2011

The Internalization of Mathematics Stereotypes in Elementary School Children

Amanda Schell
Wilfrid Laurier University

Your file Votre référence
ISBN: 978-0-494-75382-8
Our file Notre référence
ISBN: 978-0-494-75382-8

The internalization of mathematics stereotypes in elementary school children

by

Amanda Schell

Bachelor of Arts, Psychology, Wilfrid Laurier University, 2009

THESIS

Submitted to the Department of Psychology

in partial fulfillment of the requirements for

Master of Arts

Wilfrid Laurier University

2011

Abstract

This study investigated mathematics-related gender stereotypes, internalization of these stereotypes in 2nd and 5th grade children, and whether there is a relationship between internalization and the underperformance of girls in mathematics. There is evidence that gender stereotype internalization occurs throughout elementary school with an increasing impact on girls' mathematical competence and performance (e.g., Muzzatti & Agnoli, 2007). However, there has been no definite determination with respect to the point at which this process begins. Parents and teachers have displayed gender stereotypical beliefs concerning children as young as three years of age (Lee & Schell, under review). This factor can influence children's attitudes towards mathematics (Bleeker & Jacobs, 2004). The current study included 37 second graders (18 boys, 19 girls) and 27 fifth graders (12 boys, 15 girls). Each student completed tasks designed to measure gender stereotypical beliefs of their own abilities, perceptions of their parents' beliefs, internalization of occupations and activities related to masculine and feminine domains, and an assessment of their actual mathematics ability. Parents and teachers were asked to complete a questionnaire to ascertain their gender-stereotypical beliefs of the students' academic abilities. The findings revealed that children had not internalized mathematics gender stereotypes, girls did not underperform, and adults did not display stereotypical beliefs regarding children's academic competencies. These results may be described by a myriad of explanations such as gender stereotype flexibility, girls' equivalent or higher level of performance in academics, and time of data collection. Implications for future research will be discussed.

Acknowledgements

I would like to thank the students, parents, teachers, and principal of Williamsburg Public School. The study would not have been possible without their cheerful participation. I am additionally grateful to the teachers and Principal of Williamsburg P.S. for their kindness and patience during the data collection process.

An immense amount of gratitude goes to my advisor, Dr. Joanne Lee, for her continued dedication and patience throughout this research project. Thank you to my committee members, Dr. Donna Kotsopoulos and Dr. Eileen Wood, for their reviews and feedback on my initial proposal and final document. Thank you to Dr. Christine Alksnis, the external thesis examiner, for her invaluable feedback. Additionally, I am very grateful to Samantha Makosz for her many hours of help with data collection.

Table of Contents

List of Tables

List of Figures

List of Appendices

The internalization of mathematics stereotypes in elementary school children

In the workforce, a high level of mathematical skill leads to an increase in wages even more so than verbal skills (Mitra, 2002). This wage difference was found for both blue and white collar jobs, and also for both men and women. Moreover, people in mathematics and science related careers reported higher levels of job satisfaction than other careers (Lloyd, Walsh, & Yailagh, 2005). However, fewer women than men pursue careers or higher degree studies in mathematics (Lindsay & Almey, 2006). For example, according to Statistics Canada, in 2008, there were 8,214 university degrees awarded in the field of mathematics, computer and information sciences. Women attained only 2,496 of these degrees which accounted for only 30% of the total degrees awarded in this field (Statistics Canada, 2010). Mathematics-related gender stereotype attitudes have been shown to negatively affect women's mathematics performance (Ambady, Shih, Kim, & Pittinsky, 2001; Fryer & Levitt, 2010; Kiefer & Sekaquaptewa, 2007a; 2007b; Muzzatti & Agnoli, 2007).

It is important to investigate why this gender gap exists and when it begins to emerge in order to increase female participation in these fields. The current study represents an exploratory investigation in understanding gender differences by examining the formation of gender stereotypes in mathematics with elementary school students. Specifically, this study investigated mathematics-related gender stereotypes, internalization of gender stereotypes in 2nd and 5th grade children, and whether there is a relationship between internalization and underperformance of girls in mathematics in later grades of elementary school.

Prior to exploring these issues, it is important to define what is meant by terms such as gender and gender stereotypes. Therefore, the first section of this introduction section defines key terms that are used throughout this thesis. This is followed by a review of the literature for each of the key concepts and finally hypotheses underlying the current thesis are summarized.

Defining Terms

Although the terms "gender" and "sex" are often used interchangeably in popular media or among lay populations, these terms have more selective meanings within academic literatures. Most researchers agree that sex refers to the biological distinction between being male or female (Blakemore, Berenbaum, & Liben, 2009) whereas gender typically represents a cultural understanding of what it means to be masculine and feminine (Blakemore et al., 2009). For the purpose of the present study, it is the cultural rather than biological understanding that is relevant.

A second important term that is pertinent to the present study is the term "stereotype." A stereotype refers to a cognitive construct used to organize information (Wood, Groves, Bruce, Willoughby, & Desmarais, 2003). Stereotypes can include beliefs that specific characteristics are possessed by all members of a group (Arnett, 2007). Cognitively, gender stereotypes can help to organize information and assign characteristics about what it means to be male or female (Arnett, 2007; Wood et al., 2003). Socially, gender stereotypes can lead to the promotion or restriction of behaviours or attitudes which can have either desired or undesirable outcomes (e.g., Ambady et al., 2001). Gender stereotypes are not necessarily representative of reality (Blakemore et al., 2009). For the purpose of the proposed study, gender stereotype will refer to commonly held beliefs about gender that are not always a true representation of reality.

The literature on gender stereotypes utilizes a variety of terms to describe children's gender stereotype knowledge and attitudes. For the purpose of the current study, the definitions described by Signorella and Liben (1985) will be used. Knowledge of gender stereotypes pertains to children's understanding that certain objects, traits, and activities are assigned to either men or women (Signorella & Liben, 1985). The term "knowledge" will be used synonymously with perception, as children's perceptions indicate gender stereotype knowledge. Attitudes concerning gender stereotypes are evident when children endorse the gender stereotype of the culture (Signorella & Liben, 1985). The term "attitude" will be used synonymously with internalization of gender stereotypes throughout the current study.

Literature Review

Development of the understanding of gender. Gender categorization starts early in life, as shown by children as young as nine months of age being able to distinguish between female and male faces (Leinbach & Fagot, 1993). Using a habituation task, Leinbach and Fagot (1993) demonstrated that infants ranging from five to 12 months of age displayed knowledge of gender categories at about nine months old. The researchers found that the task became more difficult for the children when markers of gender such as long hair for women and gendered clothing were removed, but children at 12 months of age were still able to categorize the faces by gender in this more difficult condition. The knowledge of gender of others has been shown to emerge prior to the knowledge of self gender (Thompson, 1975). For example, children at the age of 24 months have been found to have knowledge of gender categories through the use of a picture sorting task; however, they were not able to correctly categorize their own picture into the appropriate gender. Thompson (1975) reported that children are more aware of their own gender by 30 months of age, and children are well aware of others' genders as well as their own by 36 months.

Gender stereotypes. Gender stereotypes can include conventional beliefs about what it means to be male or female and may include factors such as personality, behaviour, appearance and occupation (Blakemore et al., 2009; Six & Eckes, 1991). These stereotypes are not necessarily always false. Blakemore and her colleagues (2009) described the notion of "kernel of truth," in which a stereotype may be based on qualities that are associated with a particular gender, such as women being more likely to wear nail polish. However, they also state that some gender stereotypes are exaggerations to the point that they become false (Blakemore et al., 2009). This exaggeration is true of gender stereotypes related to academic achievement, most notably that boys are more mathematically competent than girls. This is evident in Statistics Canada's report of university degrees awarded in 2008. Women accounted for a majority of students in the following areas: education (78%), health, parks, recreation and fitness (77%), social and behavioural sciences, and law (67%), and humanities (64%), and a substantially lower number of students in areas such as mathematics, computer and information sciences (30%), and architecture, engineering and related sciences (22%) (Statistics Canada, 2010). These stereotypes are not innate, as is evident through Eccles (1987) expectancy model which demonstrated that girls learn to expect less from themselves in mathematics, which subsequently leads to lower mathematics performance and less interest to pursue education or careers in the field.

Development of gender stereotypes in mathematics. Knowledge of mathematics-related gender stereotypes emerges prior to gender stereotype attitudes. At the beginning of elementary school, children are more likely to report an overestimation of their academic abilities, including their mathematics ability (Bouffard, Marcoux, Vezeau, & Bordeleau, 2003; Burnett, 1996; Fredricks & Eccles, 2002; Freedman-Doan et al., 2000; Jacobs, 1991; Miller, Lurye, Zosuls, & Ruble, 2009). This indicates that at the beginning of elementary school, children may not yet have the knowledge of gender stereotypes in the field of mathematics. For example, in their longitudinal study with French-Canadian children, Bouffard and her colleagues (2003) reported that students' competency ratings of their mathematics and reading abilities at grade one decreased at grade three, demonstrating that children have higher competency ratings at the beginning of elementary school. Similar phenomenon was also obtained by Freedman-Doan and her colleagues (2000) of younger children having higher competency levels than older children. The results of their study of first, second, and fourth grade children displayed that younger children (69.7% of 1st graders) were optimistic about their ability to improve in academic domains, including mathematics. However, a number of older, fourth grade children were less likely to believe they could improve their achievement in their worst subject area (55.8%).

Children in early elementary school may be less likely to demonstrate knowledge of gender stereotypes relating to mathematics, but gender stereotype knowledge becomes evident as children progress through elementary school (Burnett, 1996; Freedman-Doan et al., 2000; Kurtz-Costes, Rowley, Harris-Britt, & Woods, 2008; Lloyd et al., 2005). Kurtz-Costes and colleagues (2008) studied fourth, sixth, and eighth grade students. The results demonstrated that even though the girls performed as well as or better than their male peers in mathematics and science, their ratings of self-competence in these fields were significantly lower than the ratings of self-competence completed by the boys. Lloyd and her colleagues (2005) studied a comparable age group of fourth and seventh grade children. Girls reported lower levels of confidence than boys in mathematical ability while performance levels did not significantly differ. In addition to this, there were gender differences in the attributions that children made for their successes and failures: specifically, girls were less likely to attribute their success to internal factors such as their own ability (Lloyd et al., 2005).

Other studies dispute children's knowledge of gender stereotypes at the elementary school level (Kenney-Benson, Pomerantz, Ryan, & Patrick, 2006; Martinot & Desert, 2007; Paulsen & Johnson, 1983; Skaalvik, 1990). Martinot and Desert (2007) reported an example of this discrepancy in how children perceive their abilities. In a sample of fourth and seventh graders, it was found that all girls and older boys believed that girls were more mathematically competent than boys. This result was evident even when gender was made salient for the boys. These specific results may be somewhat indicative of the higher rates of gender equality in Europe, however similar results have also been reported in a North American context (Kenney-Benson et al., 2006; Paulsen & Johnson, 1983). Kenney-Benson and colleagues (2006) demonstrated that children who were measured at grade five and again at grade seven reported no gender differences in mathematics efficacy beliefs, and no gender differences in actual mathematics performance.

Although there is a disconnect in the literature pertaining to when children display mathematics-related gender stereotypes, Steele (2003) demonstrated that children do display gender stereotype knowledge prior to gender stereotype attitudes. When 1st through 4th grade girls were asked to sort mathematics-related pictures, they were more likely to place these pictures into the male-related groupings (Steele, 2003). Additionally, when both boys and girls were asked to draw pictures of an adult mathematician, the children were more likely to draw a man. However, this was not true when the children were asked to draw the mathematician as a child, as they were then more likely to draw a child of the same sex as themselves (Steele, 2003).

The process of internalizing gender stereotypes is ongoing throughout elementary school and it is unknown when exactly children report gender stereotype attitudes. Gender stereotype attitudes have been shown to negatively affect girls' and women's mathematics performance, even if women do not explicitly report endorsing the stereotypes (Ambady et al., 2001; Brown & Josephs, 1999; Dick & Rallis, 1991; Else-Quest, Hyde, & Linn, 2010; Fryer & Levit, 2010; Kiefer & Sekaquaptewa, 2007a; 2007b; Muzzatti & Agnoli, 2007).

The theory of psychological disengagement of African Americans in academic domains parallels the effects of gender stereotype attitudes. This theory has evolved from the phenomenon that African American children identify high academic achievement as a "White thing" and thus lose interest in academics, develop more negative attitudes toward academics, and eventually perform with less academic success (Fordham & Ogbu, 1986;

Strambler & Weinstein, 2010). The theory of psychological disengagement can be relevant to mathematics stereotypes in that both can be viewed as social constructs. Towards the end of elementary school girls become inclined to avoid mathematics, attribute their failures in mathematics to a lack of ability, and feel that further effort would not lead them to success in mathematics (Dickhauser & Meyer, 2006; Stipek & Gralinski, 1991). These findings demonstrate that gender stereotypes towards mathematics have been internalized and have affected girls' attitudes towards the subject, as is similar to psychological disengagement in African American children. Although it is not certain when mathematics stereotypes are internalized, it is well documented that by the time students reach high school, the idea that mathematics is a male-dominated subject is prevalent, and this has an influence on girls' participation in mathematics courses and their actual mathematics performance (Bleeker & Jacobs, 2004; Dick & Rallis 1991; Else-Quest et al., 2010; Hyde, Fennema, Ryan, Frost, & Hopp, 1990).

