Children's Gender Stereotypes About Math: The Role of Stereotype Stratification[1]

JENNIFER STEELE[2]
Harvard University

Women underperform on standardized math tests compared to men. However, girls perform as well if not better than boys in math. In this paper, it is proposed that this discrepancy may be explained in part by a process of stereotype stratification, in which targets of a negative in-group stereotype view themselves as a member of a subgroup to which the stereotype does not apply. Two experiments with elementary-school children provide support for this theory. In Study 1, girls placed advanced math pictures with males more often than basic math pictures. In addition, girls rated men as liking and as being better at math than women, but viewed boys and girls as being equal on these variables. In Study 2, girls were more likely to draw a man when told a story about an adult mathematician, but were more likely to draw a girl when told of a child mathematician. The social and educational implications of these findings are discussed.

> If the cure for cancer is forming in the mind of one of our daughters, it is less likely to become a reality than if it is forming in the mind of one of our sons. Until this changes, everybody loses. (Sadker & Sadker, 1994, p. 14)

For decades, researchers have been concerned with a serious and pervasive gender discrepancy: Women are less likely than men to pursue careers in math and science (Betz, 1997). For example, Betz reported that women receive less than 17% of the doctoral degrees awarded in math and in the physical sciences. In addition, Steele (1997) noted that women "occupy only 10% of the jobs in physical science, math, and engineering" (p. 615). This problem, sometimes referred to as the *math and science pipeline* (American Association of University Women [AAUW], 1999), has serious implications for the future of every girl who may be considering the possibility of becoming a mathematician or a scientist.

[1]This research was supported by a doctoral fellowship from the Social Sciences and Humanities Research Council of Canada and a Graduate Society Summer Fellowship from Harvard University. I would like to thank Shawn Badali, Dave King, Casey Moeller, Liz Moriarty-Ambrozaitis, and Ted Mussman for their help with data collection.

[2]Correspondence concerning this article should be addressed to Jennifer Steele, who is now at the Department of Psychology, University of Waterloo, 200 University Avenue W., Waterloo, ON N2L 3G1, Canada. E-mail: jsteele@watarts.uwaterloo.ca

Journal of Applied Social Psychology, 2003, **33**, 12, pp. 2587-2606.

Women's underrepresentation in the field of mathematics has been of particular interest to numerous social, developmental, and educational psychologists. This is not surprising, in light of the fact that math is perhaps the most consequential field of study for any individual who hopes to pursue education beyond the high school years. Standardized tests including the Scholastic Achievement Test (SAT), Graduate Record Examination (GRE), and the Medical College Admissions Test (MCAT) all have challenging mathematical components that serve as gatekeepers to the land of elite colleges and universities. Even in more applied arenas of social work and education, courses in statistics often serve as the critical criterion in determining who will be allowed to pursue these fields of study. For this reason, math has been described as a critical filter for young people who are attempting to pass into lucrative and prestigious occupations (Hyde, Fennema, Ryan, Frost, & Hopp, 1990).

Research has shown that many of the gender differences in math ability do not emerge until the high school and college years. In fact, a meta-analysis by Hyde, Fennema, and Lamon (1990) of 100 studies revealed that in elementary and middle school, girls exhibit a slight superiority in math performance, compared to boys. However, the advantage enjoyed by girls in the early years disappears in high school, college, and beyond, with men performing moderately better than women in mathematics. Although Hyde, Fennema, and Lamon concluded that gender differences in math performance have declined over the past 30 years, it is obvious that they still exist. This is perhaps best illustrated by the gender discrepancy that remains in math SAT scores, where, in 1997, men scored an average of 30 points higher than women (AAUW, 1999).

Possible Explanations for the Math Pipeline

Numerous explanations have been proposed to account for the gender disparities observed in mathematics. One early theory that has received limited empirical support is the genetic explanation proposed by Benbow and Stanley (1980). These researchers found gender differences favoring boys in the math performance of junior high school students who had all received comparable formal education. From their data, Benbow and Stanley expressed a preference for "the hypothesis that sex differences in achievement in and attitude toward mathematics result from superior male mathematical ability, which may in turn be related to greater male ability in spatial tasks" (p. 1264).

Other researchers have identified a host of social factors that appear to contribute to women's willingness and ability to enter into math-related fields of study. For instance, according to Eccles (1987), women's decision not to pursue mathematics stems from a variety of factors, including their relative expectation for success in the field as compared to others, and the relative amount of effort they expect to exert in order to be successful. Specifically, through path analysis, Eccles and Jacobs (1986) concluded that "grades and plans to continue taking mathematics are predicted most directly by the students' . . . estimates of their mathematical abilities, their perceptions of the value of mathematics courses, and their levels of math anxiety" (p. 374). Such perceptions and expectations also may be highly influenced by personal lifestyle choices (i.e., how easy it will be to integrate family and career), as well as the extent to which they perceive or expect to experience gender discrimination (Steele, James, & Barnett, 2002).

