Open access • Report • DOI:10.18235/0000446

Under the "Cloak of Invisibility" : Gender Bias in Teaching Practices and Learning Outcomes — Source link

Marina Bassi, Rae Lesser Blumberg, Maria Mercedes Mateo-Berganza Diaz

Published on: 18 May 2016

Bassi, Marina; Blumberg, Rae Lesser; Mateo Díaz, Mercedes

Working Paper

Under the "Cloak of Invisibility": Gender Bias in Teaching Practices and Learning Outcomes

IDB Working Paper Series, No. IDB-WP-696

Provided in Cooperation with:
Inter-American Development Bank (IDB), Washington, DC

IDB WORKING PAPER SERIES Nº IDB-WP-696

Under the "Cloak of Invisibility": Gender Bias in Teaching Practices and Learning Outcomes

Marina Bassi
Rae Lesser Blumberg
Mercedes Mateo Díaz

Inter-American Development Bank
Education Division

May 2016

Under the “Cloak of Invisibility”: Gender Bias in Teaching Practices and Learning Outcomes

Marina Bassi*
Rae Lesser Blumberg**
Mercedes Mateo Díaz*

*Inter-American Development Bank
**University of Virginia

May 2016

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library
Bassi, Marina, 1975-
Under the “cloak of invisibility”: gender bias in teaching practices and learning outcomes / Marina Bassi, Rae Lesser Blumberg, Mercedes Mateo Diaz.
p. cm. — (IDB Working Paper Series ; 696)
Includes bibliographic references.
1. Sex differences in education-Chile. 2. Sex discrimination in education-Chile. 3. Women-Education-Chile. 4. Teaching-Social aspects-Chile. 5. Learning-Social aspects-Chile. I. Blumberg, Rae Lesser. II. Mateo Diaz, Mercedes. III. Inter-American Development Bank. Education Division. IV. Title. V. Series.
IDB-WP-696

http://www.iadb.org

Under the "Cloak of Invisibility"
Gender Bias in Teaching Practices and Learning Outcomes

Marina Bassi, Rae Lesser Blumberg, and Mercedes Mateo Díaz

Abstract[*]

This paper analyzes gender bias in teaching in low-performing schools in Chile. To carry out the analyses, the authors used videotaped classes for fourth graders and coded 237 tapings. Results show a general (although not uniform) bias in teachers' actions that resulted in less attention to female students. Gender bias had an even greater effect in classrooms where the teachers had worse interactions with students. Results show that less effective teachers (according to the Classroom Assessment Scoring System, or CLASS) show a larger gender bias. Greater gender bias is also correlated with lower scores for girls in Chile's standardized test (Sistema de Medición de la Calidad de la Educación, or SIMCE). With a few exceptions, the measures of gender bias in teacher–student interaction do not show statistically significant correlations with the test scores of boys.

JEL Codes: O12, J16, I2
Keywords: gender bias, gender gap, teacher quality, student learning

[*] Rae Lesser Blumberg is a William R. Kenan, junior professor of sociology at the University of Virginia. Marina Bassi is advisor to the Vice-President for Sectors and Knowledge at the Inter-American Development Bank (IDB). Mercedes Mateo Díaz is education specialist at the IDB. We are grateful for the assistance and guidance of many people. Gloria Burgoa, Olivia Feldman, Daniel Alonso and Ana Reynoso provided outstanding assistance for the literature review and analysis of the data. At the University of Virginia, Barbara Selby, helped us with her expertise to find helpful sources; and Nancy Kechner, an information technology expert, helped us to solve challenging problems with the coding of the videos. We are also grateful to David Sadker, Andy Morrison, and Norbert Shady for their invaluable insights; to the anonymous peer reviewers; and the participants of the APEC Annual Meeting and a seminar on gender and education organized by the Chilean Ministry of Education and the Ministry of Finance, both held in Santiago, Chile. All remaining errors are our own.

1. INTRODUCTION

In recent years, there has been increasing attention in the literature to systematic differences in educational outcomes by gender (Bos, Ganimian, and Vegas, 2014; Guiso et al., 2008; OECD, 2014b; Straus, 2015). In general, international test results show that boys perform better in math and science, while girls score higher in reading and language. According to results of the OECD's Programme for International Student Assessment (PISA) in 2012, girls attained on average 38 score points more than boys in reading, whereas boys attained on average 11 score points more than girls in mathematics (OECD, 2015). The gender gap in reading holds across all OECD countries.

Math results are especially important as they are good predictors of future achievement; school test scores and results in university access exams determine, amongst other factors, career decisions and will contribute to differences in employability, occupational segregation in the labor market, and differences in earnings (Bassi, Busso, and Muñoz, 2015; Bos, Ganimian, and Vegas, 2014; Heckman, 2011; Mizala, 2014; Murnane et al., 1995; Ñopo, 2012; Paglin and Rufolo, 1990).

Beyond the direct consequences of gender bias in terms of learning, there are also indirect implications for the transmission of cultural values to the younger generations. Culture plays an important role in shaping social norms and preferences that will later affect labor outcomes. There is a significant body of literature focusing on how cultural factors explain low female labor force participation (FLFP) (Fernández, 2013; Fernández and Flogli, 2009; Fogli and Veldkamp, 2011). Evidence shows that changes in cultural values will affect behavior in different domains (Harrison and Huntington, 2000; Inglehart and Norris, 2003). This paper looks at how the education system contributes to recreate and perpetuate gender stereotypes that will affect learning as well as other social outcomes during the lifecycle.

What lies behind gender disparities in performance? Learning gaps seem to increase with age, and they cannot be explained by parental background and investment in the child, unobserved ability, teacher gender, or other observable variables of classroom environment (Bharadwaj et al. 2015). Teaching quality is likely to be an important part of the explanation, as evidence has consistently shown it to be one of the key components of learning in general (Araujo et al. 2016; Bruns and Luque, 2015). A high quality teacher could represent a gain for students of one additional school year (Rivkin, Hanushek y Kain, 2005), and effective teachers can also have long-term positive impacts on outcomes such as university enrollment and income (Chetty, Friedman, and Rockoff, 2011).

This paper identifies whether gender stereotypes are present in teaching practices. It also associates the presence of these stereotypes to measures of teacher quality and differences in students' learning outcomes. The paper analyses a sample of fourth graders and their teachers in Chile to examine what happens inside the classroom in terms of gender biased teacher–student interactions, aiming to understand to what extent these biases are associated to different learning outcomes of boys and girls.

There is a broad literature on gender bias in teaching—especially for developed countries—and it is possible to find many patterns within this literature. Davis (2000) presents a broad review of research in this topic. Various articles point to a prevalence of gender bias in favor of male students consistent across subject areas and school environments (Biraimah, 1989; Brady and Eisler, 1995; DeVoe, 1991; Jones, 1989a; Jones, 1989b; Sadker and Sadker, 1994). The most common biases are found under the form of teachers' giving more attention to boys than girls (AAUW, 1992; LaFrance, 1991; Sadker and Sadker, 1994; Sadker, Sadker and Stulberg, 1993); segregation in the classroom or calling attention to a specific characteristic of students, such as gender, race, or ethnicity (Davis, 2000); and preexisting beliefs and expectations about students' abilities, skills, and competence in different subjects (Fennema et al. 1990; Lavy and Sand, 2015; Li, 1999; Robinson and Lubienski, 2011; Robinson, Lubienski, and Copur, 2011; Robinson-Cimpain et al., 2014; Schwartz and Sinicrope, 2013; Shakeshift 1995; Tiedemann, 2002; Van Duzer, 2006).

The literature indicates that getting more of a teacher's attention—whether positive (e.g., the teacher responding to or working one-on-one with the student) or negative (e.g., the teacher disciplining the student)—has consequences for students' performance (Lavy and Sand, 2015; Sadker and Sadker, 1994).

This study builds on the work of Sadker and Sadker (1994), expanding their coding categories to analyze if teachers in fact pay more attention to students by gender and if that greater attention is associated with better learning outcomes. We examine almost 590 hours of videotaped classes for fourth grade students in Chile. Following the variables defined in Sadker and Sadker (1994) (remediation, praise, criticism, and acceptance), we "quantify" teachers' interactions (episodes) with girls and boys during these classes. In addition, for each class we track the amount of time that teachers spend with girls versus boys, count the number of interventions of each gender, and analyze the teachers' (positive or negative) responses. We then link these data with students' scores in national standardized tests (e.g., Sistema de Medición de la Calidad de la Educación, or SIMCE) and with teachers' performance measured through the Classroom Assessment Scoring System (CLASS).

The results show a general (although not uniform) bias in teachers' actions that resulted in less attention to female students. Quantitative results show that, on most of the conceptual variables coded, girls received less attention from teachers than boys. Furthermore, in terms of time spent (recorded in seconds, minutes and hours), the results also show de facto teacher bias in the form of under-attention to girls; that is, when teachers address the class, they tend to focus more often on boys than girls. Moreover, boys tend to participate more (i.e., they call out answers more often that their female peers). These results mostly hold when adjusting for the fact that our sample includes slightly more boys than girls (53.2 percent boys). The differences are particularly large in terms of teacher criticism, call out (students' participation), and teacher response. When we compare the entire distribution of a gender bias index based on the variables measured, the results show that there are statistically significant differences in favor of boys, even after adjusting by the number of boys in the classroom.

The extent of gender bias in the classroom correlates with the quality of teacher–student interactions in three domains: emotional support, class organization, and instructional support.[1] The instrument used to rate the interactions is CLASS, with scores ranging from 1 (bad interactions) to 7 (excellent interactions). The results from this study show that worse teachers (according to CLASS) demonstrate larger gender bias. This association is larger in the emotional support and class organization domains, but is also present in the instructional support domain.

Another important finding of this paper is the relationship between gender bias in the classroom and student performance, measured by SIMCE scores. Girls whose teachers demonstrate greater gender bias tend to perform worse in all subjects. With a few exceptions, the measures of gender bias do not show a statistically significant correlation with SIMCE scores for boys.