Gender stereotype attitudes have been shown to affect actual mathematics performance in children in early elementary school (Ambady et al., 2001; Fryer & Levitt, 2010; Muzzatti & Agnoli, 2007). Ambady and colleagues (2001) demonstrated that when gender was made salient for children age five to seven years old, as well as 11- to 13-year olds, girls' mathematics performance decreased, whereas boys' performance increased. Fryer and Levitt (2010) also indicated the presence of the gender gap in mathematics in their nationally representative sample. They report that no gender differences were found in children's mathematics performance at the onset of grade one. However, by the third grade, girls had fallen approximately two-tenths of a standard deviations behind their male peers.

The effects of attitudes towards mathematical stereotypes on students in high school is demonstrated by Else-Quest and colleagues (2010) who found evidence for the gender stereotype that males are more mathematically competent than females in their meta-analysis of teenagers ages 14 to 16 years old. In their work, they reported that gender differences in mathematical performance were small; however boys were more positive towards mathematics than girls. Else-Quest and colleagues' (2010) study included a cross-national sample, and they noted that the gender gap in mathematics was related to gender equity within the country (i.e., women's access to education).

Sherman (1980) also conducted a longitudinal study following students from grade eight through to grade 11. These students demonstrated the negative effects of gender stereotype attitudes in mathematics. In grade eight, there were similar performance levels in mathematics for both genders. However, by the time the students reached the 11^{th} grade, girls' attitudes towards mathematics had declined far more significantly than their male peers. The 11^{th} grade boys were more confident in mathematics, thought mathematics to be more useful, and also regarded mathematics as more of a male-domain than did their female peers. More importantly, girls' mathematical performance had declined and boys were outperforming girls. These studies suggested that girls had internalized gender stereotypes by high school and that they believed they were not as competent as their male peers.

Gender stereotype attitudes can affect women's mathematical performance at the post-secondary education level, even when they do not explicitly endorse them (Kiefer & Sekaquaptewa, 2007a; 2007b). It was reported that implicit stereotypes affect college students' mathematical performance (Kiefer & Sekaquaptewa, 2007a). Implicit stereotypes refer to stereotypes which are at a subconscious level and are not explicitly endorsed by the individual (Kiefer & Sekaquaptewa, 2007a). Kiefer and Sekaquaptewa's (2007a) study involved women enrolled in a post-secondary level calculus class. It was found that the women who were more gender-identified performed less well in the mathematics course and had less interest in pursuing a mathematics-related career, than those who were less gender-identified. The study highlighted the adverse impact of internalized attitudes of gender stereotypes relating to mathematics even when women were not explicitly aware of these stereotypical beliefs. There are other studies which support the conclusion that even women who pursue high levels of mathematics education are not immune to the negative effects of gender stereotype attitudes and as a result, some of these women underperform in mathematics (Brown & Josephs, 1999; Kiefer & Sekaquaptewa, 2007b; Schmader, 2002).

Expectations for success in mathematics, as well as self-efficacy in mathematics, have been linked with the decision-making process of whether to continue pursuing advanced mathematics and the importance placed on mathematics with respect to career goals (Eccles, 1987; Fredricks & Eccles, 2002). Low mathematical achievement at the beginning of high school has been shown to deter adolescents from further pursuit of a mathematics education, thereby eliminating the potential of pursuing many prestigious careers (Shapka, Domene, & Keating, 2008). This demonstrates that internalizing gender stereotypes which results in lower expectations of girls in the field of mathematics can produce negative effects that reduce career options and inhibits the development of their full potential.

Development of the social construction of gender stereotypes. The internalization process begins with the knowledge of gender stereotypes, which can be learned from adult role models during childhood. Initially, parents are the most important figures in children's lives and they are the primary sources from which to model behaviour (Bowlby, 1982). Parental encouragement and provision of mathematical and scientific materials have been shown to result in an increased level of children's participation in mathematics, science, and computer activities (Simpkins, Davis-Kean, & Eccles, 2005).

A recent study demonstrated that some parents have mathematical-related gender stereotypical beliefs when children are as young as three years of age and therefore relate to their children in accordance with those beliefs (Lee & Schell, under review). Parents have reported stereotypical beliefs that boys are more competent in mathematics and science than girls, despite a lack of any gender differences in their elementary school children's performance (Jacobs, 1991; Lee & Schell, under review; Parsons, Adler, & Kaczala, 1982; Raty & Kasanen, 2007; Tenenbaum, 2009; Tiedemann, 2000). Parental stereotypical beliefs influence how children view their own abilities, and impact children's expectations for future success, amount of effort required to do well, and importance placed on, and interest in the subject (Bleeker & Jacobs, 2004; Parsons et al., 1982; Tenenbaum, 2009; Tenenbaum & Leaper, 2003). Bleeker and Jacobs (2004) demonstrated that mothers' negative perceptions of their children's potential mathematics success would predict children's negative feelings toward mathematics-related careers later in life (i.e. children would have less interest in a mathematics-related career). These results illustrated the influence that parents exercise over their young children. Gender stereotypical beliefs can also have a greater impact on perceptions of mathematics ability than past mathematical performance (Ambady et al., 2001; Dick & Rallis, 1991; Muzzatti & Agnoli, 2007; Parsons et al., 1982).

Besides parents being significant role models, teachers can also be crucial role models in children's lives. The literature on the role of teachers in socializing the development of gender stereotypes is less conclusive and is scarcer than literature studying the role of parents. Studies have reported that teachers perceive boys to be more competent in mathematics than girls in preschool, elementary and high school (Blakemore et al., 2009; Bleeker & Jacobs, 2004; Lee & Schell, under review; Tiedemann, 2000). Blakemore and her colleagues have found that teachers treat preschool boys and girls differently. In preschool classes, teachers have been found to respond with a more gentle approach when girls are acting out and with a less gentle approach when boys are acting out. These findings indicate that gender stereotypes influence teachers' treatment of children starting as early as preschool.

Teachers continue this differential treatment into elementary school and high school where they are more likely to encourage boys to pursue mathematics through university and into a career (Blakemore et al., 2009; Dick & Rallis, 1991). Despite this differential treatment, it appears that girls rely more heavily on their teachers' opinions than do boys, and use this opinion to make decisions in terms of future education and career choices (Dickhauser & Meyer, 2006). In spite of the findings of teachers' being influential in students' choices, Helwig, Anderson and Tindal (2001) reported that teachers do not influence the development of mathematics-related gender stereotypes in children in mid-elementary school. These studies display the disparity in the research on teachers' influence concerning gender stereotypes. Overall, there is more evidence to suggest that teachers do influence the development of gender stereotypes (Blakemore et al., 2009; Bleeker & Jacobs, 2004; Lee & Schell, under review; Tiedemann, 2000).

By examining the relationship between teachers' mathematics-related gender stereotypes and children's perceptions of their abilities while investigating their actual abilities in mathematics, we can better understand another contributing factor in the development of how children internalize gender stereotypes. The inclusion of teachers' beliefs also allows for the comparison between parents' and teachers' beliefs, and how these significant adults influence the development of children's gender stereotypical attitudes.

Proposed Study

Children become aware throughout the early elementary school years that mathematics has been accepted as a male-dominated domain (Steele, 2003). The acquisition of this knowledge of gender stereotypes results in children developing gender stereotypical attitudes toward the middle and end of elementary school that girls cannot perform as well as boys in mathematics (Stipek & Gralinski, 1991). The goal of the current study is to extend previous research by examining when the internalization of these gender stereotypes occurs, and whether this internalization triggers the underperformance in the field of mathematics by girls.

Hypotheses

There are five hypotheses associated with this study.

Hypothesis 1. The first hypothesis examined the degree of stereotype internalization exhibited by the children. The examination of the degree of internalization requires ascertaining the gender stereotypical attitudes about oneself versus others. It was expected that older children (5th graders) would report higher levels of internalized gender stereotypes than younger children (2nd graders). Studies indicate that young children have knowledge of gender stereotypes (Trautner et al., 2005; Martin & Ruble, 2009); therefore it was expected that all children in this study would have acquired this knowledge, but would differ in the degree gender stereotypical attitudes according to their age. It was hypothesized that a child would develop gender stereotypes for others prior to reporting these beliefs on a personal basis (i.e., internalizing these stereotypes to their belief system). Therefore, it was expected that the older children (5th graders) would be most likely to report gender stereotypical beliefs for both others and themselves.

Hypothesis 2. This study explored the relationship of gender stereotypes and actual mathematics performance. Previous literature has reported that girls begin to underperform in comparison to their male peers at grade three (Fryer & Levitt, 2010). It was hypothesized that gender differences would emerge in children's mathematics performance at the fifth grade but not in the second grade.

This hypothesis was based on the theory of psychological disengagement of elementary and high school African American students in academic domains (Fordham & Ogbu, 1986; Strambler & Weinstein, 2010). Gender differences were expected in fifth grade children because psychological disengagement has been found at this age with minority students (Strambler & Weinstein, 2010), so it was anticipated that fifth grade children had internalized gender stereotypes and these attitudes would have affected their mathematics performance, especially for girls. Second grade children were not expected to have internalized gender stereotypes, so this would not yet have had an effect on mathematics performance.

Hypothesis 3. The study explored parents' stereotypical beliefs. In light of research revealing that parents hold gender stereotypical views for mathematics ability of children as young as three years old (Lee & Schell, under review), it was expected that parents would rate boys as being more mathematically competent than girls across both grades. Additionally, it was expected that parents' stereotypical beliefs would be related to children's level of internalization. Therefore, parents' stereotypical beliefs are expected to positively correlate with children's gender stereotypical attitudes.

Hypothesis 4. The study explored teachers' stereotypical beliefs. Similar to parents, previous research has demonstrated that teachers rate boys as being more mathematically competent than girls starting at a young age (Lee & Schell, under review), therefore it was expected that boys' mathematics abilities would be rated higher than girls. Additionally, it was expected that teachers' stereotypical beliefs would be correlated with children's gender stereotypical attitudes, however the correlation was expected to be stronger between parents and children than the correlation between teachers and children because parents would only have to have knowledge of one child's abilities, whereas teachers would be rating many children's abilities so they may be less accurate.

Hypothesis 5. The last hypothesis examined the relationship between parents' and teachers' stereotypical beliefs. Research has demonstrated that parents and teachers are often similar in their ratings of children's academic abilities (Karkkainen, Raty, &

Kasanen, 2010). Therefore, it was expected parents' and teachers' stereotypical beliefs would be positively correlated.

Method

Participants

There were 64 children who completed the study. The children were in second and fifth grade, and there were 37 second graders (19 girls, 18 boys) and 27 fifth grade children (15 girls, 12 boys). The mean age of the second grade children was 7.38 years (*SD*= .492), and the mean age of the fifth grade children was 10.63 years (*SD*=.492). Parental education was converted to a number scale, with one being equal to high school level and four being equal to graduate school level. The mean education level of both mothers and fathers was college level ($M_{\text{mother}} = 2.09$, $SD = .921$; $M_{\text{father}} = 2.15$, $SD = .989$). The children were recruited from one local school in the Kitchener-Waterloo area. Prior to participating in the study, written consent was obtained from the school Principal, the teachers and all parents, and the children were asked for oral assent. The school received $3 for each participating family.

The parents and teachers of each participating child were asked to complete a 10-minute questionnaire pertaining to the child's literacy and mathematical abilities (see Appendix G and H). Teachers were compensated for time spent completing questionnaires based on the average stipulated hourly rate for the grade they were teaching. There were seven teachers who completed questionnaires for the current study. Five of the teachers were grade two teachers (4 women, 1 man) and two grade five teachers (1 woman, 1 man).

Materials and Procedure

The participating children completed four tasks: the picture cards task, the rabbit family task, the Children's Occupations, Activities, and Traits (COAT) questionnaire (Liben & Bigler, 2002), and subtests of the Woodcock-Johnson III Tests of Achievement battery (WJ III ACH) (Woodcock, McGrew, & Mather, 2001). Parents completed one survey regarding their child and teachers completed one survey for each participating child in their classroom. The parent and teacher questionnaires provided an indication of the extent of the gender stereotypes that these significant adults hold.

Child Tasks. Four tasks were administered to each child. The tasks were presented in a counterbalanced order by the author and one other trained female research assistant. The children completed the WJ III ACH with the author and the picture cards task, the rabbit family task, and the COAT questionnaire with the other research assistant. The child provided his/her oral assent after the researcher described each of the four tasks.

The picture cards task was designed to measure the child's stereotypical beliefs concerning his or her own abilities. Seven picture cards were presented to each child. The cards depicted a child of the same sex as the participant performing the following activities: reading, doing mathematics, drawing, listening to music, playing outside, playing with Lego blocks, and playing a board game. Each child was asked to name the type of activity depicted on each of the cards. The child was then asked to indicate which of the activities he or she was best at. The chosen card was removed, and the child was asked again which of the remaining activities he or she is best at. This continued until the child had given each of the activities a ranking from one to seven, with one indicating the highest perception of ability in the specified activity and seven indicating the lowest perception of ability in the specified activity. If our hypothesis is supported, a girl who has a high level of internalized gender stereotypes would rank her ability in mathematics closer to seven and her ability in reading closer to one. In contrast, a boy who has internalized mathematics stereotypes would rank his ability in mathematics closer to one, and his reading ability closer to seven.