The Effect of Stereotypes

More recent explanations for the underrepresentation of women in mathematics have centered on the role of stereotypes. For example, Jacobs and Eccles (1992) examined the effect of mothers' gender stereotypes and expectations on children's self-appraised ability in math. In a study of 1,500 mothers and their 11- and 12-year-old children, the authors found that mothers' stereotypes influenced their expectations for their children, which, in turn, affected children's self-appraised ability in mathematics. The researchers note that such gender stereotypes and expectations can lead to self-fulfilling prophecies, in which children unwittingly confirm the stereotypes and expectations of their mothers.

Theory and research on stereotype threat has provided another explanation for why women may not remain identified with the field of mathematics to the same extent as men (Aronson, Quinn, & Spencer, 1998; Brown & Josephs, 1999; Inzlicht & Ben-Zeev, 2000; Spencer, Steele, & Quinn, 1999; Steele, 1997). According to Steele, to succeed in school, "one must be identified with school achievement in the sense of its being a part of one's self-definition" (p. 613). Steele further argued that this type of identification might be difficult for women to retain in math-related fields because there exists a societal stereotype that devalues them in these domains. The effects of stereotype threat among women in math have been demonstrated most aptly in a study conducted by Spencer et al.

In one of Spencer et al.'s (1999) experiments, male and female college students were asked to take a challenging math test. Students in one condition were told that the test showed no gender differences, while those in the other condition were told that the test generally showed gender differences. The purpose of this manipulation was to decrease or increase, respectively, the salience of the negative gender stereotype about women and math. Consistent with the theory of stereotype threat, women's math test performance was worse than men's when the test was described as showing gender differences, but not when the test was described as showing no gender differences.

This study by Spencer et al. (1999) demonstrates how negative gender stereotypes can not only lead women to disidentify with the field of math, but also how such negative gender stereotypes can impair women's performance in math. As

Steele (1997) explained, "where bad stereotypes about these groups apply, members of these groups can fear being reduced to that stereotype. And for those who identify with the domain to which the stereotype is relevant, this predicament can be self-threatening" (p. 614). In order for the effects of stereotype threat to occur, however, it is necessary for participants to be aware of a negative gender stereotype about females and math. After all, women in the threat condition in Spencer et al.'s experiment were not told that women generally perform worse on the test, but were told simply that a gender difference existed and were left to draw their own conclusions.

If stereotypes and stereotype threat contribute to women's attitudes toward and performance in mathematics, why are girls not underperforming in mathematics? It seems logical that girls would suffer the same consequences of being stereotyped as women. One possible reason for this disparity is that elementary-school-aged girls are completely unaware of a gender stereotype about math. However, given the literature on gender-role development, this explanation does not seem plausible. From a very young age, children are highly sensitive to gender-relevant information (Eisenberg, Martin, & Fabes, 1996). By the age of 3 years, children show a preference for gender-typed toys; by the age of 5 years, children express a preference for gender-stereotype-consistent occupations. Indeed, between about the ages of 3 and 6 years, children are not only able to identify culturally based gender stereotypes, but they also endorse them quite readily (Golombok & Fivush, 1994).

Stereotype Stratification

The theory of stereotype stratification provides an alternative explanation for why girls may be unaffected by this negative gender stereotype about math ability. The term *stereotype stratification* is used to refer to the process of cognitively viewing oneself as a member of a subgroup to which the stereotype does not apply. For example, when presented with stereotypical information about females and mathematics, girls may view the stereotype as being true for women (i.e., the age subgroup to which they do not belong) but not true for girls (i.e., the age subgroup to which they do belong). In essence, it is proposed that girls develop a gender stereotype about men's and women's ability in mathematics, but do not apply this same belief to boys and girls.

By stratifying the stereotype along an age dimension, girls may be in a position to develop specific beliefs about men's superiority in mathematics, without developing a comparable belief about boys' math ability. In addition, because children learn basic math, whereas adults learn more advanced math, girls may express their stereotype by associating basic math with either males or females, while simultaneously associating advanced math more with males.

Subtyping and Subgrouping

Consistent with the notion of stereotype stratification, research has shown that when people are faced with exceptions to an out-group stereotype, they will recategorize the exceptions, as opposed to changing the stereotype entirely (Kunda & Oleson, 1995; Weber & Crocker, 1984). Allport (1954) was perhaps the first to recognize people's capacity to maintain a stereotype in the face of disconfirming evidence. He noted:

> There is a common mental device that permits people to hold to prejudgments even in the face of much contradictory evidence. It is the device of admitting exceptions . . . When a fact cannot fit into a mental field, the exception is acknowledged, but the field is hastily fenced in again and not allowed to remain dangerously open. (p. 23)

Since then, researchers have identified two distinct mechanisms that demonstrate people's ability to subdivide larger groups: subtyping and subgrouping. When faced with exceptions to out-group stereotypes, researchers have found that there is a general tendency to subtype the exceptions. *Subtyping* is "the process by which group members who disconfirm, or are at odds with, the group stereotype are mentally clustered together and essentially set aside as 'exceptions to the rule'" (Maurer, Park, & Rothbart, 1995, p. 812). Weber and Crocker (1983) found that when stereotype-disconfirming evidence was concentrated among a group of individuals, raters would subtype these members into their own category, while maintaining the more global stereotype. Similarly, Kunda and Oleson (1995) found that participants had a strong tendency to use neutral attributes as a dimension for subtyping in an attempt to explain away stereotype-disconfirming individuals.