[1] These three domains are subdivided into 11 dimensions. Emotional support includes the dimensions of positive climate, negative climate, regard for student perspectives, and teacher sensitivity. Classroom organization includes the dimensions of effective behavior management, instructional learning formats, and productivity. Instructional climate includes the dimensions of language modeling, concept development, analysis and inquiry, and quality of feedback.

The type of data collected for this study does not identify the causes of teacher behavior. Teachers might be (consciously or unconsciously) following socially established gender stereotypes, or simply be reacting to student behavior, spending some of the extra time to discipline boys who are being unruly or to attend to those who are considered high participators (e.g., by shouting out answers or otherwise attempting to draw attention to themselves). Also, we cannot establish pure causality between differences in SIMCE scores based on gender and teacher bias. The association we find can in fact reveal the effect of teacher bias on student learning; however, teachers could also be paying more attention and responding more frequently to more demanding students—independently of their gender—which in this case happen to be boys. Although the literature on gender differences in the classroom strongly points to gender bias as the reason behind larger gaps in tests scores among boys and girls, it is important to acknowledge this source of endogeneity.

The fact that gender biases are observed in all subjects raises the question of why girls do better in language. An important factor behind these results could be that the process of learning is different in math than it is in other subjects. Reading can be improved by continued reading (something students can do on their own), whereas mathematical thinking requires "engaging students in posing and solving problems" (Fite, 2002). The fact that certain subjects are more teacher-dependent than others could explain part of these differences. A recent meta-analysis of the effects of early math and literacy programs on skills suggests that programs implemented with parents and those implemented in centers produce similar results in terms of learning when it comes to literacy; for math, however, learning results are stronger when programs are implemented in the centers (Naslund et al., forthcoming).

Finally, parents also have a role in shaping their children's reading and work patterns at home and in developing their aspirations for their future. PISA data shows that, even at the same level of performance, parents have higher expectations of their sons to work the science, technology, engineering, and mathematics (STEM) fields (OECD, 2015).

The main contribution of this paper is the documentation of patterns of gender-biased behavior among teachers in one of the countries with the greatest differences in test scores by gender in the world. The correlation of gender bias to both CLASS and SIMCE scores strengthens the argument and invites researchers to develop studies to identify causality with other data sources. The paper is organized as follows. Section 2 discusses the apparent paradox between education outcomes and labor outcomes in Chile. Section 3 reviews the related literature. Section 4 presents the sample characteristics and methodology. Section 5 provides the findings and analyzes the correlation between teacher gender bias and student academic performance, as well as between teacher gender bias and measures of teacher quality. The final section offers conclusions to the analyses herein, as well as considerations for future policy implications and alternatives.

2. LITERATURE REVIEW

Since the 1990s, there has been a growing body of evidence looking at factors driving student performance, accompanied by an increasing use of education indicators and student assessments. Among the different school attributes, this paper concentrates on one element: teacher–student interactions, with a particular focus on how these interactions vary between girls and boys.

Evidence indicates that girls and boys enter school with similar abilities, but differences in learning outcomes start to appear as they advance in grade levels and years of schooling. Gender bias tends to influence the performance of girls in math and science and of boys in reading and writing. One of the most influential studies of gender bias in the classroom, the meta-review of empirical research by the American Association of University Women (AAUW 1992), concludes that in the United States, girls and boys start primary school with equally strong interest and achievement in math and science. By sixth grade, however, girls have fallen behind, and by the end of twelfth grade, girls have less confidence in their abilities in these subjects than boys. Also, Robinson and Lubienski (2011) find that the math gender gap does not exist when children start kindergarten, but grows to nearly 0.25 of a standard deviation by third grade.

In a study of academically gifted students in the United States, Olszewski-Kubilius and Turner (2002) find that boys begin to outperform girls in math starting in third grade, although by only a small amount. By fifth and sixth grade, however, the differences were significant for boys achieving very high scores. For this study, both boys and girls named math as their favorite subject. However, more girls thought their best academic strengths were in language, and more boys named math and science as their strongest subjects.

On an international level, Guiso et al. (2008) find that the gender gap in math scores among 15-year-olds on the PISA tests narrows as the level of gender equality rises.[2] The gap is widest in Korea and Turkey, the two countries with the worst Gender Gap Index (GGI) scores, and shrinks with improving GGI scores. In terms of the three nations with the best GGIs, there is virtually no gap in Norway and Sweden, and the gap slightly favors girls in Iceland. Since this study, there have been more places outside of Scandinavia where the math gender gap has reversed or appears to be vanishing all together. The PISA 2012 results show that girls in Thailand, for example, outscore boys by 14 points in math—the highest margin in East Asia and Southeast Asia and the third highest gap in favor of girls among the 65 countries and economies that participated (OECD, 2014b).

The results of PISA 2012 point to widening deficiencies in reading and writing between boys and girls. For example, the gender gap in reading performance—favoring girls—increased in 11 nations from 2000 to 2012 (Straus, 2015). The gender gap seems to be wider in working-class/low socioeconomic status (SES) boys than for middle-class boys, according to Entwistle Alexander, and Olson (2007). These authors point out that middle-class parents are more likely to encourage their sons to engage in activities that enhance school performance in these areas, such as reading for pleasure or playing board games. Two Australian studies make similar points. Teese et al. (1995) argue that boys from middle-class families are doing better than girls from working-class families, and Collins et al. (2000) find that socioeconomic status makes a bigger difference than gender in performance in English.

Teachers' expectations matter, whether they are about boys' likelihood of academic failure (Van Duzer, 2006) or girls' competence in math and science (Shakeshift, 1995). Schwartz and Sinicrope (2013) indicate that many teachers have preexisting biases about girls' deficiencies in mathematics. They point out that only 27 percent of undergraduates studying education view

[2] The study measures the countries' gender equality levels using the World Economic Forum (WEF) Gender Gap Index (GGI), first introduced in 2006. The GCI quantifies the magnitude of gender-based disparities and provides country rankings across four areas: health, education, economy, and politics. The calculation of the gender gap in health is based on two components: sex ratio at birth and the gap between women's and men's healthy life expectancy. The calculation of the gender gap in education is based on four components: access ratios of women to men in primary, secondary, and tertiary education levels and female to male literacy rates. The calculation of the gender gap in economic participation is based on four components: female labor force participation over male labor force participation; earnings disparity between men and women at similar job positions; labor income for women in relation to men; and proportion of female professional and technical workers. The calculation of the gender gap in politics is based on two components: the ratio of women to men in parliamentary positions and the ratio of women to men in terms of years in an executive office (prime minister or president) for the last 50 years.

girls' math abilities positively. Backing this, Fennema et al. (1990) find that teachers attribute math competence in male students to ability but in female students to effort. Similarly, Tiedemann (2002) finds that primary school teachers view average-achieving girls as less capable in math and having to work harder than average-achieving boys. Previous studies, such as Li (1999), observe math classrooms and find that teachers tend to see math as a male domain and have higher expectations of boys. Robinson, Lubienski, and Copur (2011) examine whether teachers in a national sample rate male math proficiency higher than female math proficiency, even when they (a) behave similarly, (b) have similar approaches to learning, and (c) have the same past and current test scores. These authors find that teachers rate the math skills of girls lower than the math skills of boys, which they find to be a consistent pattern throughout primary school, even when the girls performed as well as the boys. Their study suggests that if teachers believed that girls had the same abilities in math as boys with similar performance levels, girls would lose about 40 to 75 percent less ground in math achievement from the end of kindergarten to the end of fifth grade.

Schumow and Schmidt (2013) use observational methods as part of their study of the effect of teachers' beliefs and practices on the performance of girls in U.S. high school science classes. Girls and boys in this study achieved the same grades in science, but the girls rated themselves as significantly less competent and had less positive attitudes toward the subject. The authors find that teachers spend an average of 39 percent more time addressing male students than their female counterparts, a finding not explained by student initiation. They also find that teachers address boys more often than girls about content/knowledge issues. Although the teachers in this study verbally denied there were gender differences in science performance, the observations reveal that they held implicit beliefs suggesting gender bias.

More recently, Lavy and Sand (2015) examine the short- and long-term impacts of teachers' gender bias on the academic achievements of girls and boys in Israel during middle and high school and on their decision to take advanced level courses in math and sciences during high school. The main finding is that teachers' bias favoring boys have a positive effect on the achievements of boys and a negative effect on those of girls. The teachers seem to unconsciously discourage the female students by underestimating their abilities, while overestimating the skills of their male classmates. This was revealed through the results of two exams given to three groups of students from sixth grade through high school.[3] Instructors who did not know the students graded one exam, while instructors who did know them graded the second. In the first case, the girls outscored the boys, while in the second, the boys outscored girls. This effect was only true for the subjects of math and science. Moreover, the boys who were encouraged when in elementary school scored significantly higher in their middle school and high school matriculation exams than the girls, even though the girls had outperformed the boys on the exams graded by the instructors who did not know them. The impact continued in high school, when researchers found that girls who had been discouraged by their elementary school teachers were much less likely than boys to take advanced math and science courses. Another key result is that teachers' biases at early stages of education have long-term implications for female and male occupational choices and earning potentials, because taking advanced courses in math and science is a prerequisite for post-secondary schooling in STEM fields.

Teachers' gender bias affect students' self-perceptions. Banjong (2014) focuses on math proficiency and attitudes among fourth through seventh graders in the United States, finding no significant differences between male and female performance levels. Nonetheless, boys label math as one of their best subjects and feel successful at it, while girls largely express the opposite. Finally, complementing studies that focus on teacher practices, the literature on gender bias in the classroom also considers textbooks and other factors that potentially affect differences in learning, such as class size, density, and configuration.[4] Chile is the site of the first study of gender bias in textbooks. Magendoza (1970) provides a path-breaking analysis of the depiction of females in Chilean textbooks, finding them underrepresented and portrayed in stereotyped roles.[5] Since then, findings about gender bias in textbooks have been remarkably similar worldwide: (i) females are underrepresented, whether measured in lines of text, proportion of named characters (human or animal), mentions in titles, or citations in indexes; and (ii) females and males are depicted in highly gender-stereotyped ways in the household as well as in the

[3] The authors compile a dataset using different sources of administrative data and national exams and school records. With the consolidated dataset they can track students from primary school through middle and high school. To measure teachers' biased behavior, they look at differences between non-blind and blind assessments, comparing students' tests scores from the blind national external exam (Growth and Effectiveness Measures for Schools-GEMS) and the non-blind internal school exams.