The rabbit family task (Lee & Schell, 2010; Schell & Lee, 2009) was meant to ascertain the extent of the child's knowledge of gender stereotypical beliefs held by his or her parents. Two separate toy rabbit figure families, consisting of a father rabbit, mother rabbit, brother rabbit, and sister rabbit, was presented to the child one family at a time. The first rabbit family set represented a gender neutral family in that no stereotypical gender markers were apparent to separate the female from male members of the family (See Figure 1). The second rabbit family was the gender stereotypic family, as the male and female members of the family were wearing clothing which was indicative of their gender (e.g., mother and sister rabbits in dresses) (See Figure 2). Children always saw the gender neutral family first followed by the gender stereotypic family.

The gender neutral rabbit family required children to draw upon their own gender stereotypes without any cues whereas the gendered rabbit family primed for stereotypes by providing the clothing cues. In each case, after the family was introduced the child was given four of the cards viewed in the picture cards task (i.e., reading, doing mathematics, drawing, and playing with Lego blocks). The child was presented with the gender neutral father rabbit and asked *"Which of the activities does father rabbit think that brother rabbit is best at?"* The child assigned rankings to each of the four cards, with one being the activity that father rabbit thinks that brother rabbit is best at, and four being the activity that father rabbit thinks brother rabbit is worst at. This task was repeated with the gender neutral mother. The order of presentation of the parent rabbits was counter-balanced between mother and father rabbit. To downplay the juvenile nature of the task for the fifth grade children, they were told that younger children had been asked the same questions, and we wanted to know what older children, such as themselves, thought as well.

The child was then shown the gendered rabbit family and asked: "*This is another father rabbit. What do you think he would think brother rabbit is best at?*" The child again assigned rankings for both parents for each of the children's abilities. Again, the presentation order of the mother and father rabbit was counter-balanced. The two rabbit families provided a comparison to determine whether the presence or absence of cues impacts on children's use of stereotypes when assigning the picture cards a value.

To measure children's internalization of gender stereotypes, Liben and Bigler's (2002) Children's Occupations, Activities, and Traits (COAT) questionnaire was administered. This questionnaire measured attitudes about gender stereotypes by inquiring about children's beliefs towards typically male and typically female occupations, activities, and traits. In addition, the questionnaire examined gender stereotype attitudes towards others and towards the self. COAT is the children's version of the adult questionnaire (OAT), and is appropriate for children between 6 to 14 years of age. The short version of the COAT was used as time was constrained due to the combination of the attention span of the younger participants and the number of tasks involved in this study. There were two subscales of this measure, the first concerning attitudes toward others (COAT-AM), and the second pertaining to attitudes about the self (COAT-PM). Each of the subscales measured gender stereotypical attitudes concerning others or the sex-typing of the self respectively, in three domains of occupations (COAT-AM: Appendix A; COAT-PM: Appendix D), activities (COAT-AM: Appendix B; COAT-PM: Appendix E), and traits (COAT-AM: Appendix C; COAT-PM: Appendix F). The Cronbach's alpha indicates high internal consistency for the short version of the COAT versus the full version of the COAT. The short and long version (respectively) on the COAT-AM compare as follows: Occupations subscale (feminine items: .86 vs. .81; masculine items: .96 vs. .84), Activities subscale (feminine items: .91 vs. .83; masculine items: .88 vs. .83), and Traits subscale (feminine items: .93 vs. .84; masculine items: .95 vs. .85). The Cronbach's alphas for the short and long version (respectively) on the COAT-PM are as follows: Occupations subscale (feminine items: .91 vs. .82; masculine items: .90 vs. .78), Activities subscale (feminine items: .88 vs. .83; masculine items: .86 vs. .80), and Traits subscale (feminine items: .81 vs. .82; masculine items: .83 vs. .67) (Liben & Bigler, 2002).

The two subscales (COAT-AM and COAT-PM) consisted of 75 items, 25 items in each of the three domains. The 25 items per domain were further divided into ten masculine items, ten feminine items, and five neutral items. Examples of items on the COAT-AM questionnaire included asking the children who they thought should be a police officer (occupations), iron clothes (activities), and be dominant (traits). The COAT-PM asked children about their own attitudes in three domains, and included items such as: how much would you want to be a librarian (occupations), how often do you build forts (activities), and how much a trait, such as emotional, is like them. Each subscale (COAT-AM and COAT-PM) took approximately ten minutes to complete. The length of administration of the questionnaire may have been slightly longer for the children in 2nd grade, as they required clarification on some of the items (i e., definitions of certain terms such as geography).

As suggested by Liben and Bigler (2002), the COAT-PM about gender-related feelings concerning the self was administered prior to the COAT-AM to avoid biasing the children's responses about their attitude towards gender stereotypes. The younger children were also shown pictures for the COAT-AM to aid them with their answers. The six picture cards depicted three girls' (or women) faces, three boys' (or men) faces, and both two girls' (women) and two boys' (men) faces to match the COAT-AM response options. The dependent variable of the COAT-AM was the proportion of gender stereotypical responses. Therefore, to score this measure, the number of feminine items assigned to only girls was added to the number of masculine items assigned to only boys, and then divided by the total number of gendered items (60 items) to get a proportionate stereotypical score. Higher scores on this measure would indicate greater gender stereotyping. The COAT-PM was divided by feminine and masculine scores, and number of stereotype responses the children made will be indicative of their feminine and masculine scores. Therefore, girls with a high feminine score (close to 4) and boys with a high masculine score would indicate a high level of gender stereotype internalization.

The final task for the children was the Woodcock-Johnson III Tests of Achievement battery (WJ III ACH) which was a measure of their actual mathematics and literacy ability (Woodcock et al., 2001). The WJ III ACH provided age- and grade-based norms against which to compare the children's performance (Mather & Woodcock, 2001). The children completed the Broad Math cluster of the WJ III ACH. The Broad Math cluster contained the Calculation subtest (Test 5), the Math Fluency subtest (Test 6), and the Applied Problems subtest (Test 10). The Calculations subtest (Test 5) involved performing mathematical computations such as addition, subtraction, multiplication, division, and combinations of these operations. This subtest also measured abilities in areas such geometry, trigonometry, logarithmic and calculus problems, as well as involving decimals, percents, fractions, and negative numbers. This subtest took about ten to fifteen minutes to complete. The Math Fluency subtest (Test 6) is a three-minute timed test in which the child's ability to perform simple addition, subtraction, and multiplication problems was assessed (Mather & Woodcock, 2001). The Applied problems subtest (Test 10) included problems which required the analysis of information and solving of problems involving extraneous information. These items required the child to correctly identify the appropriate information and calculations utilized to solve the problem. This test took approximately 15 to 20 minutes to complete. These tests were age appropriate as there were specific start points depending on grade level. The raw scores were recorded with a correct item receiving one point. The total duration of the Broad Math cluster was between 28 to 38 minutes.

Children's literacy skills were assessed using the Passage Comprehension subtest (Test 9) and the Word Attack subtest (Test 13) (Mather & Woodcock, 2001). The literacy component was included as a control variable to compare against mathematics performance in order to ensure whether outcomes are a function of general ability or stereotypes. The Passage Comprehension subtest (Test 9) involved matching a picture representing a word with the actual picture; matching a picture with a phrase; and identifying missing words that belong in a short passage. This subtest took approximately 15 to 20 minutes to complete. The Word Attack subtest (Test 13) required the participants to sound out non-words or low-frequency words that are phonically consistent with English orthography (Mather & Woodcock, 2001). The difficulty increased with each item. This subtest took approximately five to ten minutes to complete. The duration of the literacy subtests was approximately 20 to 30 minutes. The total time duration of the five WJ III subtests being used was between 48 to 68 minutes.

Parent questionnaire. The primary caregiving parent provided signed consent for his or her child to participate in the study and each parent completed a short questionnaire about his or her child's literacy and mathematics abilities (See Appendix G). The parent questionnaire consisted of nine items: three literacy items and six numeracy items. The literacy items rated the child's abilities in vocabulary, reading unfamiliar words, and reading comprehension. The mathematics items consisted of understanding numerical relations and mathematical operations, geometry, mathematical reasoning and analysis, and applied problem solving. Each parent was asked to rate his or her child on a 1 (definitely not as good as children this age) to 5 (well above children this age) Likert-type scale. The parent questionnaire also contained demographic information including the child's age and gender, number and ages of any siblings, parents' highest level of education, and whether the mother, father or other caregiver of the child completed the questionnaire. The questionnaire took approximately about five to ten minutes to complete.

Teacher questionnaire. Teachers were also asked to complete a short questionnaire for each child who had parental permission to participate (See Appendix

H). The teacher questionnaire contained the same nine items as the parent questionnaire regarding children's individual literacy and mathematics abilities. The demographic information in the teacher questionnaire included teacher's gender, highest level of education, additional professional training, number of years as a teacher, and number of years that the teacher has taught the particular grade. The questionnaire took approximately five to ten minutes to complete.

Results

The current study examined children's perceptions of their abilities, children's actual abilities, and teachers' and parents' perceptions of children's abilities. Each section is outlined below.

Children's Perceptions

The first hypothesis of this study examined the degree of children's internalization of mathematical gender stereotypes. To test this hypothesis, the picture cards task data were analyzed to ascertain if children displayed gender stereotypes concerning their own abilities. This task provided the opportunity to analyze children's perceptions specific to their mathematics and reading ability. In this task, the children ranked the cards from one to seven, meaning that a lower ranking (closer to one) indicated higher perception of ability. The descriptive information such as means, standard deviations, and ranges are presented in Table 1.

Mann-Whitney 2-Independent tests were conducted according to the students' grade level (grade 2 and grade 5). The analyses revealed significant gender differences in grade two children's rankings of their reading ability. Grade two boys' ($M = 4.5$, $SD = 1.58$) ranked their reading ability significantly higher (closer to seven) than grade two girls' (*M* = 2.84, *SD* = 2.36), which indicates the younger boys were less confident in their reading ability ($U = 86.000$, $p = .008$) There were no significant gender differences between grade two children's rankings of their mathematics ability ($U = 141.500$, $p = .375$), grade five children's ranking of their reading ability ($U = 78.500$, $p = .579$), or grade five children's ranking of their mathematics ability ($U = 84.00$, $p = .780$) Therefore, only younger children displayed stereotypical views of their reading abilities, but not their mathematics abilities. Older children did not report stereotypical perceptions of their reading or mathematics abilities.

Wilcoxon signed ranks *t*-tests were also conducted for the picture cards task to compare the activities in which girls and boys separately rank themselves stronger and weaker. These analyses revealed that grade two boys' mathematics rankings (*M* = 3.00, *SD* = 1.58) were significantly lower than their reading ability (*M* = 4.50, *SD* = 2.14), which is indicative of higher confidence in their mathematics abilities ($Z = -2.177$, $p = .031$). There were no significant differences in the rankings of reading and mathematics ability for grade two girls ($Z = -.997$, $p = .324$), grade five boys ($Z = -.119$, $p = .932$), or grade five girls ($Z = -.631$, $p = .280$). Therefore only the younger boys reported perceptions of their ability in line with the gender stereotype that boys are better in mathematics than literacy.

Mann-Whitney tests were also used to analyze gender differences in children's rankings from the Rabbit Family task, a test used to measure children's knowledge of adults' gender stereotypical beliefs. Again for this task, students ranked the activities from one to four, with one indicating a higher perception of ability. The tests were conducted to analyze differences between girls' and boys' rankings of the gender neutral

Father and Mother Rabbits, as well as the gender stereotypic Father and Mother Rabbits to determine if one gender reported more awareness of parents' stereotypical beliefs. Means and standard deviations are presented in Table 2, and a lower mean indicates higher perception of competence. The children ranked what each parent rabbit thought about each child rabbit's abilities in mathematics and reading. There were no gender differences in the younger or older children's rankings of any of the parent-child dyads in regards to mathematics or reading abilities (See Table 3 for the complete Mann-Whitney results). Therefore, children did not report knowledge of adult's stereotypical beliefs towards mathematics and literacy.

Mann-Whitney analyses were also performed to see if children's rankings were influenced by the gender stereotypic clothing of the gender stereotypic Rabbit family. There were no significant differences between children's rankings of the two rabbit families. Therefore, the stereotypic clothing of the gendered Rabbit family did not influence children's stereotypic beliefs.

Correlations were performed between the Neutral and Gendered Rabbit Family rankings to ascertain if children's rankings were consistent across the two Rabbit families[1]. There were positive correlations between the Neutral Father's and Gendered Father's ranking of brother's reading ability (r = .251, p = .046), Neutral Father's and Gendered Father's ranking of brother's mathematics ability (r = .320, p = .01), and Neutral Father's and Gendered Father's ranking of Sister's reading ability (r = .296, p = .018). There were no other significant correlations between the Neutral Rabbit family and the Gendered family's ranking of the rabbit children.

[1] [1] Pearson bivariate correlation is not the appropriate statistical test but was used here because there are no easily available non-parametric correlation tests for categorical data.

To continue testing the first hypothesis, the children's responses on the COAT questionnaires were analyzed to determine if there were gender or grade differences in children's gender stereotypical attitudes. Three proportionate scores were calculated, as was done by Liben and Bigler (2002). A total masculine and feminine score was calculated for the COAT-PM, and a total proportionate score was calculated for the COAT-AM.

The COAT-PM was the personal measure which assessed children's self-endorsement of gender stereotypes pertaining to occupations, activities, and domains. The COAT-PM masculine proportionate score was calculated by adding the total number of points on the masculine items from all three domain questionnaires. The total score of those masculine items was then divided by the total number of masculine items for all three domains (i.e., 10 masculine items per domain for a total of 30 masculine items). The same was done with the feminine items to create the total proportionate feminine score (Liben & Bigler, 2002). Separate masculine and feminine scores needed to be calculated for the COAT-PM because the items do not load on the same factors (Liben & Bigler, 2002).