The literature on subtyping provides some evidence that people are willing to "re-fence" individuals who behave in a manner that is contradictory to an out-group stereotype. But what about within groups? Subgrouping is a process that generally occurs among members of both in-groups and out-groups. A *subgroup* is an identified cluster of similar people within a larger group. For example, Brewer and her colleagues (Brewer, Dull, & Lui, 1981; Brewer & Lui, 1984) have shown that both young and older people have subgroups of the elderly, such as *grandmothers*, *elder statesmen*, or *senior citizens*. Once a subgroup has been formed, Brewer et al. have suggested that the category will be "represented cognitively by prototypes—actual or imaginary instances of the category that contain attributes most representative of items outside the category" (p. 656).

In short, the proposed process of stereotype stratification incorporates aspects of both subtyping and subgrouping. *Subtyping* occurs when a person recategorizes an individual, or group of individuals, into a smaller cluster because they are exceptions to an out-group stereotype. *Subgrouping* occurs when a person identifies smaller clusters of people (e.g., businesswomen, feminists, homemakers) within a larger group (women) that is either an in-group or an out-group. Stereotype stratification, by contrast, can only occur when a person is confronted with a negative in-group stereotype. In response, the person may use one of the subgroups to which he or she belongs as a means of subtyping himself or herself away from the stereotype. In the case of the gender stereotype about females and mathematics, girls may apply the stereotype to women but not girls, or to advanced math but not basic math.

The Present Research

The purpose of the present research is to examine the process of stereotype stratification by investigating children's gender stereotypes about math. In Study 1, girls' gender stereotypes about mathematics were examined to determine whether they would (a) be specific to women but not to girls; and (b) apply to advanced math, but not basic math. Study 2 was designed to investigate whether this process of stereotype stratification would emerge in boys' and girls' prototypes of a mathematician.

Study 1

In the first study, the development of a gender stereotype about mathematics was investigated among first- through fourth-grade girls, using indirect and direct measures. Recent findings have suggested that people are often unwilling or unable to express the stereotypes they hold (Wegner & Bargh, 1998). Accordingly, researchers have begun to study stereotyping using implicit (or indirect) and explicit (or direct) measures (Banaji & Greenwald, 1994; Greenwald & Banaji, 1995). In this study, indirect and direct measures were used to examine two different expressions of stereotype stratification.

In order to assess gender stereotypes more implicitly, girls were first asked to sort pictures of people by gender, and were then asked to place pictures depicting advanced or basic math problems and solutions in either the male or the female pile. By asking girls to first categorize pictures by gender, they were put into a forced-choice situation that allows us to focus specifically on their gender-stereotype knowledge. The purpose of this task is (a) to determine whether girls are aware of a gender stereotype about math by assessing whether girls generally associate math with males more than with females; and (b) to determine whether this is particularly true for pictures of advanced math and not of basic math.

> *Hypothesis 1.* Girls will group math-related pictures with males more often than with females.

> *Hypothesis 2*. Consistent with the theory of stereotype stratification, girls will place pictures depicting advanced math with males more often than pictures depicting basic math.

To assess girls' stereotype more directly, an explicit measure was designed to examine whether girls hold a stereotype about women and math that they do not apply to girls.

> *Hypothesis 3*. Girls will view men as having a greater interest and ability in math than women.

> *Hypothesis 4*. Consistent with stereotype stratification, boys and girls will be perceived as having a comparable interest and ability in math.

Method

Participants

A total of 42 female elementary school students participated in this study. Most were White, middle- to upper-class, and all were enrolled in a public school located in a wealthy suburb of Boston. The girls were in first ($n = 9$), second ($n = 13$), third ($n = 11$), and fourth ($n = 9$) grades. Consent for the girls' participation was obtained from their parents, teachers, principal, and school board; and the study was conducted on school premises, before and during regular school hours.

Procedure

Each girl took part in an individual 20-min testing session conducted by a male ($n = 17$) or a female ($n = 25$) experimenter. A second experimenter of the opposite gender remained in the room to operate the video camera and to observe their responses. Once seated, each girl had her photograph taken and was asked to choose either a math activity book or an English activity book as a token of appreciation for her participation.

Indirect measure. Each girl was first asked to complete an indirect measure designed to assess her gender stereotypes about advanced math versus basic math. The procedure, which expanded on a methodology used by Kagan (1964) and Mullen (1990), consisted of a two-part picture-sorting task.

In the training phase, girls were taught to sort pictures by gender. Nonsense syllables ("DEP" and "GIB") were used as grouping labels in order to allow children cognitive flexibility when placing the testing items. At first, they were shown four pictures (a man, a woman, a boy, and a girl, respectively), and they were shown that the man and the boy belonged in one group (labeled "DEP" or "GIB" alternately), and that the woman and girl belonged in the other group (labeled "GIB" or "DEP" alternately). The girls were then shown 12 additional pictures of racially diverse men (3), women (3), boys (3), and girls (3) and were asked to classify each as either a "DEP" or a "GIB." They were given corrections for pictures that were sorted incorrectly. When necessary, this sorting task was repeated up to two times, until the experimenter felt that the girls understood the basis for sorting the two groups of pictures.