[4] The relationship between class size and overall student performance remains debatable and few studies exist on whether or not differences in student performance produced from differences in class sizes are gender-based; see Musua and Migosi (2013) for an overview. The overall—albeit preliminary—lesson here is that large, overcrowded classrooms seem to affect girls more, especially in math and other STEM fields.

[5] This was a year before Uren (1971) and Trecker (1971) published the first studies in the United States.

occupational division of labor and in their general actions, attitudes, and character traits (Jassey, 1998). Textbooks through the 1990s portray women as disproportionately passive and dull in comparison with men, and present them in only the most stereotypically feminine of occupations (Blumberg 2007, 2008, 2015a). Recent research on gender bias in textbooks in Chile shows continuing underrepresentation of females and stereotyping of both male and female occupational roles and character traits (Blumberg 2015a; Covacevich and Quintela-Dávila, 2014; Duarte, 2010).

The combination of two forms of the "hidden curriculum"—gender bias in textbooks and gender bias in teaching—could comprise twin and often invisible obstacles. These barriers can dampen students' ambitions and nudge girls toward traditional feminine roles and occupational pursuits that will affect short-term learning outcomes and long-term labor and social outcomes.

3. Education, Learning, and Labor Outcomes in Chile

Using data for Chile, this paper analyzes gender gaps in education with an innovative approach, focusing on what happens inside the classroom to determine whether teachers interact systematically differently with boys and girls. Chile presents a great paradox with respect to gender. It has basically full equality in terms of health and survival and parity in literacy and enrollment in education; however, according to the 2015 Gender Gap Index, it ranks 123 out of 145 countries with respect to women's economic participation and opportunity. More importantly, the significant gap in terms of economic equality between genders lowers its overall ranking to 73 (WEF, 2015), which is a significant decline from its position of 90 in 2006.

This paradox unveils a number of visible issues; but, as this paper argues, there is more to it than meets the eye. Chilean girls do not seem to be getting the same learning returns as their male classmates based on school attendance. Based on the Programme for International Student Assessment (PISA) 2012, Chile's gender gap in math scores was the largest among member countries of the Organisation for Economic Co-operation and Development (OECD), with boys outperforming girls (OECD, 2012).[6]

[6] The OECD launched PISA in 1997, with the objective to develop regular, reliable, and policy relevant indicators on student achievement. PISA 2012 is the 5th survey, which assesses competencies of 15-year-olds in reading, mathematics, and science (with a focus on mathematics) in 65 countries and economies. For more information, see http://www.oecd.org/pisa/home/.

The percentage of females completing tertiary education in Chile is one of the highest amongst OECD countries; yet, the percentage of those that participate in the labor market is one of the lowest. It is also interesting to note that the completion rate is higher for females than males in Chile at almost all education levels (Bassi, Busso, and Muñoz, 2015).

According to the results in PISA 2012, the gender gap in math learning outcomes in Chile is the largest among OECD countries. Boys perform better than girls in mathematics in 37 of the 64 countries surveyed, and Chile ranks below average in overall mathematics performance, as well as in equity in education outcomes (OECD, 2014b). School test scores and results in university entrance exams, amongst other factors, determine career decisions that later translate into occupational segregation in the labor market, and math results, in particular, are a good predictor of future earnings (Bassi, Busso, and Muñoz, 2015; Bos, Ganimian, and Vegas, 2014; Heckman, 2011; Mizala, 2014; Ñopo, 2012). Understanding what happens in schools might provide some explanations for these large gender gaps.

4. SAMPLE AND METHODOLOGY

4.1. School Samples

The analysis of gender bias in teacher practices for this study is based on videotapes from the end of 2012. The sample includes 237 tapes from the classrooms of 137 academically low-performing schools in Chile (according to SIMCE scores from 2009). These schools belong to a random sample designed for an impact evaluation of a program implemented in 2011 by Chile's Ministry of Education to improve learning outcomes in math and language of students from prekindergarten to fourth grade (Bassi, Meghir, and Reynoso, 2015). Eligible schools comprised public and subsidized private schools complying with two main additional criteria: (i) the school average SIMCE score for the years 2005–09 in math and language should be below the national average (250 points); and (ii) there should be at least 20 students per level on average from prekindergarten to fourth grade. SIMCE scores were available for the year of the tapings for the schools, students, and teachers. In addition, the same videos were coded using the renowned CLASS instrument, which provides good measures of specific dimensions of teaching quality (Pianta et al., 2008).

Since this is not a representative sample, results cannot be inferred to all Chilean schools. Descriptive statistics show that the gender gap in these low performing schools is higher than the average; it is therefore more likely to detect bias in low performing schools. Yet, the large gender gap observed both in PISA and national exams is representative of all schools. Therefore, besides a question of magnitude, there is no reason why the gender dynamics described in this paper should be different between schools. What we are seeing in teachers' interactions in the poorest schools can also explain what is going on, at a different level, in other schools.

Each of the 237 videotapes shows a single teacher instructing fourth grade students. The program does not include any component or activity explicitly addressing gender bias in teaching. Thus, a particular advantage of this study is that, since the videotapes were made for another purpose, they proved amenable for coding gender bias in the classroom, without any reason to believe the teachers might be altering their behavior in this aspect. The filming strictly followed the protocol of the upper elementary version (fourth to sixth grade) of CLASS (Pianta et al. 2008). As previously discussed, CLASS measures teacher–student interactions, and there are several studies that link better student outcomes (both in learning and the development of socioemotional skills) with teacher scores.[7] One teacher in each of the fourth grades of the 137 schools in the sample was videotaped for four pedagogic hours (including 69 math classes and 168 classes pertaining to other subjects, mainly language arts).[8]

Table 1 presents the main descriptive statistics for our sample of schools, most of which are public schools with students of medium-low socioeconomic status (according to the SIMCE classification). Approximately 63 percent of the schools have only one fourth grade class, while the rest have no more than three classes in that level. The average number of students in the classrooms surveyed is 29 with a gender distribution of 53 percent boys and 47 percent girls.

The average 2012 SIMCE score of the schools in the sample, for combined subjects, is 243 (Table 2), which is below Chile's national average of 262. The score gap by gender is 10 points in language arts (favoring girls) and 3 points in social sciences (favoring boys), both similar to the gap at the national level. In math, the gap is 5 points (favoring boys), 3 points higher than Chile's average gap.

[7] For example, Araujo et al. (2016) briefly review this literature for the United States and present a study for children in kindergarten in Ecuador.
[8] Some videos included classes for more than one subject. In those cases, each subject was coded separately.

Table 1: Descriptive Statistics of Schools in the Sample

	Frequency	**%**
Administrative dependency		
Municipal (public)	88	64.7
Private, subsidized	48	35.3
Socioeconomic level		
Low	25	18.4
Medium-low	88	64.2
Medium	24	17.5
Number of 4th grade basic level courses by school		
1	85	62.5
2	43	31.6
3	8	5.9
Average number of 4th grade basic level students by school		
<25	39	28.7
25–34	71	52.2
>34	26	19.1
Number of teachers recorded by school		
1	104	76.5
2 or more	32	23.5
Average number of 4th grade basic level students by school	29.2	
Average percent of boys by school	53.3%	
Total	**137**	

Source: Authors' elaboration based on schools' administrative data and SIMCE results.

Table 2: Average 2012 SIMCE Scores of Schools in the Sample

	Sample schools			Average Chile		
Subjects	Total	Boys	Girls	Total	Boys	Girls
Language arts	250	244	255	267	261	273
Mathematics	242	244	239	261	262	260
Social sciences	238	239	236	258	259	257
Combined	243	243	243	262	260	263

Source: Authors' elaboration based on schools' administrative data and SIMCE results.

A teachers' questionnaire, including questions on their educational background, experience, and tenure, complemented the videos. Most teachers in the schools studied were female and had achieved a university degree (Table 3). Their experience in teaching was homogeneously distributed; almost the same number of teachers had less than 5 years of experience, between 5 and 10 years, between 11 to 24 years, or more than 25 years.

Table 3: Descriptive Statistics of Teachers in the Sample

	Frequency	%
Gender		
Female	161	92.5
Male	9	5.2
NA	4	2.3
Type of institution where degree was received		
Professional institution	15	8.6
University	156	89.7
Regular school	3	1.7
Specialization/post graduate studies		
Yes	102	59.3
No	69	40.7
Teaching experience (in years)		
<5	43	24.7
5–10	40	23
11–24	46	26.4
>25	45	25.9
Tenure (in years)		
<3	43	24.7
3–4	41	23.6
5–13	46	26.4
>14	44	25.3
Total	**174**	

Source: Authors' elaboration based on schools' administrative data.