The COAT-AM proportionate score is a measure of stereotypic responses for others, with a higher score indicating greater stereotyping. The COAT-AM proportionate score is a single score and was calculated by adding up the number of "only women/girls" responses on the feminine items, and all the "only men/boys" responses on the masculine items. The sum of these items is then divided by the total number of gendered items, resulting in one proportionate score for the COAT-AM (i.e., 20 gendered items per subtest for a total of 60 gendered items) (Liben & Bigler, 2002).

An independent samples *t*-test was performed comparing girls and boys, with each of the three proportionate scores as the dependent variable. These analyses revealed that children did not report stereotypical sex-typing of self [COAT-PM $_{masculine}$: t (62) = -.749, p = .456; COAT-PM $_{feminine}$: t (62) = .710, p = .480], nor did they report gender stereotype beliefs of others [COAT-AM: t (62) = -1.024, p = .310].

A repeated measures analysis of variance (ANOVA) was performed for each of the COAT-AM and COAT-PM questionnaires. A within-subjects 3 (domain: occupation, activity, trait) x 2 (item type: masculine and feminine) x between subjects 2 (participant gender: male and female) x 2 (grade: 2nd and 5th) was conducted with the dependent variable being the children's scores on the COAT-AM questionnaire (see table 4 for means and standard deviations for the COAT-AM). There was a significant main effect of domain [F (1, 60) = 93.874, p < .001, η^2 = .610]. Bonferroni analyses revealed that children reported greater stereotypic responses for activities (M = .4927, SD = .031) on the COAT-AM, than occupations (M = .411, SD = .027) (p = .001) or traits (M = .121, SD = .020) (p < .001). There was also a significant interaction between domain and item type [F (1, 60) = 6.103, p = .016, η^2 = .092], with children displaying the most stereotypical responses on the masculine activities (M = .507, SD =.035) and feminine activities (M = .478, SD = .034), and least stereotypical responses on masculine traits (M = .102, SD = .018) and feminine traits (M = .140, SD = .025) (see figure 3). Post hoc analysis of the domain by item type interaction did not reveal any simple main effects. There were no other significant main effects or interactions for children's responses on the COAT-AM. This demonstrates that children displayed more gender stereotype attitudes of others for activities than for occupations or traits, however older children did not demonstrate a higher level of gender stereotype attitudes on the COAT-AM than younger children.

A separate within-subjects 3 (Domain: occupation, activity, trait) x 2 (item type: masculine and feminine) x between subjects 2 (gender: male and female) x 2 (grade: 2nd and 5th) ANOVA was also conducted for the COAT-PM with the dependent variable being the children's scores on the COAT-PM questionnaire (see table 5 for means and standard deviations). The COAT-PM scores are an average rating between one and four. The scores are calculated by adding the responses from all the masculine items together and then dividing that total by 10, which is the total number of masculine items per domain, to give the average rating between one and four. The same was done with the feminine items. This creates an average masculine and feminine score for each of the three domains.

There were significant main effects of domain [$F(1, 60) = 69.886, p < .001, \eta^2 = .538$], in which children reported greater self-endorsement of traits than occupations ($p > .001$) or activities ($p > .001$), and greater self-endorsement of occupations than activities ($p > .001$) (see table 6 for means and standard deviations). There was also a main effect of item type [$F(1, 60) = 4.029, p = .049, \eta^2 = .063$], which revealed that children reported greater self-endorsement of masculine ($M = 2.399$, $SD = .050$) than feminine ($M = 2.310$, $SD = .048$) items ($p = .049$).

These main effects were qualified by a significant interaction between domain and item type [$F(1, 60) = 29.499, p < .001, \eta^2 = .330$], in which children demonstrated higher levels of self-endorsement for feminine traits ($M = 2.842$, $SD = .062$) than masculine traits ($M = 2.766$, $SD = .053$), self-endorsed feminine activities ($M = 2.002$, $SD = .063$)

more than masculine activities (*M* = 1.994, *SD* = .069), but showed greater endorsement of masculine occupations (*M* = 2.436, *SD* = .071) than feminine occupations (*M* = 2.086, *SD* =.072) (see figure 4). Post hoc analysis of the domain by item type interaction revealed a simple main effect of occupations ($p = .001$). There was also a significant interaction between item type and participant's gender [$F(1, 60) = 57.271, p < .001, \eta^2 = .488$] in which girls demonstrated higher levels of self-endorsement for the feminine items ($M_{mas} = 2.297$, *SD* = .067; $M_{fem} = 2.543$. *SD* = .065) and boys self-endorsed masculine items (*M* = 2.50, *SD* = .073) more than feminine items (*M* = 2.077, *SD* = .070)(see figure 5).

There was an interaction of domain, item type, and grade [$F(1,60) = 4.370, p = .041, \eta^2 = .068$], in which second grade children self-endorsed masculine occupations and activities more than the corresponding feminine items, but endorsed feminine traits more than masculine traits (see table 5 for means and standard deviations). Fifth grade children also endorsed masculine occupations more than feminine ones, but conversely showed greater endorsement of feminine activities and masculine traits (see figure 6) (see table 6 for means and standard deviations). Post hoc analyses of the interaction between domain, item type and grade revealed a simple main effect of grade 2 children's endorsement of occupations ($p = .008$) and traits ($p = .001$), but no significant simple main effects for the fifth grade children. The final interaction was between domain, item type, and participant's gender [$F(1, 60) = 27.884, p < .001$ $\eta^2 = .371$] in which girls self-endorsed the feminine items in each of the three domains and boys self-endorsed the masculine items in each of the three domains more than the feminine items (see figure 7) (see table 6 for means and standard deviations). Post hoc analyses of the domain by item type by gender interaction revealed a simple main effect of boys' scores on the occupation items ($p > .001$) and activity items ($p = .003$), as well as simple main effects of girls' scores on the occupation items ($p =$ 043), activity items ($p = .001$), and traits items ($p = .033$). These interactions demonstrate that children did display gender stereotype attitudes concerning themselves in the domains of occupations, activities and traits.

Children's Actual Abilities

The second hypothesis examined the relationship between children's gender stereotypic attitudes and mathematical performance. The WJ III ACH was the measure of children's actual ability in mathematics and literacy. Each correct item on the subtests received a score of one; therefore higher scores were indicative of higher ability in both mathematics and literacy. A 2 (gender: boys and girls) x 2 (grade: 2^{nd} and 5^{th}) ANOVA was conducted for each of the three mathematics subtests, as well as the two literacy subtests (see table 7 for means, standard deviations and range, and table 8 for age and grade estimates). A separate ANOVA was conducted for each subtest to analyze gender differences on the specific skill set that the subtest measured.

There was a significant main effect of gender for the mathematics subtest Test 10: Applied problems [$F(1, 60) = 7.863, p = .007, \eta^2 = .118$], as well as for the literacy subtest Test 9: Passage comprehension [$F(1, 60) = 6.650, p = .012, \eta^2 = .100$]. Overall, girls outperformed boys both on Test 10: applied problems ($M_{Girls} = 97.059$, $SD =$ 10.685; $M_{Boys} = 90.500$, $SD = 8.970$) and Test 9: passage comprehension ($M_{Girls} =$ 105.818, $SD = 12.548$; $M_{Boys} = 95.267$, $SD = 15.102$). There were no other significant main effects or interactions for any of the five subtests.

Children's actual abilities were also analyzed according to children's report card grades. There were three literacy strands (reading, writing, communication) and five mathematics strands (number sense, measurement, algebra, geometry, and data management) included on the report cards which were completed by the teachers. The report cards were collected at the end of the school year so there were grades from each of the three reporting terms, for a total of three grades per strand. Every strand was not completed at each reporting term, so not all strands had a total of three grades. Because of this, average grades were calculated for each of the three literacy strands and five mathematics strands. These average grades were first created by converting the number grades to a 12-point GPA scale (e.g., $A^+ = 12$, $A = 11$). The converted grades were then averaged across the three reporting terms for one average grade per strand.

A repeated measures 2 (grade: 2 and 5) x 2 (gender: girls and boys) x 3 (literacy strand: read, write, communication) ANOVA was conducted with children's average literacy report card grades as the dependent variable. There was a main effect of literacy strand [$F(1,60) = 25.840$, $p < .001$, $\eta^2 = .301$] and a main effect of gender [$F(1,60) = 9.918$, $p = .003$, $\eta^2 = .142$]. These main effects were qualified by a significant interaction between literacy strand and gender [$F(1,60) = 5.785$, $p = .019$, $\eta^2 = .658$]. Girls outperformed boys on reading grades ($M_{girls} = 8.618$, SD = 1.623; $M_{boys} = 7.178$, $SD = 1.863$), writing grades ($M_{girls} = 8.010$, $SD = 1.908$; $M_{boys} = 6.550$, $SD = 1.903$), and communication grades ($M_{girls} = 8.235$, $SD = 1.799$; $M_{boys} = 7.428$, $SD = 1.442$) (See Figure 8). A post hoc ANOVA of the literacy strand by gender interaction revealed a significant effect of gender for reading grades ($p = .002$) and for writing grades ($p = .003$). There was also a significant interaction of literacy strand and grade [$F(1,60) =$

11.030, $p = .002$, $\eta^2 = .155$]. Second grade children had higher literacy report card grades than fifth graders in all literacy strands except for communication (See Figure 9). A post hoc ANOVA of the literacy strand by grade interaction revealed no significant simple main effects for the literacy strands.

Repeated measures ANOVA was also conducted with children's average mathematics grades. A 2 (grade: 2 and 5) x 2 (gender: girls and boys) x 5 (mathematics unit: number sense, measurement, geometry, algebra, data management) ANOVA was conducted with children's average mathematics report card grades as the dependent variable. There was a significant main effect of gender [$F(1,58) = 5.420$, $p = .023$, $\eta^2 = .085$) in which girls ($M = 8.712$, $SD = .359$) outperformed boys ($M = 7.499$, $SD = .378$). There was also a significant interaction between mathematics strand and grade ($F(1,58) = 14.147$, $p < .001$, $\eta^2 = .196$]. A visual inspection of the interaction revealed that grade two children had higher report card grades than grade five children on all mathematics units except for data management (See Figure 10).

Correlational analyses were also conducted to determine if there was a significant relationship between scores on the COAT questionnaires and mathematical performance. A negative correlation may indicate the effect of gender stereotype attitudes on mathematical performance. The three COAT total proportionate scores were used for these correlations. The correlational analyses revealed a significant negative correlation between grade two boys' COAT-PM Feminine proportionate score and their performance on the WJ Test 10 Applied problems ($r = -.506$, $p = .032$), indicating that grade two boys who had a higher score on the Applied problems subtest had a lower level of self-endorsement on the feminine items of the COAT-PM. There was a significant positive relationship between grade five boys' COAT-PM Masculine proportionate score and their performance on the WJ Test 5 Math calculations ($r = .894, p < .001$), indicating that grade five boys who had a higher score on the Math calculations subtest, had a higher level of self-endorsement on the masculine items of the COAT-PM There were no other significant correlations among the total proportionate scores of the COAT-AM, COAT-PM, and the WJ subtests. This indicates that gender stereotype attitudes were related to mathematics performance, but only for the boys.

Parents' and Teachers' Perceptions

The third and fourth hypotheses explored the relationship between parents' and teachers' stereotypical beliefs. The responses from the parent questionnaires on the three literacy items were averaged to create a mean parent literacy score for each child. Similarly, parents' responses on the six mathematical items were also averaged to create a mean parent numeracy score for each child. The same was done for the teachers' responses (means, standard deviations and ranges are presented in Table 9).

To test the hypotheses that parents and teachers would report gender stereotypical beliefs pertaining to children's abilities, first a repeated measures 2 (adults: parents and teachers) x 2 (child's gender: male and female) x 2 (child's grade: 2^{nd} and 5th) ANOVA was conducted for the adults' literacy ratings. The dependent variable was the parents' and teachers' average literacy ratings of the children. There was a significant main effect of adult's literacy rating [$F(1, 60) = 4.790, p = .033, \eta^2 = .074$]. Parents' ratings ($M = 3.289$, $SD = .787$) were significantly higher than teachers' literacy ratings ($M = 3.120$, $SD = .899$). No other significant effects were found for adults' literacy ratings.

A repeated measures 2 (adults: parents and teachers) x 2 (child's gender: male and female) x 2 (child's grade: 2nd and 5th) ANOVA was also conducted for the adult's average numeracy ratings of the children's abilities. There was a significant interaction between adults and gender [$F(1, 60) = 5.040$, $p = .028$, $\eta^2 = .077$], in which parents and teachers rated girls' mathematics abilities ($M_{parents} = 3.128$, $SD = .850$; $M_{teachers} = 3.322$, $SD = 1.025$) higher than boys' mathematics abilities ($M_{parents} = 3.008$, $SD = .843$; $M_{teachers} = 2.924$, $SD = .947$), but teachers rated girls' abilities higher than parents rated girls mathematics abilities (see figure 11). No post-hoc analyses of interactions were conducted as each variable has two levels. No other significant effects were found for adults' numeracy ratings.