Upon successful completion of the training portion of the indirect measure, each girl began the testing phase. Girls were told that the experimenter would not be looking at the next set of pictures, but would instead record their responses by looking at a number on the back of the card. They were further told,

> I will show you a card, and I want you to tell me whether it is a GIB or a DEP. If you're not sure of your answer, that's okay. There are no right or wrong answers for these pictures. I just want to know what group you think each goes with the best.

The test pictures consisted of eight pictures depicting math[3] and eight gender-neutral pictures.[4] The eight math pictures consisted of four pictures depicting basic math problems and solutions and four pictures depicting advanced math problems and solutions. The 16 testing pictures were presented to the girls one at a time in random order, and the number of pictures that the girls placed in each pile was recorded.

Direct measure. Following the first task, the girls completed a more explicit task designed to measure her awareness of a stereotype about females and math. This measure consists of two subscales. The interest subscale is designed to assess the extent to which members of specific groups are perceived as liking math. Girls were shown five stick faces with different expressions (a big smile, a

[3]To verify that the chosen math pictures were representative of math from the perspective of the child, 5 of the participants completed an additional task before returning to class. Each child was individually shown the original 16 testing pictures and was asked to pick out the math pictures. All 5 children correctly chose the eight math pictures used in the experiment.

[4]Gender-neutral pictures were chosen based on ratings from 48 racially diverse Harvard University undergraduates (31 females, 17 males). Participants viewed a total of 25 pictures for 5 s each. All of these pictures had been judged by the author to be gender-neutral. Roughly half of the participants ($n = 21$) were asked to state whether each picture was male or female. The other half ($n = 27$) were asked to rate, on a 7-point scale ranging from 1 (*men/boys*) to 7 (*women/girls*), whether each picture was more male or more female. The eight pictures used in the present experiment were chosen because the mean on the scale hovered most closely around the middle value of 4 (*M*s ranged from 3.74 to 4.30), and roughly an equal number of participants labeled each picture as male and as female. Pictures included party hats, a light bulb, musical notes, a mug, pencils and tape, a bell, stars, and a piano.

small smile, no smile or frown, a small frown, or a big frown) and were asked to point to the face that best represented what they thought in response to seven questions. The first three questions were practice questions, designed to give the girls experience with the interest subscale. Upon completion of the three practice questions, the girls were told that they would now be asked questions about math and that there were no right or wrong answers for these questions. The four test questions were "How much do most girls like math?"; "How much do most boys like math?"; "How much do most men like math?"; and "How much do most women like math?"

The girls also completed an ability subscale designed to assess the degree to which members of specific groups are perceived as being good at math. They were asked to use a rating scale ranging from *very good* (full tube) to *sort of good but sort of not good* (half full tube) to *not very good* (empty tube), and to point to the tube that best represented what they thought in response to three practice questions and four test questions. The four test questions were "How good are most boys at math?"; "How good are most girls at math?"; "How good are most women at math?"; and "How good are most men at math?"[5]

Results and Discussion

Excluded Data

Of the 42 girls, 4 failed to recognize that the pictures were to be grouped by gender during the training phase of the implicit task. Accordingly, they were excluded from the analyses that follow. Hence, the final sample consists of 38 girls in first (n = 8), second (n = 11), third (n = 11), and fourth (n = 8) grade.

Indirect Measure

Hypothesis 1 stated that girls would place math-related pictures with males more often than with females. In order to test this hypothesis, the number of math pictures that each girl placed in the pile of female pictures was first counted as a continuous variable. The girls were given a Total Math score ranging from 0 (*not at all female*) to 8 (*completely female*) that reflects the total number of math pictures they had placed in the female pile versus the male pile. A one-sample t test was then performed on the girls' Total Math scores. If the girls had been randomly placing a math-related picture in either the male pile or the female pile, then the average Total Math score should not have differed from 4. The one-sample t test was reliable, $t(37) = -3.81$, $p = .001$, indicating that, as predicted,

[5]Children in Study 1 also completed a math test as part of the procedure. However, as these results are not the focus of the current paper, they have not been included. Please contact the author for further details.

Table 1

Explicit Ratings by Girls for the Interest and Ability Subscales: Study 1

	How much do most _____ like math?		How good are most _____ at math?	
	M	*SD*	*M*	*SD*
Men	4.13	0.74	4.50	0.56
Women	3.87	0.84	4.29	0.61
Boys	3.55	1.20	4.00	0.81
Girls	3.64	0.85	3.92	0.63

Note. Explicit ratings of interest and ability were rated on a 5-point scale ranging from 1 to 5, with higher scores indicating greater perceived interest or ability in math.

girls grouped math-related pictures with males more often than with females (M = 2.87, SD = 1.83). This suggests that girls are generally aware of a gender stereotype about mathematics.