5. Coding Categories and Strategy

Sadker and Sadker (1994) originally developed the basic coding criterion used herein. To the best of our knowledge, only a few studies have used videotapes to look at gender bias in teacher classroom practice (e.g., Davis, 2000; Sadker and Sadker, 1994). Sadker and Sadker (1994) utilize observation and analysis of videotapes to study gender bias in teachers' differential use of attention in their interactions with female versus male students in the fourth, sixth, and eighth grades in four states and the District of Colombia in the United States. They establish empirically that (i) praise, criticism, and remediation all are better forms of attention; (ii) boys receive more of this better kind of attention; (iii) boys are eight times more likely than girls to call out during a discussion (i.e., shout out answers even when they are not called upon); (iv) teachers are less likely to reject boys' behavior, although it is against classroom rules;[9] (v) teachers interact more with male students; (vi) girls receive more "acceptance," which is a bare acknowledgment (e.g., "uh-huh" or "okay") of a student's statement or work; and (vii) girls who receive less attention from their teachers may come to underestimate their abilities and lose motivation. The authors' conceptual framework is refined and completed herein with other dimensions gathered from literature on gender bias in the classroom, including time spent with girls versus boys and the level of control a teacher has over his or her classroom. The coding scheme used in this paper consists of the following variables:

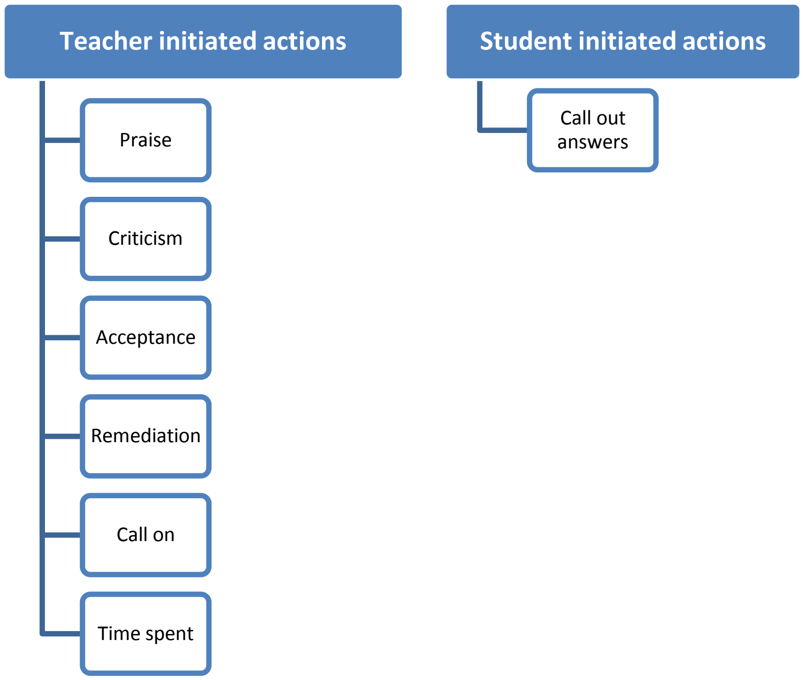
- The four Sadker and Sadker (1994) variables:
  - Praise: after correct answers given, unsolicited, or general. Examples include: "Good job." "That was an excellent paper." "I like the way you're thinking."
  - Criticism: negative comments or discipline, giving an explicit statement that something is not correct. Examples include: "No, you've missed number four." "This is a terrible report."
  - Remediation: helping a student, encouraging him or her to correct a wrong answer, or expanding and enhancing his or her thinking. Examples include: "Check your addition." "Think about what you've just said and try again."
  - Acceptance: acknowledgement of correct answers given when called on, when the student calls out, or during quiet work. Examples include: "Uh-huh." "Okay."
- Calling on by name: when the teacher calls a student by name or asks him or her to answer a question or speak to the class/participate.
- Time spent: calculated as how much time the teacher spends with individual students or groups of students of a single gender (divided into time segments).[10]
- Students call out: students shouting out answers when they are not called on by the teacher.

This study classifies classroom behaviors by those initiated by the teachers and those initiated by the students (Figure 1). In addition, the coding sheets include a section where coders can indicate their perception of the level of control observed in the classroom, which is ranked

[9] The teachers' rationale was that the boys tended to be more demanding and their tones and attitudes obliged teachers to respond.
[10] We used the video viewer to accurately aggregate the codes for time spent by the teacher with girls versus boys.

from 1 to 3, from poor control to good control respectively, as well as the level of gender bias, ranked from 1 to 3, from no obvious gender bias to significant/obvious gender bias, respectively.

Figure 1: Coding Scheme



Source: Authors' elaboration; categories are a refined and expanded version of Sadker and Sadker (1994).

Except for time spent (which accounts for the time the teachers spend with girls versus boys), coding consists of counting specific events that take place during a video and classifying them under a noted category (e.g., praise, criticism, remediation, among others). In this study, the coder only counted the specific events when he or she *clearly* observed them in the video. We used a formative method, starting the coding with a pilot sample of 10 percent (19 videos). We took this approach not only as a test of the coding variables but also to search for other dimensions that merited coding in the overall sample, involving the remaining 90 percent (169 videos). In addition to adding variables based on a given teacher's control of the classroom and any gender bias demonstrated by that teacher, we disaggregated subjects into math and language arts/other subjects. This was based on the "embedded" nature of some of the language instruction and the similarity of non-mathematics instruction.

The teachers taught the math classes in a distinct manner to the other subjects, using considerable board work and never resorting to reading out loud. The language arts classes were not the only classes that focused on language; for example, grammar and other aspects of language instruction were also incorporated into subjects such as social studies. Every tape contained some language instruction, however, and all the subjects except math used similar pedagogical techniques that included reading out loud by the teacher, individual students, and often the whole class in unison.

Five carefully trained coders coded 237 videotapes between October 2014 and November 2015. Although coders received thorough training ahead of time, as well as a clear template to ensure they applied homogenous criterion, video coding naturally involves a subjective component that might result in measurement error. To address this potential problem, the coders double-coded part of the videos, as done in comparable studies (Araujo et al., 2016; Brown et al, 2010). Two coders independently coded 41 percent of the videos and the inter-coder reliability rate was 90.7 percent,[11] which is a good result compared to other studies.[12]

6. RESULTS

6.1. Gender-Biased Variables

It is important to reiterate that the videotapes were not made with any intent to study teachers' gender bias in the classroom. Thus, it is highly unlikely that the teachers modified their behavior with respect to this variable because they knew that they were being videotaped. Table 4 shows the raw measures of the variables mentioned above—that is, counts of episodes (for the Sadker and Sadker variables, call on and call out) and time spent as recorded by the coders. The table separates the results by subject (math and others, which includes mainly language arts). All variables in both subjects show a bias in favor of boys. In math, differences are statistically significant in criticism, acceptance, call on, and call out. In other subjects (mainly language arts), all differences are statistically significant with the exception of praise.

Although we do not have a measure of which students were actually present in the classroom on the day of the filming, we know the class composition by enrollment. Since gaps may be reflecting, in part, the fact that there are a few more boys than girls among the

[11] The inter-coder reliability rate assumes that two codings do not coincide if the difference in their scores is more than one standard deviation (by category). When this criterion is modified to 0.5 standard deviations, the inter-coder reliability rate remains above 80 percent.

[12] For example, Brown et al. (2010) report an inter-coder reliability rate of 83 percent and Araujo et al. (2014) report a rate of 93 percent.

classrooms observed, we repeat the measures controlling for the number of students of both genders.

Table 4: Differences in Measures by Gender and Subject *(raw measures)*

Raw measures							
		Boys		Girls		Ttest	
Variables	Obs	Mean	Std. dev.	Mean	Std. dev.	Diff.	pval
	Math						
Praise	69	2.181	2.662	1.558	2.366	0.623	0.148
Criticism	69	6.152	5.731	2.261	2.896	3.891	0.000
Remediation	69	5.319	7.239	3.71	5.047	1.609	0.132
Acceptance	69	12.529	14.294	7.659	5.209	4.87	0.009
Call on	69	11.949	10.19	8.333	6.72	3.616	0.015
Time spent	69	1.833	2.516	1.469	2.176	0.364	0.364
Call out	69	14.442	22.688	5.819	7.197	8.623	0.003
	Other						
Praise	168	2.833	3.931	2.414	3.144	0.42	0.281
Criticism	168	5.739	6.368	2.459	4.226	3.28	0.000
Remediation	168	4.988	6.649	3.798	5.964	1.19	0.085
Acceptance	168	14.603	11.629	11.38	8.82	3.223	0.004
Call on	168	17.755	14.112	13.105	11.836	4.65	0.001
Time spent	167	1.702	2.37	1.199	1.982	0.502	0.036
Call out	168	13.846	13.362	7.063	7.487	6.784	0.000

Source: Authors' analysis based on primary data.

Table 5 shows the differences in measures and subject by student after this adjustment. With the exception of praise in non-math subjects, all gaps remain positive (in favor of boys); differences in criticism and call out are still significant, and call on, acceptance, and time spent are no longer statistically significant. In the case of math, gaps in criticism, acceptance, and call out response remain statistically significant. In sum, teacher's bias in favor of boys seems to be robust to the adjustment for the number of boys in the classroom.

Table 5: Differences in Measures by Gender and Subject *(measures per student)*

Measures per student *(by corresponding gender)*							
		Boys		Girls		Ttest	
Variable	Obs	Mean	Std. dev.	Mean	Std. Dev.	Diff.	pval
	Math						
Praise	63	0.175	0.236	0.163	0.299	0.013	0.792
Criticism	63	0.469	0.461	0.191	0.271	0.278	0.000
Remediation	63	0.467	0.785	0.338	0.557	0.129	0.291
Acceptance	63	1.031	1.023	0.729	0.642	0.303	0.049
Call on	63	0.935	0.799	0.776	0.697	0.159	0.236
Time spent	63	0.161	0.279	0.132	0.21	0.029	0.505
Call out	63	1.167	1.607	0.56	0.873	0.607	0.010
	Other						
Praise	139	0.213	0.293	0.226	0.33	-0.013	0.729
Criticism	139	0.444	0.541	0.175	0.276	0.269	0.000
Remediation	139	0.39	0.568	0.334	0.554	0.057	0.399
Acceptance	139	1.152	1.067	0.98	0.881	0.172	0.143
Call on	139	1.349	1.297	1.114	1.122	0.235	0.107
Time spent	139	0.127	0.206	0.102	0.217	0.025	0.327
Call out	139	1.128	1.198	0.613	0.733	0.515	0.000

Source: Authors' analysis based on primary data.