Correlational analyses were included between the adults' questionnaire data and the rabbit family task. There were a number of significant correlations between children's rankings of the Rabbit family and parents' and teachers' ratings of the children's abilities. In the rabbit family task, a lower rank (closer to one) was indicative of a greater perception of ability, therefore a positive correlation between adults' ratings and the Rabbit family rankings would demonstrate that a higher adult competency rating is related to a lower competency ranking of the Rabbit child.

This was demonstrated through a number of positive correlations with adults' ratings and the Rabbit family, in which higher adult ratings were associated with lower rankings of the Rabbit child. These positive correlations are opposite of what would be expected from children who have been influenced by adults' gender stereotypical beliefs. It was expected, if children had been influenced by adults' gender stereotypical beliefs, that boys would be ranked higher in mathematics by the adults and this rating would be related to their ranking of the Brother rabbit's mathematics ability. Similarly, girls who had been subjected to adults' stereotypical beliefs, would be rated higher in reading and this rating would be correlated with their rankings of the Sister rabbit's reading ability. This would indicate that parents and teachers gender stereotypical beliefs are related to and may influence children's own perceptions.

The parent literacy rating of grade two boys was positively correlated to children's rankings on the Neutral Father-Brother reading dyad ($r = .589, p = .010$). This indicates that younger boys who were rated higher in literacy thought the Father rabbit would rate the brother's reading ability lower (closer to 4).

The teachers' literacy ratings were positively correlated with grade two boys' rankings of the Neutral Father-Brother reading dyad ($r = .599, p = .009$), and grade two boys' rankings of the Neutral Mother-Sister mathematics dyad ($r = .475, p = .046$), which indicates that younger boys who were rated more competent in literacy, thought the Brother would be ranked less competent in reading by the Father, and the Sister ranked less competent in mathematics by the Mother.

Teachers' literacy ratings were also positively correlated with grade five boys' rankings of the Gendered Mother-Sister mathematics dyad ($r = .682, p = .015$), which indicates that boys who were rated higher in literacy by their teacher thought the Sister would be ranked less competent in mathematics by the Mother.

Parents' numeracy ratings were positively correlated to grade two girls' rankings of the Gendered Mother-Brother mathematics dyad ($r = .477, p = .039$), which indicates that girls who were rated higher by parents in numeracy activities thought the Brother would be ranked lower in mathematics by the Mother. There was also a negative correlation which opposed gender stereotype attitudes, in which teacher literacy ratings were negatively correlated to grade five girls' rankings of the Gendered Mother-Brother reading dyad ($r = -.524, p = .045$), which indicates that girls who were rated as being more competent in literacy activities thought the Brother would be ranked higher in reading by the mother.

The only correlation which suggested the gender stereotype of boys being more mathematically competent was with teacher numeracy ratings in which teacher numeracy ratings were negatively correlated to grade two boys' rankings of the Neutral Mother-Brother mathematics dyad ($r = -.560, p = .016$), which shows that boys who were rated as having a higher mathematics ability by their teachers thought the Brother would be ranked higher in mathematics by the Mother.

These correlations demonstrate that adults did not seem to have influenced children's perceptions in a stereotypical way. These correlations are consistent with the results from the adults' questionnaires, as adults rated girls more mathematically competent than boys and there were no gender differences in adults' ratings of the children's literacy abilities. Therefore, the adults did not seem to have stereotypical beliefs concerning children's mathematical abilities, and children's perceptions were not biased in a gender stereotypical manner. The only exception to this was the grade two boys, as their teachers' rating was related to their perceptions of the Brother's mathematics ability.

Correlations were also performed between adults' literacy and numeracy ratings and children's responses on the COAT questionnaires. The three total proportionate COAT scores were used for these correlational analyses. There were positive correlations between grade five boys' COAT-PM masculine score with the parent numeracy rating ($r = .626, p = .030$), the teacher literacy rating ($r = .637, p = .026$), and the teacher numeracy rating ($r = .705, p = .011$). These correlations indicate that older boys, who were rated more mathematically competent by parents and teachers, demonstrated greater self-endorsement on the masculine items of the COAT-PM. These correlations between the adult questionnaire data and both the Rabbit family and COAT questionnaires seem to indicate boys' perceptions and attitudes were related to adults' ratings in a manner that was more consistent with gender stereotypes, which was not the same for the girls in this study.

The fifth hypothesis explored the relationship between parents' and teachers' stereotypical beliefs. There were positive correlations between the parents' average literacy ranking ($M = 3.290$, $SD = .787$) and the teachers' average literacy ranking ($M = 3.120$, $SD = .899$) [$r = .627, p < .001$], as well as between parents' average numeracy ranking ($M = 3.071$, $SD = .842$) and teachers' average numeracy ranking ($M = 3.136$, $SD = 1.002$) [$r = .802, p < .001$], which indicates that parents and teachers were consistent with each other on their ratings of children's abilities. There were also significant positive correlations between parents' and teachers' ratings of the children's abilities with performance on the WJ III ACH (see table 10), as well as between adults' ratings and children's report card grades (see table 11), indicating that parents and teachers were well informed of the children's abilities, and that they were not basing their ratings on gender stereotype beliefs.

Discussion

The goal of the current study was to examine gender stereotype internalization in elementary school age children and how internalization affects mathematics performance. The findings of the current study demonstrate that internalization of the gender stereotype that boys are more mathematically competent than girls was not evident in either second or fifth grade participants. There were no gender differences in how children ranked their perception of their mathematics abilities or how parents and teachers ranked children's mathematics abilities. Additionally, girls did not underperform on the mathematics subtests of the Woodcock-Johnson in comparison to boys, which is what would be expected if girls had internalized mathematics-related gender stereotypes.

In contrast to these main findings related to mathematics, there was evidence of the effects of gender stereotypes pertaining to boys' literacy activities (i.e., completing word passages). Specifically, there was endorsement of the gender stereotype that boys are less competent in literacy than girls. This was unexpectedly demonstrated by the younger boys in the rankings of their own reading abilities, which were lower than girls. All the boys in the current study underperformed in comparison to girls, on the passage comprehension test and the mathematics word problems test, which were the tests that required the highest level of reading comprehension. Boys' perceptions of their literacy abilities were not consistent with parents' and teachers' beliefs, as there were no gender differences in adults' ratings of children's literacy abilities.

The results of the current study suggest that the internalization of gender stereotypes pertaining to mathematics had not occurred with these children. However, it is possible that boys had internalized gender stereotypes in the literacy domain. These findings suggest that there should be an increased focus on the education of boys to address the gender gap in literacy competence. This issue has been recently identified both in popular and research literature (Below, Skinner, Fearrington, & Sorrell, 2010; Marinak & Gambrell, 2010).

Children's Perceptions

The first hypothesis examined gender stereotype internalization exhibited by second and fifth grade children. It was hypothesized that the older fifth grade children would have internalized gender stereotype attitudes, but that the younger second grade children would not. This hypothesis was only partially supported by children's responses on the personal measure of the COAT questionnaire. Children in both grades demonstrated higher self-endorsement of stereotypical items according to their own gender (i.e., a boy rated his desire to be a policeman higher than a cheerleader). Conversely, children did not report stereotypical perceptions of their own mathematics abilities, nor did they display gender stereotypes concerning others on their responses to the attitude measure of the COAT questionnaire. However, grade two boys did report lower perceptions of their reading ability, indicating that although girls may not have been influenced by mathematics gender stereotypes, boys seem to have negative beliefs regarding their literacy abilities.

The lack of gender stereotypical perceptions and attitudes found in the current study may be reflections of the knowledge young children have about the stereotypes, and how these stereotypes develop over time. Trautner and colleagues (2005) suggested that there are three primary steps of gender stereotype development which include (1) beginning awareness, (2) rigidity, and (3) flexibility. Beginning awareness is similar to the previous definition of gender stereotype knowledge. Stereotype rigidity is defined by gender stereotypes held in a rigid fashion and a lack of individual variation in levels of masculinity and femininity. A child who holds rigid gender stereotypes would say that a boy would want to play with trucks since being a boy is associated with masculinity (Trautner et al., 2005). Children reach a peak level of rigidity between five and seven years of age. The phase of stereotype flexibility begins at about the age of eight years, and this is when children have knowledge of gender stereotypes but are able use the response "both" to questions about who can or who should do stereotypical activities (Trautner et al., 2005). Although children at this age are aware of gender stereotypes, they also know that there is variation in how they can be applied. Children typically reach a peak level of flexibility between 10 to 12 years of age (Trautner et al., 2005).

These primary steps of gender stereotype development could possibly explain our current findings with this group of participants because the children fall between the age ranges of both stereotype rigidity and flexibility. Gender stereotype flexibility could explain grade five children's lack of gender stereotype attitudes. The fifth grade children were between the ages of 10 to 11 years, which is the range that children are said to reach the peak level of flexibility. Therefore, they may have had knowledge of gender stereotypes, but were applying a high level of gender stereotype flexibility in their responses to the tasks.

Similarly, gender stereotype flexibility could also help explain grade two children's perceptions of their reading ability, as well as the lack of internalization of gender stereotypes demonstrated through the other tasks. Since the second graders were between the ages of seven and eight years, these children would fall between high rigidity and the beginning of flexibility, and therefore variation may exist in their knowledge of gender stereotypes. Gender stereotypes pertaining to academics have been shown to develop later than other gender stereotypes (i.e., items, clothing, activities) (Blakemore et al., 2009; Leinbach & Fagot, 1993; Six & Eckes, 1991), and as such, grade two children may not yet have graduated into the flexibility stage of literacy-related gender stereotypes, which appeared to be the more salient stereotype for children of this age. However, because gender stereotypes of occupation, activities, and traits are more common and emerge earlier, the children were able to apply flexibility to these gender stereotypical items, and therefore did not display gender stereotypical attitudes.

Flexibility in regards to gender stereotypes is debatable because it is often measured by the response 'both" to stereotypical items. However, the inclusion of the "both" response has multiple interpretations; it could mean either that the child was responding in a flexible way to the item, or that they did not have knowledge of the gender stereotype (Banse, Gawronski, Rebetez, Gutt, & Morton, 2010). This is an important implication of the current study: it is possible that children have not internalized gender stereotypes at this age. Although it has been evidenced that internalization affects mathematics performance of girls in high school (Bleeker & Jacobs, 2004; Dick & Rallis 1991; Else-Quest et al., 2010; Hyde et al., 1990), it is still unknown when children begin to internalize this stereotype. There is little existing research which focuses solely on children's internalization of gender stereotypes pertaining to mathematics. Therefore, it is possible that internalization of mathematics gender stereotype occurs later in elementary and middle schools, perhaps between grades six and eight; more research would be needed to examine this issue.

An unexpected finding of the current research was that early elementary school-age boys were less confident in their reading abilities than girls and their literacy ability was also lower than girls. Although boys underperformed in comparison to girls on only one of the two Woodcock-Johnson literacy subtests, this subtest required the children to read a passage and correctly fill in a missing word and so a higher level of reading ability was necessary to do well. Similarly, the Woodcock-Johnson mathematic subtest which boys also underperformed in comparison to girls was the test which required a higher level of literacy ability since it involved mathematical word problems. These results were found for all boys in the current study. The boys also had significantly lower reading, writing, and communication report card grades than girls. It is possible that the younger boys' perceptions of their literacy abilities were related to their actual ability or to self-efficacy in literacy subjects, and were not influenced by gender stereotypes. Further research would need to be conducted to examine the relationship between boys' perception of their literacy abilities, self-efficacy in literacy, and actual literacy abilities, to see if boys' lower perceptions of their abilities emerge prior to gender differences in actual literacy abilities.

Children's Actual Abilities

The second hypothesis of the current study examined the relationship between children's mathematics performance and gender stereotype internalization. It was expected that at the fifth grade level, boys would be mathematically outperforming girls. This hypothesis was not supported; there were few gender differences in children's actual abilities. Where differences did exist on both the Woodcock-Johnson and mathematics report card grades, girls seemed to be the more academically competent gender. This indicates that gender stereotype attitudes did not affect children's mathematics performance.

Although girls did not perform in a gender stereotypical way on the mathematics tests, boys performed in a stereotypical way on the literacy tests, by underperforming in comparison to girls on two subtests of the Woodcock-Johnson and report card grades.

These findings regarding children's mathematics and literacy abilities may have been affected by the composition of the sample. It is important to note that the mean grade estimates for the fifth grade children were at least one grade level behind on three of the Woodcock-Johnson subtests including Test 5: Calculations, Test 6: Math Fluency, and Passage Comprehension (See table 8 for grade estimates). Given that the students were at the end of grade five, their grade estimates should have been closer to six. The mean grade estimates for the second grade students were all two or above. The low grade estimates indicates that the fifth grade students may have been low-achieving. Furthermore, age differences were not expected in children's report card grades, since grades are generally based on achievement in comparison to same age peers and the appropriate grade curriculum. However, significant age differences were found in children's report card grades, with younger children receiving higher report card grades on two of the three literacy units and four of the five mathematics units. The findings regarding report card grades, as well as the low grade estimates on the Woodcock-Johnson, suggest that the grade five participants may have been comprised of lower achieving students.

Parents' and Teachers' Perceptions

The third and fourth hypotheses, which explored parents' and teachers' stereotypical beliefs of children's academic abilities, were also not supported as neither parents nor teachers rated one gender more mathematically competent than the other. It was expected that both parents and teachers would rate boys' mathematics abilities higher than girls.