Hypothesis 2 concerned the process of stereotype stratification. It was predicted that girls would place pictures of advanced math in the male pile more often than pictures of basic math. In order to determine whether this was the case, two continuous scores were created: an Advanced Math score and a Basic Math score. Each score ranges from 0 (*not at all female*) to 4 (*completely female*), and reflects the number of advanced or basic math pictures, respectively, that the girls had placed in the female pile. The results of a paired t test provide some initial support for stereotype stratification. As hypothesized, girls were marginally less likely to group pictures depicting advanced math (M = 1.29, SD = 1.04) with females than they were to group pictures depicting basic math (M = 1.58, SD = 1.11) with females, $t(37) = -1.60, p = .12$.

Direct Measure

Hypothesis 3 stated that, consistent with the theory of stereotype stratification, girls would view men as liking math more, and as being better at math, compared to women. In addition, it was hypothesized that girls would view boys and girls as being equal in math interest and ability. Paired t tests were conducted to test this prediction, and Table 1 displays the mean ratings on this direct measure.

As expected, girls rated men (M = 4.13, SD = 0.74) as liking math more than women (M = 3.87, SD = 0.84), $t(37) = 2.37, p = .02$); and they rated men (M = 4.50, SD = 0.56) as being better at math than women (M = 4.29, SD = 0.61), $t(37) = 2.74, p = .01$. By contrast, girls viewed boys (M = 3.55, SD = 1.20) and girls (M = 3.64, SD = 0.85) as liking math to the same extent, $t(37) = -0.36$, p = .72; and they viewed boys (M = 4.00, SD = 0.81) and girls (M = 3.92, SD = 0.63) as being equally good at math, $t(37) = 0.52$, $p = .61$.

Finally, correlations between the card sorting (indirect measure) and questions about liking and ability (direct measure) reveal only a small positive relationship (rs < .29, ps > .09). Consistent with other research on implicit and explicit beliefs, there was very little support for a strong relationship between the two. In addition, the data provide little evidence of a consistent age trend.

To summarize, the results of Study 1 provide initial support for the theory of stereotype stratification. Girls expressed gender-stereotype-consistent views for women, but not for girls. On the indirect measure, girls associated advanced math with males more than basic math. Explicitly, girls viewed men as liking math more and as being better at math than were women. However, consistent with stereotype stratification, these same girls rated boys and girls as liking and as being good at math to the same extent. In short, the first study provides preliminary evidence that girls show a stronger association between advanced math and males than basic math and males. In addition, the first study provides support for the development of a negative stereotype about women and math, but not about girls and math.

These findings are limited for two reasons. First, the indirect measure used in Study 1 forced girls to attend to gender as a dimension of categorization, and hence might have elicited particularly stereotypical responses. Similarly, these findings might have emerged in part because of the potential ease with which girls could guess the purpose of the task. Second, because pictures of people from diverse ages were included in the training phase of the indirect measure, it is difficult to tell whether girls' responses were a result of their beliefs about men and women, girls and boys, or some combination of them all. These limitations are addressed in Study 2.

Study 2

The purpose of Study 2 is to examine the gender of drawings provided in response to a story about either a child or an adult who was really good at math. According to the theory of stereotype stratification, children's prototype for an adult mathematician should be male. Yet, the prototype for a child mathematician should not show this same gender bias. Consequently, the following is hypothesized:

> *Hypothesis 5.* Children will be most likely to draw a man when asked to draw an adult mathematician.

By contrast,

> *Hypothesis 6.* Children will be most likely to draw a character of their same gender when asked to draw a child mathematician.

In the absence of a gender stereotype relevant to children and math, it is believed that the gender of the drawings will be influenced by an in-group bias on the part of the children.

> *Hypothesis 7.* The sex of girls' drawings will not be affected by condition in another academic domain such as spelling.

Method

Participants

A total of 58 children (32 girls, 26 boys) who were attending an after-school program participated in this study. Children ranged in age from 6 to 10 years, with a mean age of 7.65 years (SD = 1.18). The majority of participants were White and middle- to upper-class. Consent for the children's participation was obtained from their parents and the after-school program director.

Procedure

Children were individually escorted into the testing room by either a male (n = 29) or a female (n = 29) experimenter. This experimenter remained in the room to work the video equipment, while the other experimenter tested the child. Once the child had been introduced to both experimenters, the child was told that he or she would be asked to draw a picture. The child was asked to choose one marker from a small bag. Pink and blue markers had been removed to help ensure that the color chosen did dictate the gender of the drawing.

Children were then informed that they would listen to an audiotape of a story two times and would then be asked to "draw a picture of the person in the story." Children were asked to listen very carefully, and then a pre-recorded story was played on a tape recorder. When the story had been played twice, children were told that they could begin drawing on the folded sheet of paper in front of them.

When the children finished the drawing, the sheet was turned over, and they were informed that they would now hear a different story two times. Again, each child was told that he or she would be asked to draw a picture of the character in the story, and he or she was given the opportunity to select a new marker. After the second story had been played twice, the children were told that they could begin drawing.

Stories. Each child was presented, in random order, a story about a mathematican and a speller. These stories were designed to describe a person who excelled at and enjoyed math or spelling, respectively. Prior to entering the testing area, children were randomly assigned to either a Child or an Adult condition that differed only in terms of the age of the character described. For children in the Child condition, the character described in both stories was a child, whereas for those in the Adult condition, the character described in both stories was an adult.