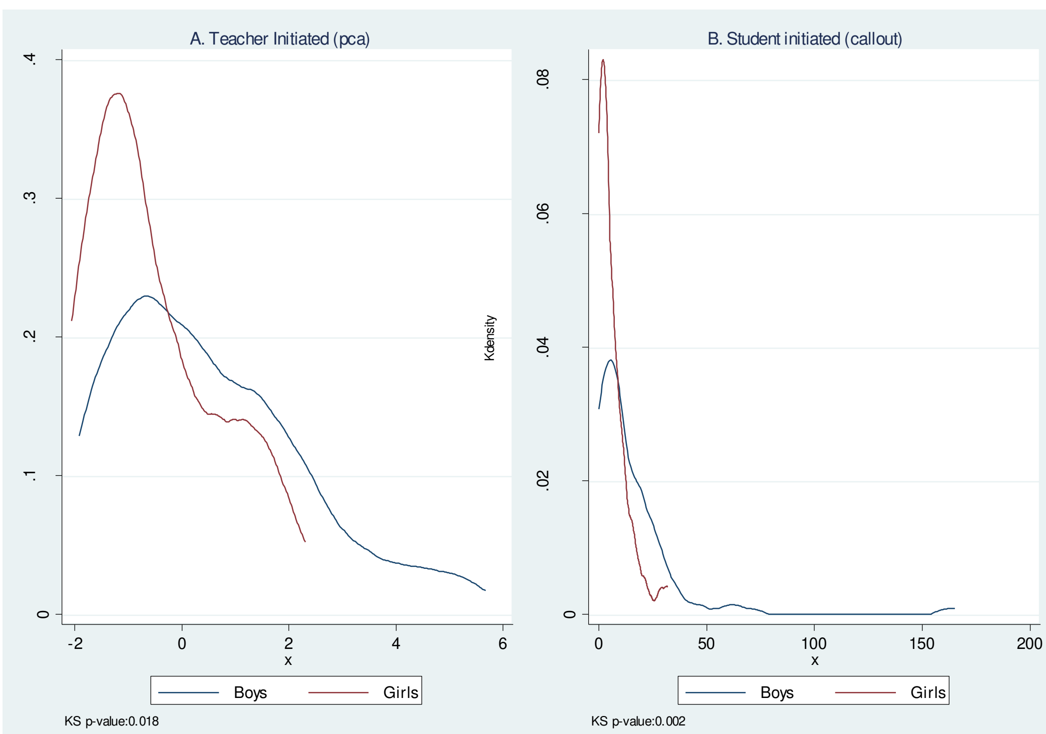
These results contrast the raw perceptions of the coders after watching the videotapes. As described above, the protocol included a question for the coders regarding the degree of gender bias observed in the teacher's practices. Interestingly, in nearly 83 percent of the videos, the coders reported that there was no obvious gender bias and in close to 13 percent, the coders reported noticing some obvious gender bias. The coders observed obvious gender bias in only two videos of the sample. Thus, the different interaction of teachers with boys versus girls seems to be quite subtle or invisible.

The previous analysis only tests for differences in the mean of the measures between boys and girls. Below, we evaluate the equality of the entire distribution of the measures across gender by graphical inspection and by performing the Kolmogorov-Smirnov test of equality of distributions. Using the variables described above, we construct two different measures. First, for all teacher-initiated variables (the four Sadker and Sadker variables, call on, and time spent), we take the first principal component of the different factors by subject and plot their distribution for boys and girls. The second measure is call out, the student-initiated variable. For all measures, we plot the distribution for girls and boys separately for math and language arts/other subjects. Then, we report the p-value of the test of equality of the principal component across gender.

Figure 2 shows the distribution for the raw measures for girls and boys in math (without adjusting for the number of students in the class). Panel A presents the first principal component of teacher-initiated actions (TIA), while Panel B illustrates the student-initiated actions (SIA) (i.e., call outs). Interestingly, TIA and SIA results show statistically significant differences in their distribution for boys and girls, with girls' distribution notably skewed to the left. This implies that there are practically no cases of high attention to girls (call ons) and there are few or no cases of high participation (call outs) of girls.

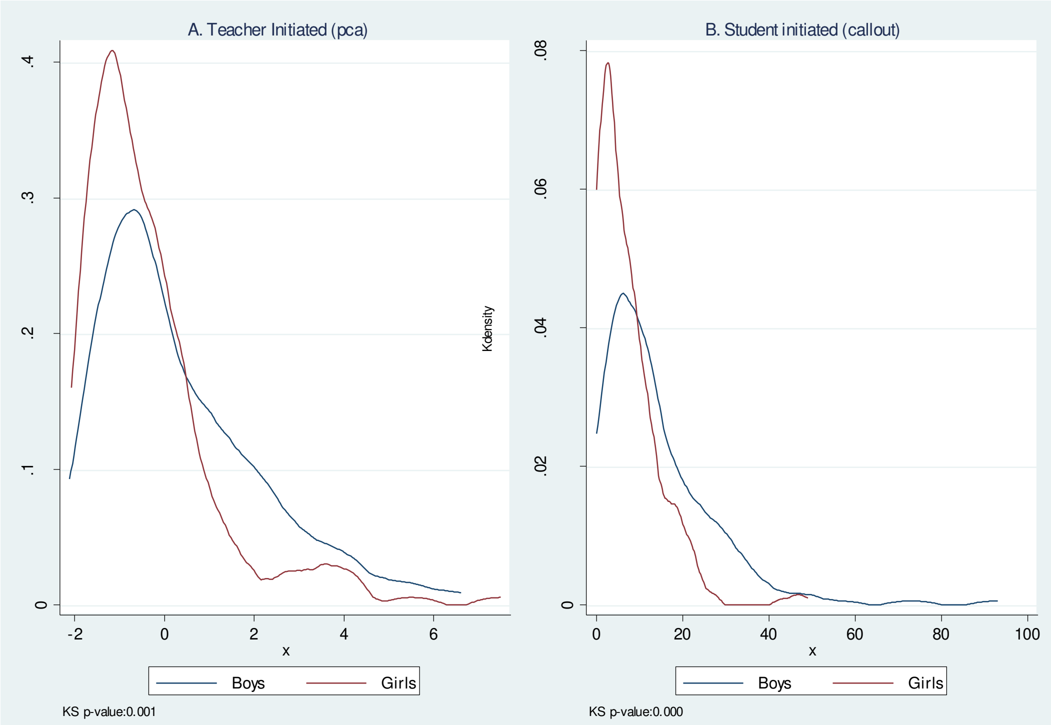
Figure 3 repeats the exercise for other subjects (mainly language arts), showing consistent differences between the distributions by gender. These time gender differences in teachers' negative responses to SIA are also significant.

Figure 2: Differences in the Distribution by Gender: Math *(raw measures)*



Source: Authors' analysis based on primary data.

Figure 3: Differences in the Distribution by Gender: Other Subjects *(raw measures)*

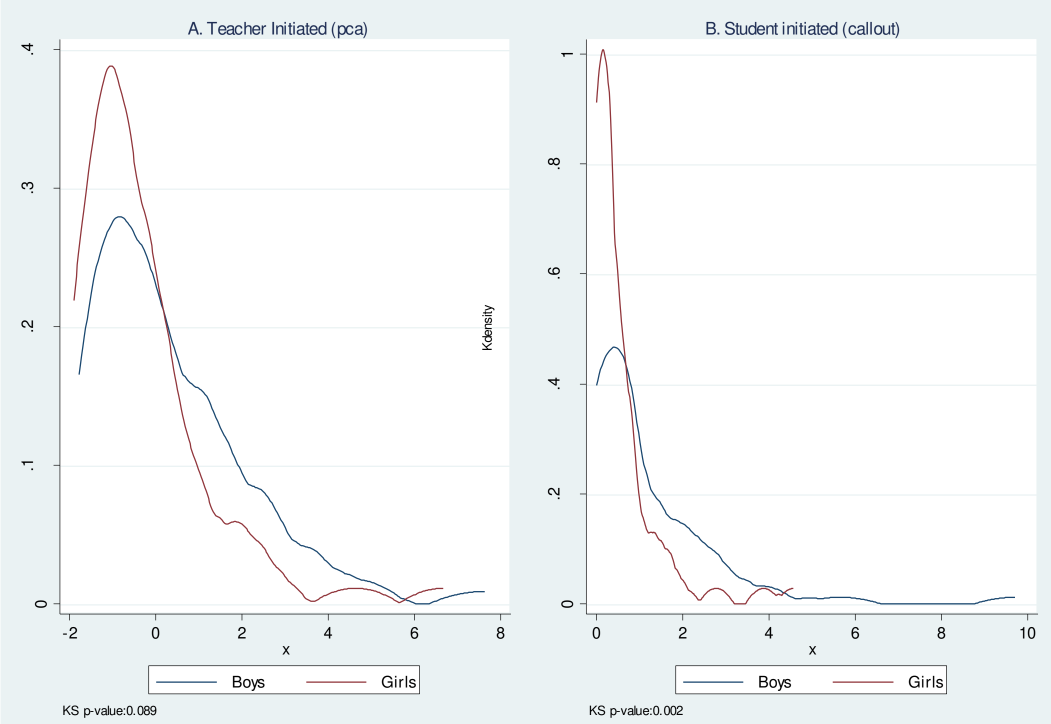


Source: Authors' analysis based on primary data.