The last hypothesis of the current study was not supported. The parents and teachers did not display stereotypical beliefs of the children's abilities, so although adults' ratings were positively correlated, they were also correlated with children's performance on the mathematics and literacy subtests and report card grades. This suggested that adults were well aware of children's abilities and that their ratings of children's abilities were not influenced by gender stereotypes.

The lack of stereotypical beliefs demonstrated by the adults was unexpected, however may have been affected by the time of school year. The parents and teachers may have been able to recall recent feedback regarding children's academic abilities at the time they completed the questionnaires. In previous research, parents and teachers completed similar questionnaires of three-, four-, and five-year-old children's reading and mathematics abilities and these adults rated boys more mathematically competent than girls, even at three years of age (Lee & Schell, under review). Data collection for this younger age cohort took place prior to distribution of the first report card of the year. At this stage, parents may have been less knowledgeable about their child's current progress in mathematics and reading, so gender stereotype attitudes may have become a bigger factor in their ratings.

The parents and teachers included in the current study completed the questionnaires in the month of June. At this point in the school year, parents would have received report cards for two school terms. Additionally, teachers were in the process of completing the third report card while data collection was taking place. The parents and teachers were most likely able to rely more on their recent knowledge of the child's abilities. Perhaps if the teachers, and maybe to a lesser extent for parents, had completed the questionnaires at the beginning of the school year, when they did not have current information regarding the child's abilities, they may rely on other social cognitive resources such as gender stereotypes when rating children's abilities.

There is no existing research which considers the time of school year where parents' and teachers' complete their ratings of children's academic abilities. Future research could compare parents' and teachers' ratings both at the beginning of the school year and again at the end of the school year to examine if these ratings change over the course of the year, or perhaps parents and teachers are simply more cognizant of children's abilities at the early- to mid-elementary school level than when the children are in preschool and kindergarten.

Limitations

There were a few limitations of the current study. The first limitation was the sample size of the study. Due to various recruitment issues and time constraints, all the participants in the study were recruited from one elementary school. The data collection sessions occurred at the end of the school year in June, and additional schools could not be approached to participate because schools and the school board have the policy not to permit research studies close to the end of the academic year. To maintain consistency in terms of the data collection time, the recruitment of participants was not continued in the new academic year in September to avoid confounding performance on the four tasks by students who might be affected by the summer vacation months, the adjustment to a new school year and new classroom teacher. Moreover, the current study was prematuredly terminated due to some issues associated with our follow-up study at the school board level.

There were only two grade five classes at this school, which resulted in a smaller sample of grade five children. There were only 12 fifth grade boys who took part in this study. Therefore, the lack of significant findings may be due to the explanations listed above, or to a sample size that was too small to reveal an effect. Power analysis revealed a range from .05 to .92, which indicates that for some of the analyses there was not sufficient power to detect an effect. For example, a larger sample of grade five participants could support whether older boys were actually equally confident in their reading abilities, or if they too may be affected by gender stereotypes in literacy

Another potential limitation of the study was the use of the COAT questionnaires and the picture cards task for the older children. The picture cards may have not been the most accurate measure of academic stereotypes for the older children since some of the activities depicted may have been too juvenile, such as playing with Lego blocks. Cards that illustrate academic subjects, such as gym, art, or geography, which are more age appropriate, may yield more accurate findings regarding children's academic gender stereotype beliefs. Similarly, the COAT questionnaire may not have been sensitive enough to measure gender stereotype internalization in mathematics. The COAT questionnaires measured gender stereotype attitudes on a more general level, and therefore may not have been specific enough to ascertain children's attitudes concerning gender stereotypes in mathematics.

Other potential limitations to the current study included the length of the child tasks and gender of the researchers. The children completed the four tasks in two sessions with each of the sessions taking approximately 45 to 60 minutes to complete. The children could have gotten bored or fatigued during the sessions, especially in regards to the mathematics and literacy tests during which their actual abilities were being tested. Attempts to combating possible fatigue and boredom were made but were difficult in this case, since the researchers were trying to keep the length of the sessions to a minimum, while limiting the disruptions to the child's classroom routine since the child needed to be absent from class. Adding another session would be disruptive to the child's learning in the classroom and may have not increased the quality of the data enough to warrant another disruption.

Another concern was that both of the researchers were female. This may have affected how children rated their own abilities, whether the female researchers increased or decreased their comfort level. Future research could counter balance the effect of researcher by including one female and one male researcher, however the majority of the elementary school teachers in the current sample, and in general, were female, so female researchers could not have adversely affected the children's performance on any of the tasks.

Future Research

The next step in research regarding gender stereotypes internalization would be to gather second, fifth, sixth, and eighth grade students. Findings including this older age group may indicate an alternative path of gender stereotype internalization than the current study suggests. Findings that are consistent with the current study would provide further support for Trautner and colleagues' (2005) primary three steps of gender stereotype development, and gender stereotype flexibility in children at this age range.

Future research in gender stereotype internalization pertaining to mathematics should also examine the gender stereotype attitudes of children in late elementary school. For example, researchers would be able to determine when children internalize attitudes and also how quickly this process occurs. A number of studies have demonstrated that by the time girls reach high school they believe that mathematics is a male-dominated field, and they show less interest in mathematics (Bleeker & Jacobs, 2004; Dick & Rallis, 1991; Else-Quest et al., 2010; Hyde et al., 1990). Research on gender stereotype attitudes of older children could pinpoint a specific age at which internalization occurs, and could also illustrate how quickly the process takes place. Liben and Bigler (2002) demonstrated that children at sixth grade had internalized the general gender stereotypes included in the COAT questionnaire, so if future research included fifth grade students who do not display stereotype attitudes, but children in sixth grade did, then it would be apparent that gender stereotype internalization occurs in the sixth grade and that it is a rapid process in which children internalize gender stereotypes over the course of only one year.

Additionally, further research with a focus on boys' literacy development is a need that has been made apparent in the current findings. Research with preschool children should examine early perceptions of literacy abilities, as well as actual literacy competence, to see if gender differences in perception emerge prior to gender differences in literacy performance. Furthermore, this research should continue to following boys' literacy development to examine if gender differences in literacy persist into high school and post secondary education. The Association of Universities and Colleges of Canada (AUCC) (2010) presented that women outnumbered men in post-secondary enrolment by 57% to 43%, and that in 2008 females comprised a greater number of graduates from all levels with the exception of the doctoral level. Further research should examine how boys' literacy development affects later academic achievement, including enrolment in post-secondary education.

The last suggestion for future research would be the most optimal but also the most difficult to complete. The most effective and accurate means of creating developmental trajectories is through longitudinal designs. In order to map a developmental trajectory of gender stereotype internalization, longitudinal research could follow children from preschool age, when adults have been known to display stereotypical beliefs for children's abilities, through elementary school when internalization is thought to take place. This line of research would be very time consuming and difficult to conduct. Cross-sectional research should primarily be conducted prior to longitudinal to fully examine the process of internalization, but longitudinal data would be very beneficial in creating a developmental trajectory of gender stereotype knowledge and attitudes.

Conclusion

The current study provides a starting point for further research on the influence of gender stereotype internalization on children's academic abilities. It appears that girls up to grade five are on par or even performed better than their male peers in both literacy and mathematics domains. The lack of gender stereotypical attitudes found in the current study differs from the majority of existing literature. This may be a result of children's actual knowledge being more salient than the gender stereotype. Children may have been aware of their own and their peers' achievement, which resulted in their perceptions being based on reality, instead of stereotypes. Older children may partake in less group work with their peers and more individual school work, therefore making them less knowledgeable regarding their peers' academic achievement. It is possible that older children's perceptions would then be based more on gender stereotypes instead of their existing knowledge. Given the potential impact that gender stereotypes have shown in previous research, the question of how stereotypes impact mathematics performance and beliefs about mathematical ability warrants further investigation.

Table 1

Means, standard deviations, and ranges of children's rankings on the Picture Cards task

Grade			N	Min	Max	*Mean (SD)*
2	Girls	Read	19	1	7	2.84 (2.36)
		Math	19	1	7	3.37 (1.77)
	Boys	Read	18	3	7	4.50 (1.58)
		Math	18	1	6	3.00 (2.14)
5	Girls	Read	15	1	7	3.27 (2.02)
		Math	15	1	7	3.67 (2.06)
	Boys	Read	12	2	7	3.67 (1.56)
		Math	12	1	7	3.42 (2.11)

Table 2

Gender Neutral and Gender Stereotypic Rabbit Family means and standard deviations

			Grade 2s		Grade 5s	
			Boys	Girls	Boys	Girls
			M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Neutral	Dad-Bro	Read	2.44	2.74	2.75	3.07
			(0.92)	(1.05)	(0.87)	(0.80)
		Math	2.28	2.05	2.92	2.40
			(1.18)	(0.85)	(1.24)	(1.12)
	Dad-Sis	Read	1.83	2.53	2.17	1.87
			(0.86)	(1.22)	(1.03)	(0.74)
		Math	2.44	2.68	2.42	2.67
			(1.10)	(0.89)	(1.08)	(1.05)
	Mom-Bro	Read	3.00	2.68	2.83	2.40
			(1.09)	(0.95)	(0.94)	(1.18)
		Math	2.50	2.89	2.75	2.73
			(1.10)	(1.05)	(1.14)	(0.96)
	Mom-Sis	Read	1.94	2.26	2.08	1 73
			(1.11)	(0.87)	(0.10)	(0.80)
		Math	2.11	2.26	2.50	2.80
			(0.96)	(1.10)	(1.00)	(0 94)
Gendered	Dad-Bro	Read	2.83	2.37	2.83	2.60
			(0.99)	(1.12)	(1.12)	(0.83)
		Math	2.28	2.47	2.25	2.27
			(1.07)	(1.17)	(1.06)	(1.16)
	Dad-Sis	Read	1.72	2.21	2.42	2.33
			(0.90)	(0.92)	(1.08)	(0.90)
		Math	2.28	2.37	2.17	2.40
			(1.18)	(0.96)	(1.03)	(0.91)
	Mom-Bro	Read	2.78	2.79	2.75	2.47
			(1.11)	(1.08)	(1.06)	(0.83)
		Math	2.50	2.11	2.92	2.13
			(1.25)	(1.10)	(1.24)	(1.19)
	Mom-Sis	Read	2.06	2.26	2.33	1.80
			(1.11)	(1.10)	(1.07)	(0.78)
		Math	2.17	2.37	2.25	2.47
			(0.99)	(1.12)	(1.06)	(1.13)

Table 3

Mann-Whitney statistics for the children's Rabbit family rankings

Grade	Rabbit family	Rabbit family dyad	Activity	Mann-Whitney *U*	*p-value*
2	Neutral	Father-Brother	Read	139.500	.327
			Math	154.000	.614
		Father-Sister	Read	116.500	.087
			Math	147.000	.449
2	Neutral	Mother-Brother	Read	138.000	.310
			Math	136.000	.274
		Mother-Sister	Read	131.000	.200
			Math	159.000	.731
5	Neutral	Father-Brother	Read	70.000	.356
			Math	69.000	.323
		Father-Sister	Read	75.500	.510
			Math	78.500	.578
5	Neutral	Mother-Brother	Read	69.500	.342
			Math	88.000	.921
		Mother-Sister	Read	72.500	.407
			Math	73.000	.389
2	Gendered	Father-Brother	Read	129.500	.205
			Math	154.000	.631
		Father-Sister	Read	118.000	.097
			Math	159.500	.735
2	Gendered	Mother-Brother	Read	170.500	.998
			Math	141.000	.364
		Mother-Sister	Read	152.000	.571
			Math	152.500	.612
5	Gendered	Father-Brother	Read	75.000	.446
			Math	89.500	1.000
		Father-Sister	Read	85.000	.799
			Math	74.000	.425
5	Gendered	Mother-Brother	Read	73.000	.397
			Math	58.000	.111
		Mother-Sister	Read	64.500	.215
			Math	79.500	.673

Table 4

Means and standard deviations for the COAT-AM by grade and gender

Subscale	Gr.2 Boys		Gr.2 Girls		Gr.5 Boys		Gr.5 Girls	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Fem C<u>O</u>AT-AM	.422	.196	.311	.197	.375	.238	.493	.291
Fem CO<u>A</u>T-AM	.528	.205	.505	.276	.367	.303	.513	.304
Fem COA<u>T</u>-AM	.156	.257	.174	.235	.117	.153	.113	.119
Mas C<u>O</u>AT-AM	.472	.232	.400	.254	.408	.239	.407	.222
Mas CO<u>A</u>T-AM	.589	.235	.558	.287	.475	.341	.407	.237
Mas COA<u>T</u>-AM	.156	.195	.121	.151	.100	.121	.033	.062

Note *Underscored O, A, and T indicate occupations, activities, and traits respectively*

Table 5

Means and standard deviations for the COAT-PM by grade and gender

Subscale	Gr.2 Boys		Gr 2 Girls		Gr.5 Boys		Gr.5 Girls	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Fem C<u>O</u>AT-PM	1.883	.573	2.547	.615	1.600	.372	2.313	.608
Fem CO<u>A</u>T-PM	1.711	.466	2.161	.575	1.842	.438	2.293	.448
Fem COA<u>T</u>-PM	2.794	.453	3.095	.481	2.633	.339	2.847	.612
Mas C<u>O</u>AT-PM	2.781	.618	2.366	.660	2.492	.423	2.107	.422
Mas CO<u>A</u>T-PM	2.094	.547	1.858	.578	2.100	.391	1.923	.598
Mas COA<u>T</u>-PM	2.633	.423	2.789	.374	2.900	.465	2.740	.410