The following is the mathematician story, heard by children in the Child condition:

> This is a story about a child who is really good at math. This child is always the first to finish every math problem, no matter how hard. And this child also really likes doing math. If there is a math problem to be done, this kid is the one to do it. This child is a really great mathematician.

The story for children in the Adult condition was identical, with the only exception being that the words "child" and "kid" were replaced by the words "grown-up" and "adult," respectively. Likewise, the speller story was identical to the story of the mathematician, except the words "math" and "mathematician" were replaced by the words "spell" and "speller." The stories were prerecorded by both a man and a woman, and the gender of the voice on the audiotape was counterbalanced across children's gender, condition, and grade.

Gender of character. The dependent measure was the gender of the character drawn by the child. In order to assess the gender of the character, a series of steps was taken. First, the experimenter attempted to determine the gender of the drawing indirectly. That is, prior to asking the children any questions, the experimenter looked at their two drawings and recorded whether the character looked male, female, or was undistinguishable. Next, the experimenter pointed to the first picture and asked the children to tell them something about the character. This procedure was repeated for the second drawing, and the gender of the pronouns used was recorded. Each child was then asked to provide a name for the character in each of the drawings; the name was recorded, and the experimenter assessed the gender of the name. Finally, the children were asked for each drawing "Is the person in this picture a girl (woman) or a boy (man)?"

Manipulation checks. After the gender of each character had been assessed, children were asked two questions about each drawing that served as manipulation checks. First, they were asked "What was the person in this picture really good at?" Next, they were asked "Was this person a grown-up or a child?"

Results and Discussion

Excluded Data

Data were excluded from the following analyses for one of three reasons. First, data were omitted if children failed to correctly answer the manipulation

Table 2

Percentage of Boys and Girls Who Drew a Male Mathematician or Speller According to the Age of the Story's Character: Study 2

	Boys		Girls	
Age of story's character	Male mathematician	Male speller	Male mathematician	Male speller
Adult	92% (13)	82% (11)	64% (14)	57% (14)
Child	90% (10)	70% (10)	31% (13)	57% (14)

Note. Numbers in parentheses are the actual number of participants out of 48 (less excluded data) who drew a male mathematician or speller.

checks. This was the case for the spelling story of 3 children (2 girls, 1 boy), the math story of 2 girls, and both stories for 1 girl.

In addition, children's data were excluded if it was impossible to determine the gender of the drawing from their responses and from the ratings of an independent rater. This was the case for 3 children (2 boys, 1 girl) in the child condition and 2 children (1 boy, 1 girl) in the adult condition. Finally, one boy's data were omitted because it was determined from the videotapes that he believed he was supposed to draw the person who told the story and not the character described in the story.

Mathematician Drawings

Table 2 presents the percentage of boys and girls who drew a male character in response to each story. In order to examine the hypotheses, chi-square analyses were performed. Across all children, there was a general bias toward drawing males (68%) over females (32%) in response to mathematican story, $\chi^2(1, N = 50) = 6.48, p = .01$. However, as can be seen in Figure 1, this was particularly true for boys (male mathematician drawings = 91%; female mathematician drawings = 9%), $\chi^2(1, N = 23) = 15.70, p < .001$; as opposed to girls (male mathematician drawings = 52%; female mathematician drawings = 48%), $\chi^2(1, N = 27) = 0.04, p = .85$.

As stated in Hypotheses 5 and 6, boys were just as likely to draw a male, regardless of whether the mathematician in the story was an adult (92%) or a child (90%), $\chi^2(1, N = 23) = 0.04, p = .85$. However, a 2 × 2 chi-square analysis reveals that girls were influenced by the age of the story's character, $\chi^2(1, N = 27) = 3.03, p = .08$. Consistent with the theory of stereotype stratification, girls were more likely to draw a girl (69%) than a boy (31%) when the mathematician

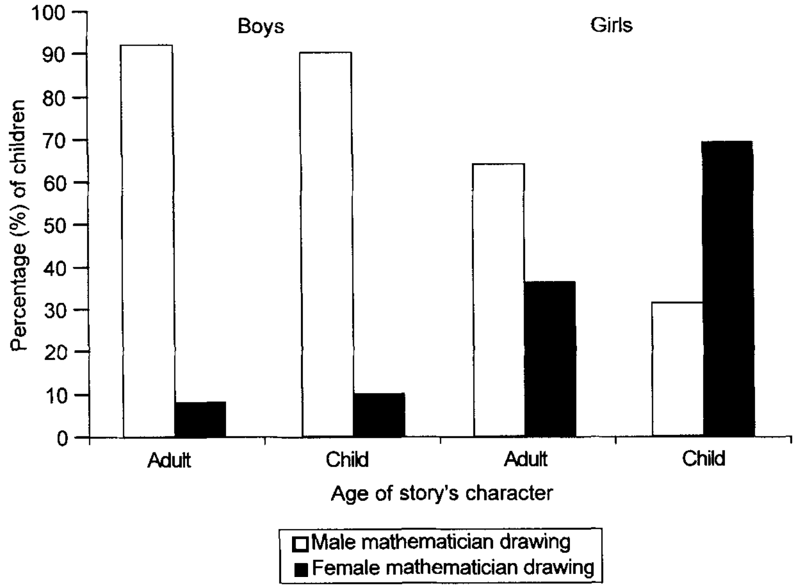

Figure 1. Percentage of boys and girls who drew a male or female mathematician according to the age of the story's character (Study 2).

was a child, but were more likely to draw a man (64%) than a woman (36%) when the story's character featured an adult (Figure 1).