To account for the fact that the classes in our sample include slightly more boys than girls, Figures 4 and 5 present the distributions for the same gender measures controlling by the number of students, instead of raw measures, for math and language/other subjects, respectively. In both math and language arts/other subjects, the differences between distributions by gender remain significant for both TIA and SIA, with skewed distributions for girls (to the left).[13]

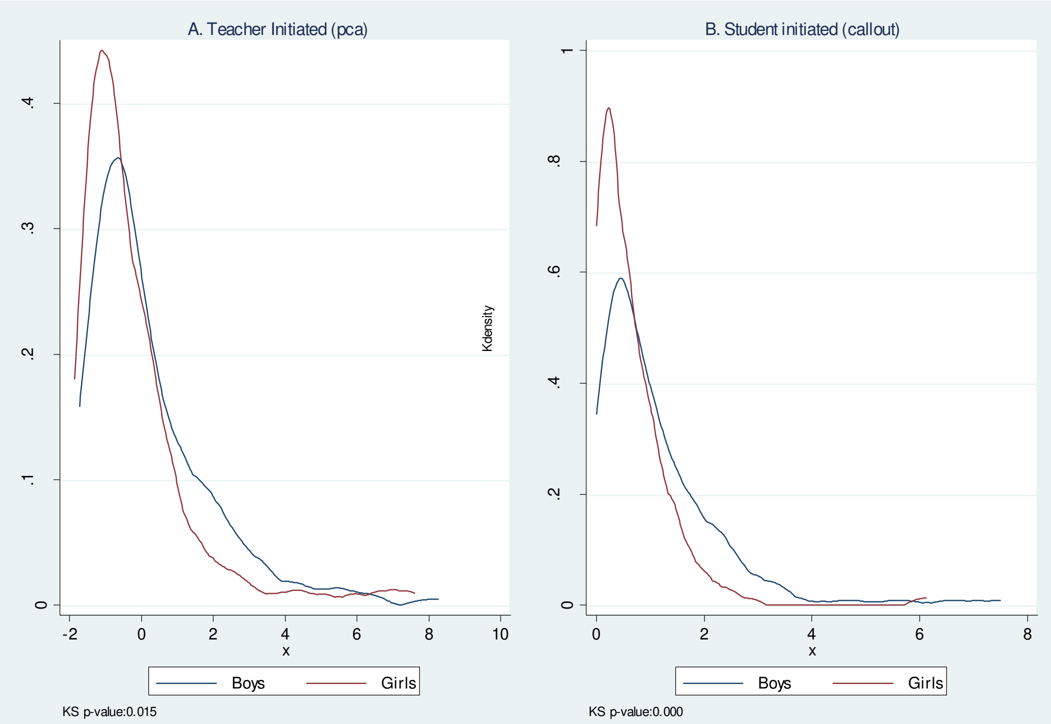
[13] When time spent is removed from the first principal component of TIA the differences are even stronger (p-value 0.000).

Figure 4: Differences in the Distribution by Gender: Math *(per student measures)*



Source: Authors' analysis based on primary data.

Figure 5: Differences in the Distribution by Gender: Language/Other Subjects *(per student measures)*



Source: Authors' analysis based on primary data.

6.2. Correlation of Gender Variables and SIMCE and CLASS Scores

The previous section finds evidence of gender bias in teacher practices and differences by gender in student behavior. Teachers in our sample spend more time with boys and give them more frequent attention. Also, boys call out answers in class more often that girls and receive a higher ratio of negative responses from teachers than their female classmates. Our data do not allow us to identify the causes of these behaviors, particularly in the case of the teachers. Teachers could be (consciously or unconsciously) responding differently to boys and girls because of gender bias, or they could be reacting to the more active participation (or greater indiscipline) of boys in class. Although it is difficult to identify what makes teachers behave systematically different with male and female students, we are able to associate gender bias estimates with measures of teaching quality.

The same 237 videotapes were coded using the upper elementary version (for fourth to sixth grade) of the CLASS instrument (Pianta et al. 2008). As mentioned earlier, CLASS measures the quality of teacher–student interactions in three main domains (emotional support, classroom organization and instructional support). Coders scored these interactions on a scale of 1 to 7 (Table 6). The average CLASS score for the teachers in our study is 3.95. There is only one case of high overall score (+5) and no cases of low general score (<=2). Scores tend to be higher in class organization and lower in instructional support.

Table 6: CLASS Scores of Teachers in the Sample

	Combined		**Emotional support**		**Class organization**		**Instructional support**	
	n	%	n	%	n	%	n	%
Low (<=2)	0	0	0	0	0	0	35	18.5
Medium (2–5)	188	99.5	187	98.9	10	5.3	154	81.5
High (+5)	1	0.5	2	1.1	179	94.7	0	0
Average	3.95		3.60		5.93		2.31	

Source: Authors' elaboration based on CLASS data.

Results show that gender bias measures are correlated with CLASS scores (Table 7). The regressor TIA is the difference between the first principal components of all teacher initiated actions (Sadker and Sadker variables, call on, and time spent) for boys minus the same measure for girls. The regressor SIA is the differences between boys and girls in calling out answers. For each regressor, the second column includes other covariates: characteristics of the school (type of administration, income decile, and experience and tenure the school principal) and characteristics of the classroom teacher (total experience, tenure in the school, and whether he or she has or is pursuing a graduate degree). We pool the subjects together and add a dummy variable that takes value 1 if the observation corresponds to math.

For the first two measures (TIA and SIA), the average CLASS score is negatively correlated with gender bias (worse teachers present higher gender bias) and is stronger when school, teacher, and student characteristics are included.

Table 7: Correlation between CLASS (pca) and Gender Bias (TIA and SIA)

VARIABLES	(1) TIA	(2) SIA
Class scores	-0.083**	-0.097*
	(0.040)	(0.049)
R-squared	0.133	0.104
Class organization	-0.356*	-0.424*
	(0.196)	(0.246)
R-squared	0.128	0.098
Emotional support	-0.387**	-0.440**
	(0.168)	(0.206)
R-squared	0.136	0.107
Instructional support	-0.277	-0.362
	(0.220)	(0.243)
R-squared	0.115	0.079
Observations	186	186

Source: Author's analyses based on primary data and CLASS data.
*Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. All specifications include covariates: characteristics of students (family income, school performance, household structure, and parents' education); characteristics of the school (type of administration, income decile, and experience and tenure the school of principal); and characteristics of the classroom teacher (total experience, tenure in the school, and whether has or is pursuing a graduate degree).*

Results are similar for the different domains in the CLASS score. First, for the class organization score, results are consistent. The correlation with the emotional support score remains negative for TIA and SIA. Lastly, for instructional support, correlations are statistically non-significant.

Overall, results indicate a correlation between measures of teacher quality and the different measures of gender bias, in particular for the emotional support score and for class organization. This finding is consistent with the literature suggesting that those gender biases are precisely affecting student self-perception and motivation (see Section 3).

More attention to some students could translate into better learning outcomes. Although we cannot identify causality, we can analyze if gender bias is correlated to better tests scores. To do this, we use the SIMCE, which applies to all schools in Chile in different grades. Since fourth graders take the test every year, we have individual test scores for 2012 of the students in our sample classrooms. Table 8 presents the results of simple OLS regressions of SIMCE scores by discipline (reading, math, and science) for girls and boys. As explanatory variables, we use the two measures of gender bias described above (TIA and SIA). In all specifications, we included other covariates to control for student, school, and teacher characteristics as well as the CLASS score to control for quality of teacher–student interactions. Along with the coefficients, we include the normalized coefficient to help interpreting the results.

In the case of the indicator for TIA, results show negative and significant correlations between these gender biases and the test scores for girls in math and science, with a larger coefficient in the case of math; that is, girls whose teachers demonstrate greater gender bias in their interactions with students present lower tests scores in sciences, but particularly in math. Specifically, for every 1 standard deviation of increase in TIA, scores on SIMCE decrease by 0.09 standard deviations in math and by 0.06 standard deviations in science controlling for student, school, and teacher characteristics.

Similarly, SIA or call out is significantly and negatively correlated with test scores only for girls in math and science. Coefficients suggest a similar impact on Science scores and a larger impact on math scores than TIA. In particular, an increase in 1 standard deviation of SIA is associated with a decrease by 0.16 standard deviations in math scores and by a 0.06 standard deviations in science scores.

In sum, results suggest a correlation between gender bias (both from teachers and in terms of student behavior) and students' test scores. Our data do not allow for causality identification, but are consistent with arguments well documented in the literature that point to the importance of teacher–student interaction in terms of better learning. In this case, gender bias seems to be a factor.

Table 8: Correlation between Gender Bias and SIMCE Scores

	Reading		Math		Science	
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Boys	Girls	Boys	Girls	Boys	Girls
TIA						
Coefficient	1.040	-1.991	1.586	-4.060**	-1.275	-2.924**
SE	(1.428)	(1.401)	(1.945)	(2.030)	(1.295)	(1.314)
Normalized coefficient	0.020	-0.041	0.038	-0.091**	-0.027	-0.065**
R-squared	0.101	0.110	0.109	0.175	0.114	0.131
SIA						
Coefficient	-0.646	-0.790	-1.452	-7.842***	0.015	-3.530**
SE	(1.710)	(1.885)	(1.299)	(1.610)	(1.635)	(1.652)
Normalized coefficient	-0.010	-0.012	-0.045	-0.161***	0.000	-0.059**
R-squared	0.100	0.108	0.109	0.189	0.113	0.131
Observations	1,299	1,198	577	560	1,288	1,188

Source: Author's analyses based on primary data and SIMCE data.
*Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. All specifications include covariates: characteristics of students (family income, school performance, household structure, and parents' education); characteristics of the school (type of administration, income decile, and experience and tenure the school of principal); and characteristics of the classroom teacher (total experience, tenure in the school, and whether he or she has or is pursuing a graduate degree).*

7. Conclusions and Policy Implications

This study finds differences in the amount and type of attention teachers devote to girls and boys in the classroom, depending on the subject area. The fact that clear patterns already appear in four grader students is important, given that gender gaps tend to increase as students move up in school grades (Bharadwaj et al., 2015). Many highly able girls and boys may wrongly conclude from such experiences that their abilities are insufficient for careers, respectively, involving math and science or literature and communication. Receiving more teacher attention (positive or negative) has consequences for students in terms of motivation, aspirations, and performance (Bauer 2000; Frawley, 2005; Sadker 1999; Sadker and Sadker, 1994; Streitmatter 1994; Wellhousen and Yin 1997), as well as concerning long-term outcomes, such as decisions about college (Lavy and Sand, 2015) and future employment possibilities and earnings.