Note. Underscored O, A, and T indicate occupations, activities, and traits respectively

Table 6

Means and standard deviations for the COAT-AM and COAT-PM

Subscale	Boys *M*	Boys *SD*	Girls *M*	Girls *SD*	Combined *M*	Combined *SD*
Fem COAT-AM	0.40	0.04	0.40	0.04	0.40	0.03
Fem COAT-AM	0.45	0.05	0.51	0.05	0.48	0.03
Fem COAT-AM	0.14	0.04	0.14	0.04	0.14	0.03
Mas COAT-AM	0.44	0.04	0.40	0.04	0.42	0.03
Mas COAT-AM	0.53	0.05	0.48	0.05	0.51	0.03
Mas COAT-AM	0.13	0.03	0.08	0.03	0.10	0.02
COAT-AM	0.42	0.04	0.40	0.04	0.41	0.03
COAT-AM	0.49	0.05	0.50	0.04	0.49	0.03
COAT-AM	0.13	0.03	0.11	0.03	0.12	0.02
Fem COAT-PM	1.74	0.11	2.43	0.10	2.09	0.07
Fem COAT-PM	1.78	0.09	2.23	0.09	2.00	0.06
Fem COAT-PM	2.71	0.09	2.97	0.08	2.84	0.06
Mas COAT-PM	2.64	0.10	2.24	0.10	2.44	0.07
Mas COAT-PM	2.10	0.10	1.89	0.09	1.99	0.07
Mas COAT-PM	2.77	0.08	2.76	0.07	2.77	0.05
COAT-PM	2.19	0.09	2.33	0.09	2.26	0.06
COAT-PM	1.94	0.08	2.06	0.08	2.00	0.06
COAT-PM	2.74	0.07	2.87	0.07	2.80	0.05

Note Underscored O, A, and T indicate occupations, activities, and traits respectively

Table 7

WJ III ACH Descriptive statistics of children's performance by gender and grade

Grade	Gender		N	Min	Max	*M (SD)*
2	Girls	WJ Test 5	19	58.00	119.00	95.579 (14.542)
		WJ Test 6	19	72.00	110.00	91.421 (10.875)
		WJ Test 9	19	79.00	119.00	98.105 (10.197)
		WJ Test 10	19	82.00	125.00	106.842 (13.729)
		WJ Test 13	19	89.00	122.00	105.895 (10.619)
	Boys	WJ Test 5	18	58.00	141.00	94.500 (20.077)
		WJ Test 6	18	60.00	116.00	88.389 (12.821)
		WJ Test 9	18	65.00	109.00	91.111 (10.163)
		WJ Test 10	18	55.00	125.00	94.000 (17.392)
		WJ Test 13	18	91.00	124.00	106.556 (9.697)
5	Girls	WJ Test 5	15	50.00	108.00	90.800 (14.537)
		WJ Test 6	15	66.00	103.00	87.200(10.073)
		WJ Test 9	15	79.00	117.00	95 733 (11 492)
		WJ Test 10	14	87.00	122.00	104.429 (11.092)
		WJ Test 13	15	86.00	132.00	106.667 (13.037)
	Boys	WJ Test 5	12	64.00	100.00	84.667 (11.292)
		WJ Test 6	12	74.00	111.00	84.750 (11.071)
		WJ Test 9	12	70.00	98.00	89.583 (7.141)
		WJ Test 10	12	72.00	114.00	97.167 (11.28)
		WJ Test 13	12	77.00	125.00	98.333 (13.64)

Table 8

WJ III ACH Age estimates and Grade estimates

Grade		Age Equivalent	Grade Equivalent
2	Test 5 Calculations	7.57	2.59
	Test 6 Math Fluency	7.35	2.37
	Test 9 Passage Comprehension	7.69	2.67
	Test 10 Applied Problems	8.34	3.21
	Test 13 Word Attack	9.87	4.80
5	Test 5 Calculations	10.13	4.99
	Test 6 Math Fluency	9.24	4.23
	Test 9 Passage Comprehension	9 47	4.65
	Test 10 Applied Problems	11.83	6.87
	Test 13 Word Attack	13.09	8.39

Table 9

Parents' and teachers' average rankings of children's mathematics and literacy abilities

		Adult	Rank type	N	Min	Max	*M (SD)*
Grade 2	Girls	Parent	Literacy	19	2.00	4.33	3.404 (0.624)
			Numeracy	19	1.33	4.00	3.026 (0.594)
		Teacher	Literacy	19	1.00	4.33	3.316 (0.857)
			Numeracy	19	1.33	4.50	3.160 (0.766)
	Boys	Parent	Literacy	18	1.33	5.00	3.083 (0.940)
			Numeracy	18	1.00	4.83	2.999 (0.864)
		Teacher	Literacy	18	1.33	4.67	2.981 (0.690)
			Numeracy	18	1.00	4.50	3.087 (0.795)
Grade 5	Girls	Parent	Literacy	15	1.00	5.00	3.578 (0.904)
			Numeracy	15	1.00	4.83	3.256 (1.104)
		Teacher	Literacy	15	1.67	5.00	3.578 (1.042)
			Numeracy	15	1.00	5.00	3.528 (1.281)
	Boys	Parent	Literacy	12	2.33	4.00	3.056 (0.494)
			Numeracy	12	2.00	4.50	3.021 (0.848)
		Teacher	Literacy	12	1.00	3.33	2.444 (0.653)
			Numeracy	12	1.00	5.00	2.681 (1.131)

Table 10

Correlations between adults' average literacy and numeracy rankings and children's WJ III ACH scores

WJ III ACH Subtest	Parent Literacy	Parent Numeracy	Teacher Literacy	Teacher Numeracy
WJ Test 5 Calculations	.354**	.538**	.371**	.571**
WJ Test 6 Math Fluency	.401**	.519**	.450**	.584**
WJ Test 9 Passage Comprehension	.487**	.497**	.615**	.501**
WJ Test 10 Applied Problems	.405**	.478**	.436**	.555**
WJ Test 13 Word Attack	.413**	.367**	.600**	.528**

Note. ** $p < .01$

Table 11

Correlations between parent and teacher average numeracy ratings and children's report card grades by unit and report term

Unit	Report term	Parent Average Numeracy	Teacher Average Numeracy
Number Sense and Numeration	1	.702**	.755**
	2	.759**	.835**
	3	.624**	.768**
Measurement	1	.492	.776**
	2	.709**	.706**
	3	.688**	.778**
Geometry and Spatial Sense	1	.504**	.777**
	2	.582**	.703**
	3	.665**	.799**
Patterning and Algebra	1	.568**	.693**
	2	.755**	.840**
	3	.610**	.736**
Data Management and Probability	1	.698**	.638**
	2	.681**	.824**
	3	.637**	.724**