Speller Drawings

According to Hypothesis 7, the results obtained for the mathematician story should not extend to the speller story. In order to verify that this was the case, the same analyses were performed on children's drawings in response to the speller story. Once again, the children had a general tendency to draw males (65%) over females (35%), $\chi^2(1, N = 49) = 4.59, p = .03$; and this tendency was primarily a result of boys' drawings (male speller drawings = 76%; female speller drawings = 24%), $\chi^2(1, N = 21) = 5.76, p = .02$, and not girls' drawings (male speller drawings = 57%; female speller drawings = 43%), $\chi^2(1, N = 28) = 0.57, p = .45$. However, as can be seen in Table 2, the age of the story's character did not have a significant effect on either boys' or girls' tendency to draw a male or a female speller, $\chi^2(1, N = 23) = 0.40, p = .53$, and $\chi^2(1, N = 28) = 0.00$, *ns*, respectively. This provides some initial indication that children's prototype for an adult or child mathematician is not the same as it is in another domain, such as spelling.

General Discussion

Women continue to be underrepresented and to underperform in mathematics throughout the college years and beyond. However, girls are performing as well as if not better than boys in elementary school mathematics. Although numerous factors contribute to this discrepancy, one specific possible explanation was examined through the present experiments. It was proposed that girls develop a gender stereotype about mathematics that is specific to women (i.e., "Men are better than women at math"), not girls (i.e., "Girls and boys are equally good at math"). This process of psychologically placing oneself in a subgroup to which the stereotype does not apply has been termed *stereotype stratification.*

Indeed, preliminary support for these hypotheses was received through two experiments. In Study 1, girls were more likely to group pictures depicting advanced math with pictures of males, as opposed to females. In addition, girls were more likely to rate men as being better at math and as liking math more than women. However, consistent with the hypotheses, girls rated boys and girls as being equal in their interest and ability in this domain. In Study 2, children's prototypical drawing of a mathematician was examined. Boys' prototype of a mathematician was consistently male, regardless of whether the story's character was an adult or a child. By contrast, in line with stereotype stratification, girls' prototype was more likely to be male when an adult mathematician was described and a girl when a child mathematician was described.

Alternative explanations for these findings exist, however.[6] Based on the notion that stereotypes reflect a kernel of truth, it is possible that these data reflect girls' knowledge of the reality discussed in the introduction, that girls generally do at least as well as boys in math, but that men are more prevalent (and hence perhaps more skilled) in the field of mathematics than are women. It is interesting to note, however, that the data from Study 2 indicate that boys did not share this cultural knowledge. It is possible that girls are more aware of this reality than are boys because it is beneficial for them in the short term. By attending to differences by age, girls are able to justifiably develop the stereotype in a way that it does not apply to them. This possibility is consistent with the theory of stereotype stratification.

Another possible explanation is that these findings reflect a strong in-group bias among children at this age. For example, Powlishta (1995) found evidence of in-group favoritism by gender among young participants in middle childhood. When viewing videotapes of various unknown children, girls predicted liking female targets more than male targets, whereas the reverse pattern was found for boys. In addition, both boys and girls rated the targets in a gender-stereotypical

[6]I would like to thank an anonymous reviewer for pointing out some of these alternative explanations.

manner. Female targets were assumed to be more feminine, and male targets were assumed to be more masculine.

In-group favoritism could help to explain some of the results from Studies 1 and 2. However, Powlishta's (1995) findings do not explain why girls did not show in-group favoritism when describing the abilities and interests of women in mathematics (Study 1) or when they depicted adult mathematicians (Study 2). Women could easily be perceived as in-group members for girls, just as men could be perceived as part of boys' in-group. In addition, if girls were basing their answers to the direct measure in Study 1 on an in-group gender bias, we might expect that they would rate girls as being significantly better than boys at math. This was not the case, however. For this reason, based on the data, the explanation that girls stratified the stereotype remains plausible, and future research in this area would serve to increase our understanding of this process.

From an educational perspective, these data suggest the need to go beyond teaching girls that they can do math, to teaching boys and girls that women can do complex mathematics. As girls become women, holding a negative stereotype about women and mathematics may be hindering girls' desire to pursue this field in adolescence and beyond. By providing young scholars with more positive female role models in math and science, children may be in a better position to challenge their current gender stereotypes about women in these domains.

From a social psychological perspective, these data underscore the need to examine how people process, and potentially overcome, stereotypes directed at their groups. It is proposed that people may overcome stereotypes by placing themselves in a subgroup to which the stereotype does not apply. However, future research will be needed to determine the contexts in which this is the case.