The design of policies addressing gender bias should consider that biases are often unconscious; teachers are not necessarily aware of the gender bias they convey in their daily classroom practice (Davis, 2000; Frawley, 2005). Biases are based on myths and beliefs that are not necessarily grounded on hard evidence or even direct experience, but they shape our everyday behavior. For example, there is a general perception that girls talk more in class than boys. In one of their studies, Sadker and Sadker (1985) show a film of a classroom discussion and ask teachers and administrators which gender talked more. The majority of teachers claimed that girls talked more that the boys, but the quantitative data showed that boys talked three times as much.[15]

Given that differences widen by grade level/age—lending more weight to social factors—interventions should start as early as possible. As shown, evidence suggests that the magnitude of gender bias is not as great before sixth grade, but begins to grow more pronounced by the time the students enter high school and even more so by the time they are ready for graduation.

Another important issue for policy is that, because there are short-term and long-term consequences, any strategy to correct for disparities should consider interventions for the flow and the stock. That implies designing mechanisms to correct teacher classroom practices so that entering students are not exposed to bias, but also implementing mechanisms to correct for existing disparities in the labor market and other long-term outcomes. Attitudes and perceptions about children's abilities are hard to change, however. Espinoza, Luz Fontes, and Arms-Chavez (2014) try to reduce high school math teachers' bias (attributing boys' success to ability and girls' success to effort) by providing them training in "an incremental theory of intelligence." The teachers briefly reversed their stereotypes, seeing girls' success as ability-based and boys' as due to effort, but relapsed to their original biased views in less than a year. Thus, efforts should be sustained over time.

Is there a menu of policy options? Specific policy recommendations need further research on the true causes behind the differentiated teachers' attention to boys and girls. Are teachers (maybe unconsciously) acting according to socially or culturally generated stereotypes? Are they simply reacting to more demanding boys (either through unruly behavior or greater participation)? In either case, one obvious recommendation, which aligns with gender bias and quality of teaching literatures, is to hire good teachers. Good teachers will have a better control of the classroom, will be more proactive, and will better guide participation in the classroom. What makes teachers effective? Teachers bring content knowledge, skills, motivation, and

[15] LaFrance and Mayo (1978) find that males talk more in everyday life as well.

classroom practice. Different combinations of these will result in different outcomes: a teacher could have accumulated a great content knowledge and be very motivated, but be ineffective in the classroom. Of the set of a teacher observable characteristics including age, level of formal education, certification, years of experience, and compensation, the evidence suggests that content knowledge (measured by test scores) and absence rates seem to be the most critical factors for learning (Bruns and Luque, 2015; Glewwe et al., 2011). In terms of teacher classroom practice,[16] the use of instructional time (e.g., time spent on academic activities "on task" versus "off task"); the instructional support (e.g., quality of feedback and concept development); the use of materials (including ICT); classroom management (e.g., discipline and organization); the ability to keep students engaged; and the emotional support provided to students (e.g., sensitivity and positive vs. negative climate) seem to be important factors in the equation of learning.

A second policy option is to raise general awareness about these issues among teachers by including exercises and training, and providing concrete examples of "DOs" and "DON'Ts" in the classroom (e.g., avoiding references to role-models or stereotypes, recognizing and awarding collaborative students, integrating students and avoiding gender-based teams, interacting equally with all students, not allowing certain students to control the discussion, and providing effective stimulus and feedback to students). Robinson and Lubienski (2011), for instance, conclude that raising awareness of—and, ideally, eliminating—teachers' biased assessments of boys as better in math (even against the evidence) may go a long way toward closing the gender gap in math. To be effective, training does not necessarily have to be specifically on gender issues. Rather, these issues should be incorporated as part of traditional training on core subjects (e.g., math, science, reading, and writing).

In terms of stimulating broad participation of all students in the classroom, certain pedagogical techniques—for example, teachers calling on the whole classroom, teachers asking for quick answers/rushing the student, and teachers accepting call outs from students even if they have not respected their turn (by raising their hands)—seem to be more problematic for girls. Raising the awareness of these practices should be combined with a review of textbooks that are used as support and reference materials for teachers and students.

[16] Two instruments have been most commonly used in studies of teacher classroom practice: Stallings and CLASS. For full description and application of Stallings instrument see Bruns and Luque (2015); for a full description and application of CLASS, see Pianta et al. (2008). See additional evidence in Bassi, Meghir and Reynoso (2015); Kane et al. (2011); Lavy (2011); Stallings, Johnson and Goodman (1985); and Vieluf et al. (2012).

At the end of the day, the driver of change should be closing the gender gap, whether it favors boys or girls. Any educational policy should take into account impacts and unintended consequences in both groups.

References

AAUW (American Association of University Women). 1992. *How Schools Shortchange Girls.* Annapolis Junction, MD: AAUW.

Araujo, M. C., P. Carneiro, Y. Cruz-Aguayo, and N. Schady. 2016. “Teacher Quality and Learning Outcomes in Kindergarten.” IDB Working Paper Series No. IDB-WP-665. Washington, DC: Inter-American Development Bank.

Banjong, D. N. 2014. “Same Performance but Different Perception: Female Stereotypes in Mathematics Emerge in Fifth Grade.” *International Online Journal of Educational Sciences* 6(2): 258–68.

Bassi, M., M. Busso, and J. S. Muñoz. 2015. "Enrollment, Graduation, and Dropout Rates in Latin America: Is the Glass Half Empty or Half Full?" *Economia journal of the Latin American and Caribbean Economic Association* 16(1): 113-156.

Bassi, M., M. Busso, S. Urzúa, and J. Vargas. 2012. *Disconnected: Skills, Education, and Employment in Latin America*. Washington, DC: Inter-American Development Bank. Available at https://publications.iadb.org/handle/11319/427?locale-attribute=en.

Bassi, M., C. Meghir, and A. Reynoso. 2015. “Teaching Practices and Students’ Learning: An Experiment in Chile.” Unpublished.

Bauer, K. S. 2000. “Promoting Gender Equity in Schools.” *Contemporary Education* 71(2): 22–5.

Bharadwaj, P., G. De Giorgi, D. Hansen, and C. Neilson. 2015. “The Gender Gap in Mathematics: Evidence from a Middle-Income Country.” Staff Report #721. New York, NY: Federal Reserve Bank of New York.

Biraimah, K. 1989. “The Process and Outcomes of Gender Bias in Elementary Schools: A Nigerian Case.” *The Journal of Negro Education* 58(1): 50–67.

Blumberg, R. L. 1984. “A General Theory of Gender Stratification.” *Sociological Theory* 2: 23–101.

-----------. 2007. “Gender Bias in Textbook: A Hidden Obstacle on the Road to Gender Equality in Education.” Paris: UNESCO 2008/ED/EFA/MRT/PI/18. Official Background Paper for 2008 Education for All Global Monitoring Report, *Education for All by 2015 – Will We Make It?* Paris, France: UNESCO.

Available at http://unesdoc.unesco.org/images/0015/001555/155509e.pdf.

----------. 2008. "The Invisible Obstacle to Educational Equality: Gender Bias in Textbooks." *Prospects: A Quarterly Journal of Comparative Education* 38(3): 345–61.

----------. 2015a. *"Gender Bias in Textbooks: Still a Near-Invisible Obstacle to Educational Equality?"* Education for All Global Monitoring Report. Official Background Paper 5.2 for EFA GMR 2015. Paris, France: UNESCO.

----------. 2015b. "'Dry" vs. 'Wet' Development and Women in Three World Regions." *Sociology of Development* 1(1): 91–122.

Bos, M. S., A. Ganimian, and E. Vegas. 2014. "América Latina en PISA 2012: Cómo se desempeñan los varones y las mujeres?" Washington, DC: Inter-American Development Bank. Available at https://publications.iadb.org/handle/11319/6355?scope=123456789/4&thumbnail=true&rpp=5&page=0&group_by=none&etal=0.

Brady, K. L. and R. M. Eisler. 1995. "Gender Bias in the College Classroom: A Critical Review of the Literature and Implications for Future Research." *Journal of Research and Development in Education* 29(1): 9–19.

Bruns, B. and J. Luque. 2015. "Great Teachers: How to Raise Student Learning in Latin America and the Caribbean." Working Paper No. 89514. Washington, DC: World Bank Group. Available at http://documents.worldbank.org/curated/en/2014/01/19798994/great-teachers-raise-student-learning-latin-america-caribbean.

Chetty, R., J. N. Friedman, and J. E. Rockoff. 2011. "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood." NBER Working Paper No. 17699. Cambridge, MA: National Bureau of Economic Research. Available at http://www.nber.org/papers/w17699.

Collins, C. W., J. Kenway, and J. McLeod. 2000. "Factors Influencing the Educational Performance of Males and Females in School and their Initial Destinations after Leaving School." Waurn Ponds, Australia: Deakin University.

Covacevich, C. and G. Quintela-Dávila. 2014. "Desigualdad de género, el currículo oculto en textos escolares chilenos." Technical Note No. IDB-TN-694. Washington, DC: Inter-American Development Bank.

Davis, K. L. 2000. “A Case Study in Gender Equity for Basic Instruction Programs.” *Physical Educator* 57(1): 47.

DeVoe, D. E. 1991. “Teacher Behavior Directed toward Individual Students in Elementary Physical Education.” *Journal of Classroom Interaction* 26(1): 9–14.

Duarte, C. 2010. “Informe final.” Presented at a seminar organized by the School Textbooks Unit, Ministry of Education of Chile, on Universidad de Chile research on “Social Representations of Gender, Generation and Interculturality in Chilean Textbooks.”

Entwisle, D. R., K. L. Alexander, and L. S. Olson. 2007. “Early Schooling: The Handicap of Being Poor and Male.” *Sociology of Education* 80(2): 114–38.

