Ganley, C. M. (in prep). Exploring early elementary school children's aspirations for STEM careers and understanding of required STEM skills.

Conlon, R.A., Ganley, C.M., & Schatschneider, C. (in prep). Examining the Relationship Between Teacher Math-Gender Stereotypes and Students' Math Outcomes.

Conference Paper & Poster Presentations

Burrell, N.**, Conlon, R.A.,** Geer, E.A., Ganley, C.M., Hart, S. (2020, October). Examining the home math environment as a mediator of the relation between parent and child math anxiety. Presented in poster session at the Home Mathematics Environment Virtual Conference.

Conlon, R.A., Barroso, C., Schoen, R., Schatschneider, C., Ganley, C.M. (2020, June). Examining the relationship between math-gender stereotyping and gender-typicality of career aspirations in children. Accepted for poster session at the Math Cognition and Learning Society (MCLS) Conference, Dublin, Ireland [Cancelled due to COVID-19 pandemic].

Conlon, R.A., Barroso, C., Schoen, R., Schatschneider, C., Ganley, C.M. (2020, April). Mathematics anxiety in kindergarten students concurrent and longitudinal relations with mathematics performance. Accepted for poster session at the American Educational Research Association (AERA) Conference, San Francisco, CA [Cancelled due to COVID-19 pandemic].

Ganley, C.M., Barroso, C., Geer, E.A., **Conlon, R.A**., Schoen, R., Schatschneider, C. (2020, April). Teacher math knowledge, attitudes, and beliefs as predictors of instructional practices and student math learning. Accepted for paper symposium at the American Educational Research Association (AERA) Conference, San Francisco, CA [Cancelled due to COVID-19 pandemic].

Conlon, R.A., Ganley, C.M., Barroso, C., McGraw, A.L., Geer, E.A. (2019, April). Exploring gender differences in the career aspirations of early elementary school children. Presented in Poster Session at FSU Department of Psychology Graduate Research Day, Tallahassee, FL.

Ganley, C. M., Barroso, C., Geer, E. A., **Conlon, R.A.,** McGraw, A. L., Schoen, R., Schatschneider, C. (2019, March). Mathematics anxiety in kindergarten students concurrent and longitudinal relations with mathematics performance. Presented in paper symposium at the Society for Research in Child Development (SRCD) Conference, Baltimore, MD.

Research Experience

Graduate Research Assistant **Summer 2020**
National Project on Achievement in Twins (NATPAT), funded by the funded by the National Institute of Child Health and Human Development
Principal Investigator: Dr Sara Hart

Graduate Research Assistant **Fall 2019-Present**
Research on Experiences, Attitudes, and Learning in Mathematics (REALM) Project, funded by the Institute of Education Sciences
Principal Investigator: Dr Colleen Ganley

Individual Research Project, Trinity College Dublin **Fall 2014-Sping 2015**
Project Title: *The Relationship Between Girls' Self-Concept and School-Type in an Irish Adolescent Population: A Mixed Methods Approach*
Advisor: Dr. Lorraine Swords

Group Research Project, Trinity College Dublin **Fall 2013-Spring 2014**
Project Title: *Exploring the Sibling Experience of Autism Spectrum Disorder*
Advisor: Dr. Siobhán Corrigan

Academic Employment

Florida State University **Fall 2019-Present**
Teaching Assistant, Research Methods in Psychology

Florida State University **Fall 2018 – Summer 2019**
Departmental Assistant, Dept. of Psychology

Volunteer/Service Experience

Big Brothers Big Sisters **Spring 2020-Present**
Mentor/ Big Sister

Skype A Scientist **Fall 2019-Present**
Classroom Speaker

Centre for Global Engagement, FSU **Spring 2019**
Global Ambassador for Ireland

Student 2 Student, Trinity College Dublin **Fall 2012 – Summer 2015**
Peer Supporter

Fighting Words Dublin **Spring 2013-Summer 2015**
Creative writing workshop facilitator

Voluntary Tuition Programme, Trinity College Dublin **Fall 2014-Summer 2015**
English, Irish, Math & Science tutor

TCD Psychological Society **Fall 2011-Summer 2014**
Chairperson & Committee Member

Awards and Funding

Watson Scholarship (2018) Florida State University	$1000
Jane West Scholarship (2018, 2019) Florida State University	$3000
Ray Fuller Prize (2015) Trinity College Dublin	€160

Professional Associations

Math Cognition and Learning Society	Student Member
Providing Opportunities for Women in Education Research (POWER)	Student Member

Professional Development and Training

Introduction to Structural Equation Modeling, Curran-Bauer Analytics	May 2020
Multilevel Modeling: A Second Course, Statistical Horizons	October 2019