Future research also should determine the conditions under which the stratification of a self-relevant stereotype or identity is likely to occur. As was noted previously, Weber and Crocker (1983) found that people subtyped members of an out-group only when the nonstereotypical members could be classified easily along some dimension. Subtyping the self away from a self-relevant stereotype might similarly require an easily identifiable grouping variable. In terms of the gender stereotype about math, girls may be provided with exactly the information necessary for stratification to occur. Given the relative dearth of women in math and science, girls may begin to develop this gender stereotype from at an early age. However, on a daily basis in the classroom, girls cannot help but notice that boys and girls perform equally well in math-related activities. It is no wonder, then, that in the face of this paradoxical information, girls may be in a position to develop a negative stereotype that is specific to women.

It also seems possible that any protection that stereotype stratification has to offer is context-dependent. Although girls may develop a gender stereotype about math in such a way that it does not apply to them, other situations may arise in which they are susceptible to the more general stereotype about females and mathematics. Take the recent study by Ambady, Shih, Kim, and Pittinsky (2001), in which Asian American girls were subtly reminded about their female identity, their Asian identity, or neither identity. Consistent with the authors' earlier work (Shih, Pittinsky, & Ambady, 1999), this study revealed that when girls in kindergarten through second grade were reminded of their female identity, they performed worse on a math test than when their Asian identity or neither identity was primed. The authors suggested that girls who were primed with their female identity were accessing the stereotype about females and mathematics, which, in turn, affected their math test performance. Interestingly, among girls in third grade to fifth grade, priming their female identity actually caused an increase in performance. Consistent with the present data, it seems possible that during those later elementary years, girls have the greatest belief that girls can do math, despite the fact that they might not be developing a comparable attitude about women's abilities.

Although the results of the present studies provide some support for the theory of stereotype stratification, future research will be needed to determine the conditions under which this process helps members of stigmatized groups to overcome negative stereotypes. Longitudinal studies are also needed to examine the attitudes of girls and women toward math and science so that we may better understand how they manage to overcome their susceptibility to the potentially detrimental effects of stereotypes directed at their group. As Sadker and Sadker (1994) noted, it is in our best interest to have talented boys *and* girls pursue higher education and careers in math and science. Hopefully, through continued research, we will identify the best strategies for making these goals a reality.

References

Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 12*, 385-390.

American Association of University Women. (1999). *Gender gaps: Where schools still fail our children*. New York, NY: Marlowe.

Aronson, J., Quinn, D. M., & Spencer, S. J. (1998). Stereotype threat and the academic underperformance of minorities and women. In J. K. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 83-103). San Diego, CA: Academic Press.

Banaji, M., & Greenwald, A. G. (1994). Implicit stereotyping and prejudice. In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario Symposium* (Vol. 7, pp. 55-76). Hillsdale, NJ: Lawrence Erlbaum.

Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science, 210*, 1262-1264.

Betz, N. (1997). What stops women and minorities from choosing and completing majors in science and engineering? In D. Johnson (Ed.), *Minorities and girls in school* (pp. 105-131). Thousand Oaks, CA: Sage.

Brewer, M. B., Dull, V., & Lui, L. (1981). Perceptions of the elderly: Stereotypes as prototypes. *Journal of Personality and Social Psychology, 41*, 656-670.

Brewer, M. B., & Lui, L. (1984). Categorization of the elderly by the elderly: Effects of perceiver's category membership. *Personality and Social Psychology Bulletin, 10*, 585-595.

Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of Personality and Social Psychology, 76*, 246-257.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11*, 135-172.

Eccles, J. S., & Jacobs, J. E. (1986). Social forces shape math attitudes and performance. *Signs: Journal of Women in Culture and Society, 11*, 367-380.

Eisenberg, N., Martin, C. L., & Fabes, R. A. (1996). Gender development and gender effects. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 358-396). New York, NY: Simon & Schuster Macmillan.

Golombok, S., & Fivush, R. (1994). *Gender development.* New York, NY: Cambridge University Press.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139-155.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect. *Psychology of Women Quarterly, 14*, 299-324.

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*, 356-371.

Jacobs, J. E., & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology, 63*, 932-944.

Kagan, J. (1964). The child's sex role classification of school objects. *Child Development, 35*, 1051-1056.

Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology, 68*, 565-579.

Maurer, K. L., Park, B., & Rothbart, M. (1995). Subtyping versus subgrouping processes in stereotype representation. *Journal of Personality and Social Psychology, 69*, 812-824.

Mullen, M. K. (1990). Children's classifications of nature and artifact pictures into female and male categories. *Sex Roles, 23*, 577-587.

Powlishta, K. K. (1995). Intergroup processes in childhood: Social categorization and sex role development. *Developmental Psychology, 31*, 781-788.

Sadker, M., & Sadker, D. (1994). *Failing at fairness: How our schools cheat girls*. New York, NY: Touchstone.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*, 80-83.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4-28.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613-629.

Steele, J., James, J. B., & Barnett, R. C. (2002). Learning in a man's world: Examining the perceptions of undergraduate women in male-dominated academic areas. *Psychology of Women Quarterly, 26*, 46-50.

Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology, 45*, 961-977.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 446-496). New York, NY: McGraw-Hill.