Espinoza, P., A. Luz Fontes, and C. Arms-Chavez. 2014. “Attributional Gender Bias: Teachers’ Ability and Effort Explanations for Students’ Math Performance.” *Social Psychology of Education* 17(1): 105–26.

Fennema, E., P. L. Peterson, T. P. Carpenter, and C. A. Lubinski. 1990. “Teachers’ Attributions and Beliefs about Girls, Boys and Mathematics.” *Educational Studies in Mathematics* 21: 55–69.

Fernández, R. 2013. “Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century.” *American Economic Review* 103(1): 472–500.

Fernández, R. and A. Fogli. 2009. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility." *American Economic Journal: Macroeconomics* 1(1): 146–77.

Fite, G. 2002. “Reading and Math: What is the Connection? A Short Review of the Literature”, *Kansas Science Teacher*, 14: 7–11.

Fogli, A. and L. Veldkamp. 2011. “Nature or Nurture? Learning and the Geography of Female Labor Force Participation.” *Econometrica* 79(4): 1103–38.

Frawley, T. 2005. “Gender Bias in the Classroom.” *Childhood Education* 81(4): 221–27.

Glewwe, P. W., E. A. Hanushek, S. D. Humpage, and R. Renato, 2011, “School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010.” NBER Working Paper 17554. Cambridge, MA: National Bureau of Economic Research.

Guiso, L., F. Monte, P. Sapienza, and L. Zingales. 2008. “Culture, Gender, and Math.” *Science* 320 (May 30): 1164–65 (plus supporting online data).

Harrison, L. E. and S. P. Huntington. 2000. *Culture Matters: How Values Shape Human Progress*. New York, NY: Basic Books.

Heckman, J. J. 2011. "The Economics of Inequality: The Value of Early Childhood Education." *American Educator* (Spring): 31–47.

Inglehart, R. F. and P. Norris. 2003. *Rising Tide: Gender Equality and Cultural Change Around the World*. New York, NY: Cambridge University Press.

Jassey, I. A. 1998. "Gender in Elementary School Texts." *Japan Quarterly* 45: 87–93.

Jones, M. G. 1989a. "Gender Bias in Classroom Interactions." *Contemporary Education* 60(4): 218–22.

----------. 1989b."Gender Issues in Teacher Education." *Journal of Teacher Education* 49(1): 33–8.

Kane T. J., E. S. Taylor, J. H. T., and A. L. Wooten. 2011, "Identifying Effective Classroom Practices Using Student Achievement Data." *Journal of Human Resources*: 587–613.

Lavy, V. 2011. "What Makes an Effective Teacher? Quasi-Experimental Evidence." Paper presented at the NBER 2011 Summer Institute Education Conference.

Lavy, V. and E. Sand. 2015. "On the Origins of Gender Human Capital Gaps: Short and Long Term Consequences of Teachers' Stereotypical Biases." NBER Working Paper 20909. Cambridge, MA: National Bureau of Economic Research.

LaFrance, M. 1991. "School for Scandal: Different Educational Experiences for Females and Males." *Gender & Education* 3(1): 1–13.

LaFrance, M. and C. Mayo. 1978. *Moving Bodies: Nonverbal Communication in Social Relationships.* Monterey, CA: Brooks/Cole.

Li, Q. 1999. "Teachers' Beliefs and Gender Differences in Mathematics: A Review." *Educational Research* 41(1): 63–76.

Magendoza, S. 1970. "Relación entre estereotipos de roles sexuales y libros de enseñanza." In *Educación y realidad socioeconómica*. Mexico, D.F.: Centro de Estudios Educativos.

Martinot, D. and M. Désert. 2007. "Awareness of a Gender Stereotype, Personal Beliefs and Self-Perceptions Regarding Math Ability: When Boys Do Not Surpass Girls." *Social Psychology Of Education* 10(4): 455–71.

Mizala, A. 2014. "Brecha de género, desempeño en matemáticas y elección de carreras." Proyecto FONDECYT. Santiago, Chile: Fondo Nacional de Desarrollo Científico y Tecnológico.

Murnane, R. J., J. B. Willett, and F. Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77: 251–66.

Musau, L. M. and J. A. Migosi. 2013. "Effects of Class Size on Girls' Academic Performance in Science, Mathematics and Technology Subjects." *International Journal of Education, Economics and Dev. lopment* 4(3): 285–95.

Naslund-Hadley, E., J Kwanghee, E. Garcia, and J., Rocha. Forthcoming. "Meta-analysis of the Effects of Early Math and Literacy Programs on Skills." Unpublished

Ñopo, H. 2012. *New Century, Old Disparities: Gender and Ethnic Earnings Gaps in Latin America and the Caribbean*. Washington DC: World Bank Publications.

OECD (Organization for Economic Cooperation and Development). 2012. PISA 2012 Data File. Interative Selection, Variable STO4Q01, Gender. Paris, France: OECD. Available at http://www.oecd.org/pisa/.

----------. 2014a. *Economic Policy Reforms 2014: Going for Growth Interim Report*. Paris. France: OECD Publishing.

----------. 2014b. *Education at a Glance 2014: OECD Indicators.* Paris, France: OECD Publishing.

----------. 2015. *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. Paris, France: OECD Publishing.

Olszewski-Kubilius, P. and D. Turner. 2002. "Gender Differences Among Elementary School-Aged Gifted Students in Achievement, Perceptions of Ability and Subject Preference." *Journal For The Education Of The Gifted* 25(3): 233–68.

Paglin, M. and A. M. Rufolo. 1990. "Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences." *Journal of Labor Economics* 8: 123–44.

Pianta, R. C., A. J. Mashburn, J. T. Downer, B. K. Hamre, and L. Justice. 2008. "Effects of Web Mediated Professional Development Resources on Teacher–Child Interactions in Pre-kindergarten Classrooms." *Early Childhood Research Quarterly* 23(4): 431–51.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73(2): 417–58.

Robinson, J. P. and S. T. Lubienski. 2011. "The Development of Gender Achievement Gaps in Mathematics and Reading During Elementary and Middle School: Examining Direct Cognitive Assessments and Teacher Ratings." *American Educational Research Journal* 48(2): 268–302.

Robinson, J. P., S. T. Lubienski, and Y. Copur. 2011. "The Effects of Teachers' Gender-Stereotypical Expectations on the Development of the Math Gender Gap." Evanston, IL: Society for Research on Educational Effectiveness.

Robinson-Cimpian, J. P., S. T. Lubienski, C. M. Ganley and Y. Copur-Gencturk. 2014. "Teachers' Perceptions of Students' Mathematics Proficiency May Exacerbate Early Gender Gaps in Achievement." *Developmental Psychology* 50(4): 1262–81.

Sadker, D. 1999. "Gender Equity: Still Knocking at the Classroom Door." *Educational Leadership* 56(7): 22–7.

Sadker, M. and D. Sadker. 1985. "Sexism in the Schoolroom of the 80s." *Psychology Today* March: 54–57

----------. 1994. *Failing at Fairness: How America's Schools Cheat Girls*. New York, NY: Scribner's Sons/MacMillan Publishing Co.

Sadker, M., D. Sadker, and L. M. Stulberg. 1993. "Fair and Square: Creating a Nonsexist Classroom." *Instructor* 102(7): 44–46, 67–68.

Schwartz, C. S. and R. Sinicrope. 2013. "Prospective Elementary Teachers' Perceptions of Gender Differences in Children's Attitudes Toward Mathematics." *School Science and Mathematics* 113(8): 410–16.

Shakeshift, C. 1995. "Reforming Science Education to Include Girls." *Theory Into Practice* 34(1): 74–79.

Shumow, L. and J. A. Schmidt. 2013. "Academic Grades and Motivation in High School Science Classrooms Among Male and Female Students: Associations with Teachers' Characteristics, Beliefs and Practices." *Journal of Education Research* 7(1): 53–71.

Stallings, J., R. Johnson, and J. Goodman, 1985. "Engaged Rates: Does Grade Level Make a Difference?" Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois. (Eric Document No. 267 018.)

Straus, B. “Boys and Literacy.” Research Apprenticeship Report. Charlottesville, VA: University of Virginia.

Streitmatter, J. 1994. *Toward Gender Equity in the Classroom: Everyday Teachers' Beliefs and Practices.* Albany, NY: State University of New York Press.

Teese, R. DAvies, M. Charlton, M. and Polesel, J. 1995. *Who Wins at School? Boys and Girls in Australian Secondary Education,* Department of Education, Policy and Management: University of Melbourne.

Tiedemann, J. 2002. “Teachers' Gender Stereotypes as Determinants of Teacher Perceptions in Elementary School Mathematics.” *Educational Studies in Mathematics* 50: 49–62.

Trecker, J. L. 1971. “Women in U.S. History High School Textbooks.” *Social Education* 34: 249–60.

United Nations. 2013. “Enrollment in Tertiary Education.” UNData. Montreal, Canada: United Nations.

Uren, M. B. 1971. “The Image of Women in Textbooks.” In V. Gornick and B. K. Moran (editors), *Woman in Sexist Society: Studies in Power and Powerlessness*. New York, NY: Signet/New American Library.

Van Duzer, E. 2006. “Gender Bias in Predictions of Boys' Academic Failure.” *Action in Teacher Education*. 28(1): 94–8.

Vichit-Vadakan, J. 2014. “Gender Analysis of School Textbooks: Conclusions and Recommendations: Report presented to the Rockefeller Foundation.” Bangkok: National Institute for Development Administration (NIDA).

----------. 2015. “National Survey of Gender Views: Thai Kindergarten, Primary, Secondary and Vocational Teachers and Upper Secondary Students: Report to the World Bank.” Bangkok: National Institute for Development Administration (NIDA).

WEF (World Economic Forum). 2015. *The Global Gender Gap Report*. Geneva, Switzerland: WEF.

Wellhousen, K. and Z. Yin. 1997. “Peter Pan Isn't a Girl's Part: An Investigation of Gender Bias in a Kindergarten Classroom.” *Women & Language* 29(2): 35-41.