Note. ** $p < .01$

Figure 1

Gender-neutral Rabbit family

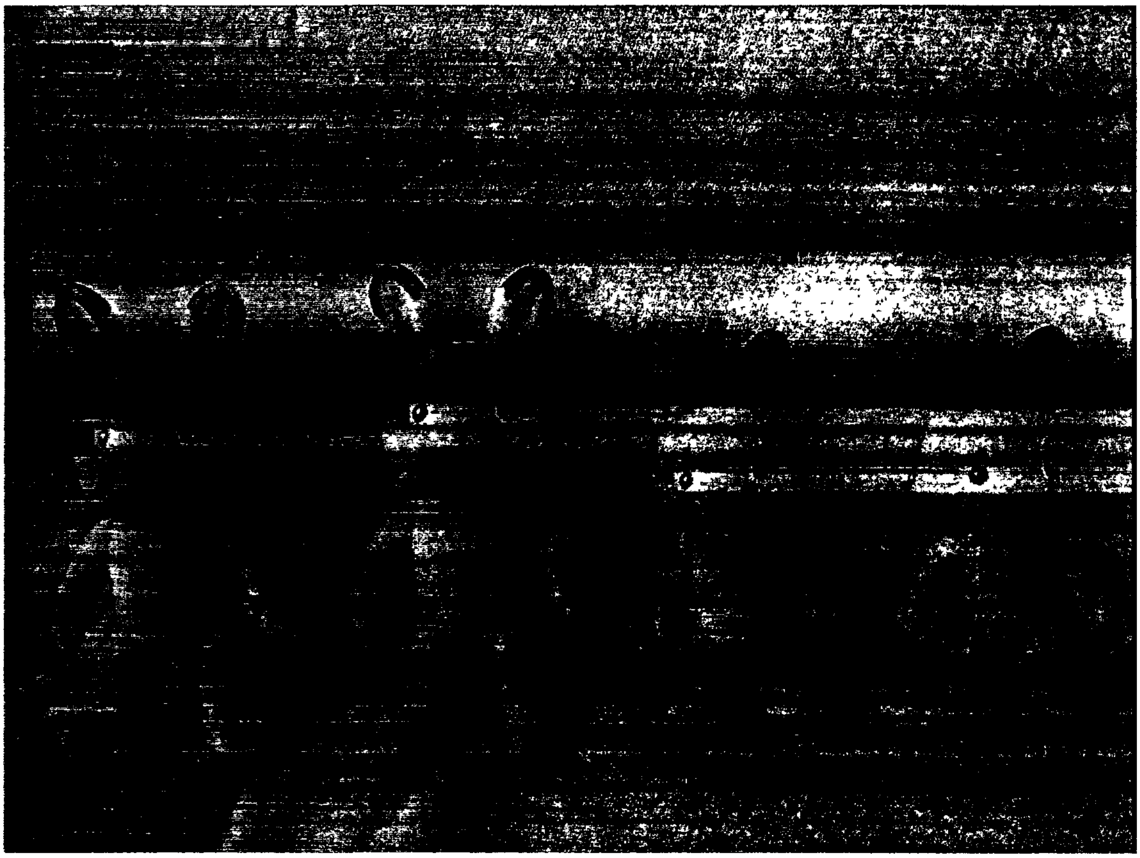


Figure 2

Gender-salient Rabbit family

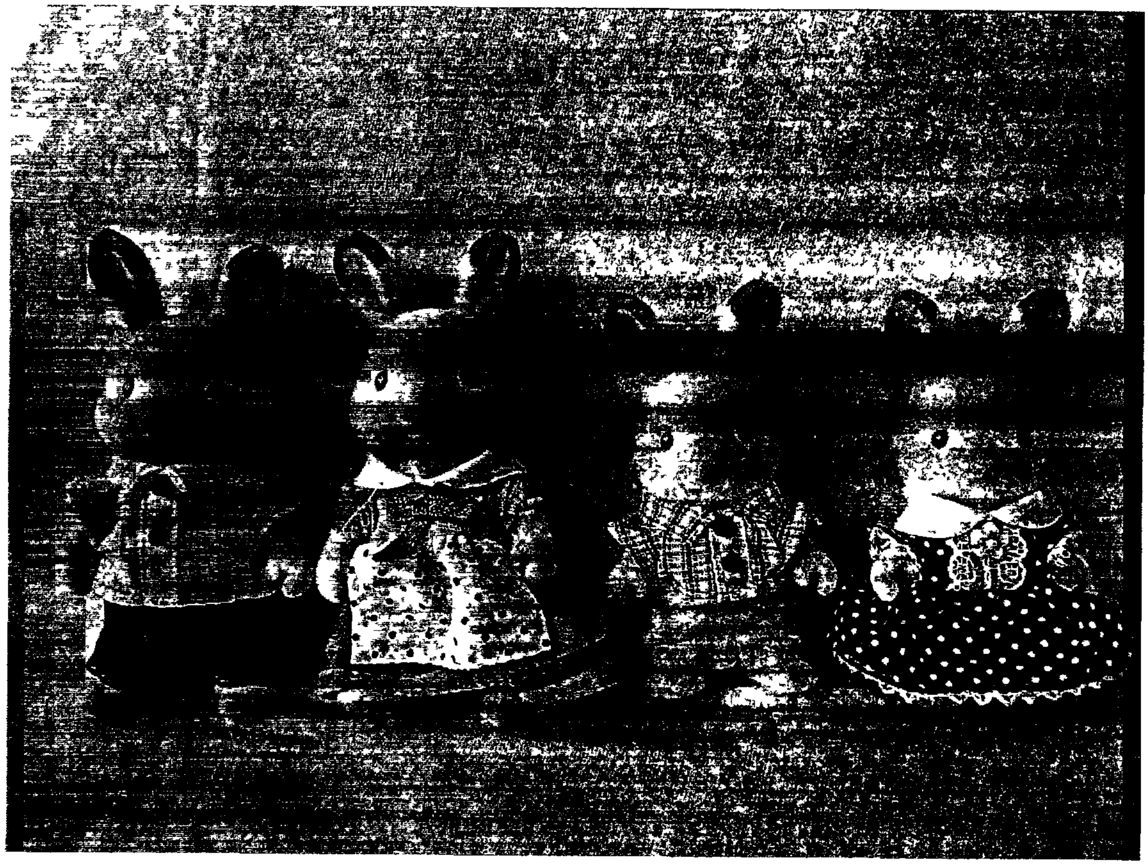


Figure 3

COAT-AM interaction. Domain x item type

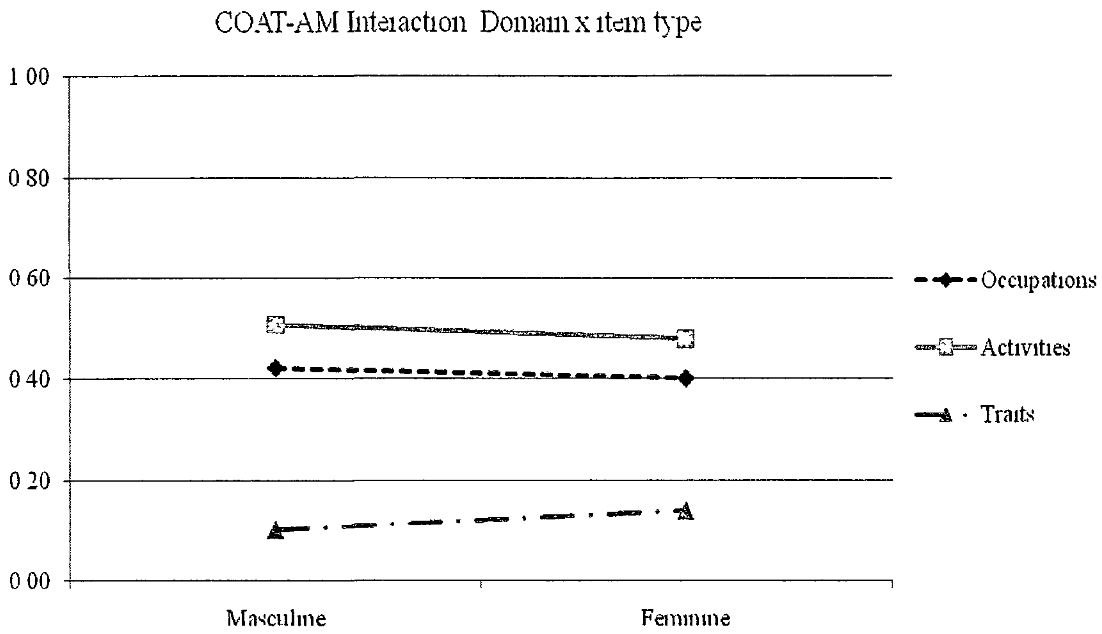


Figure 4

COAT-PM interaction Domain x item type

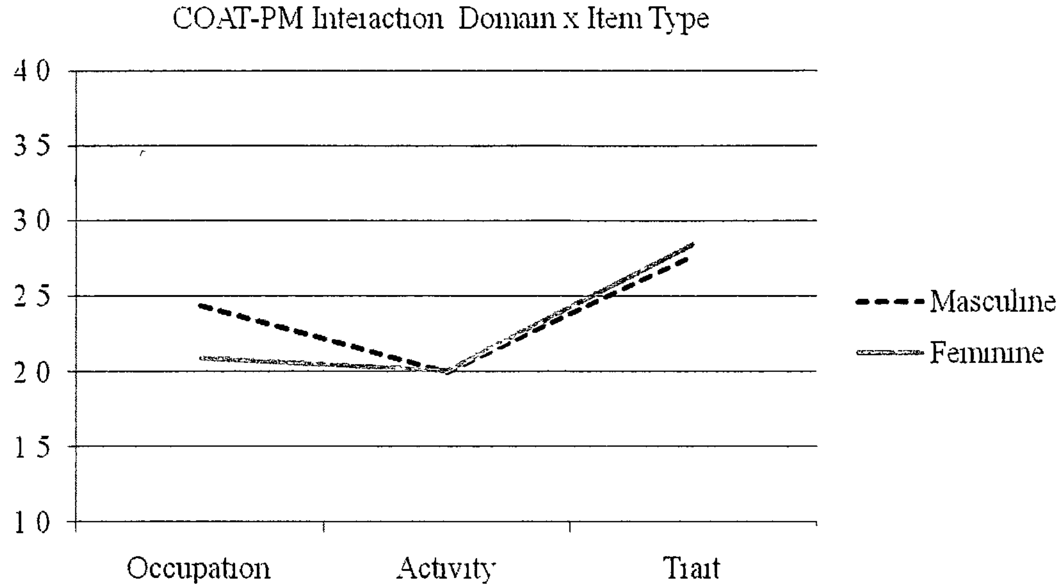


Figure 5

COAT-PM interaction Item type x gender

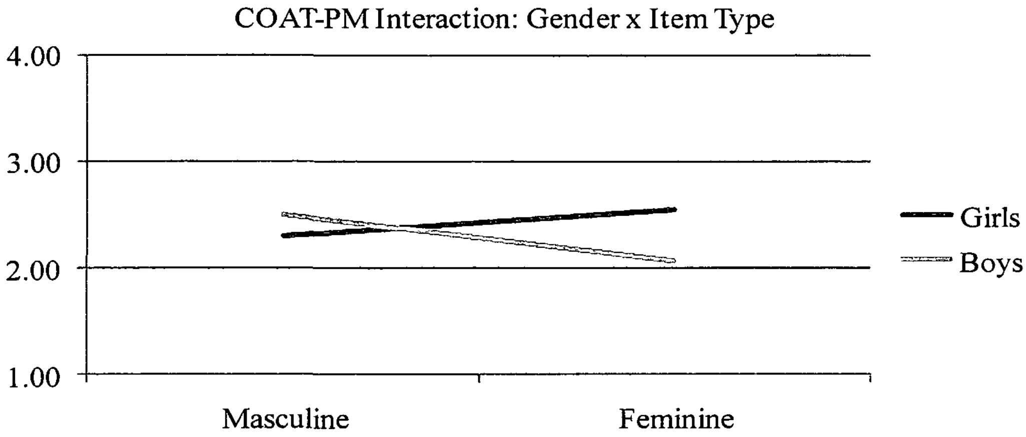


Figure 6

COAT-PM interaction Domain x item type x grade

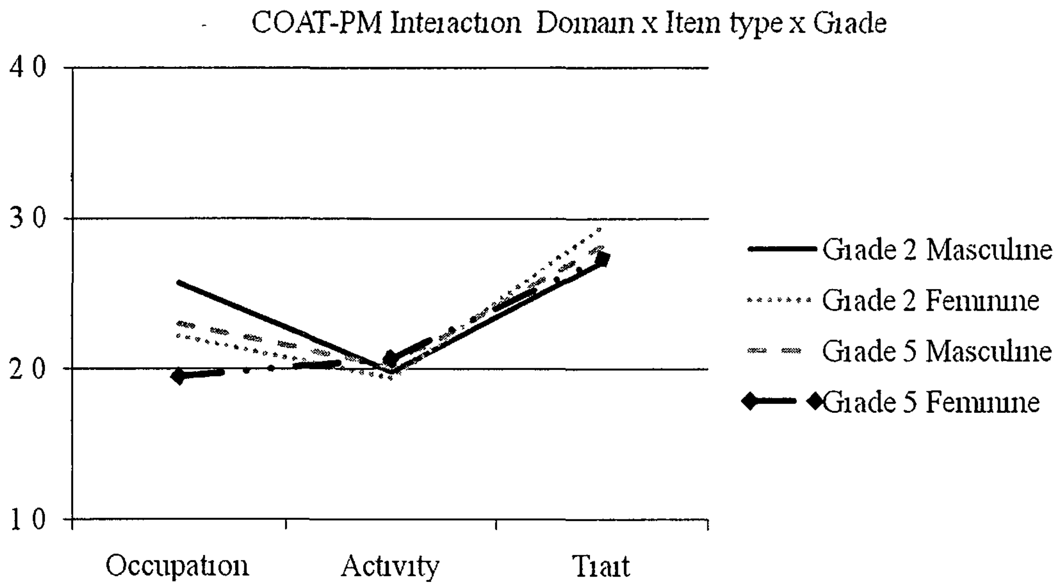


Figure 7

COAT-PM interaction Domain x item type x gender

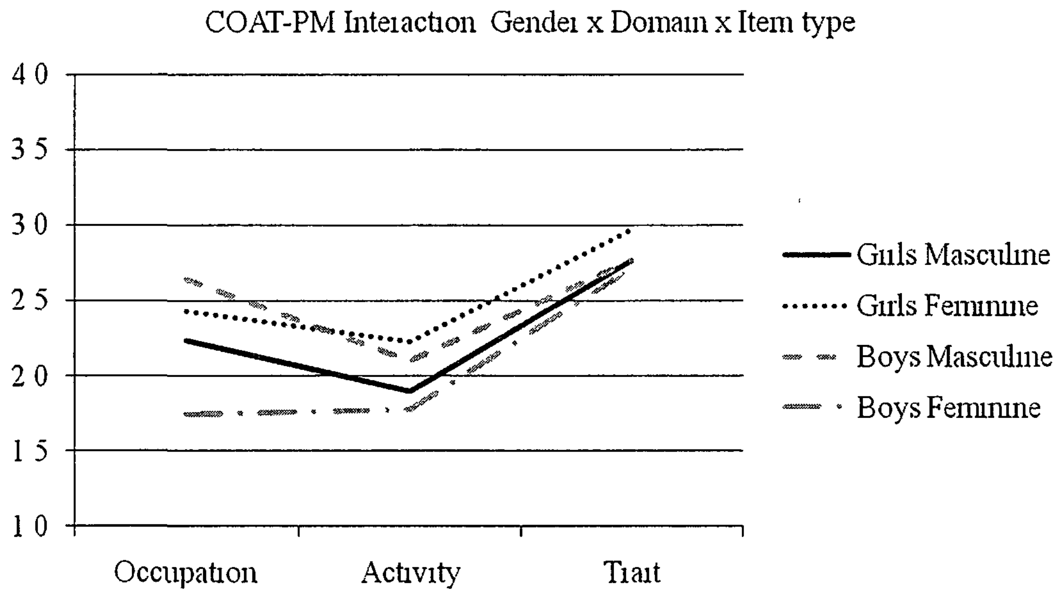


Figure 8

Children's mean literacy grades

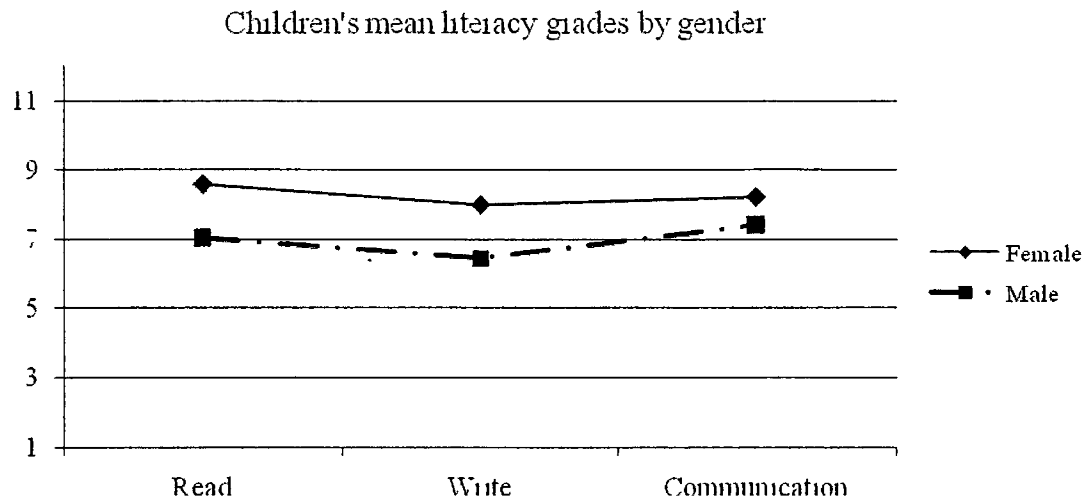


Figure 9

Children's mean literacy grades by age

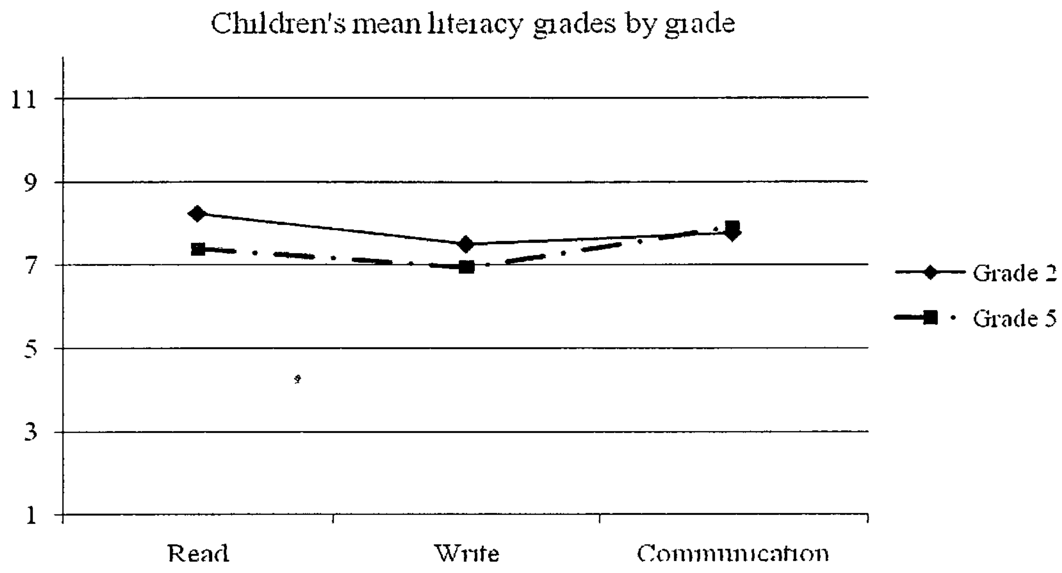


Figure 10

Children's mean mathematics grades

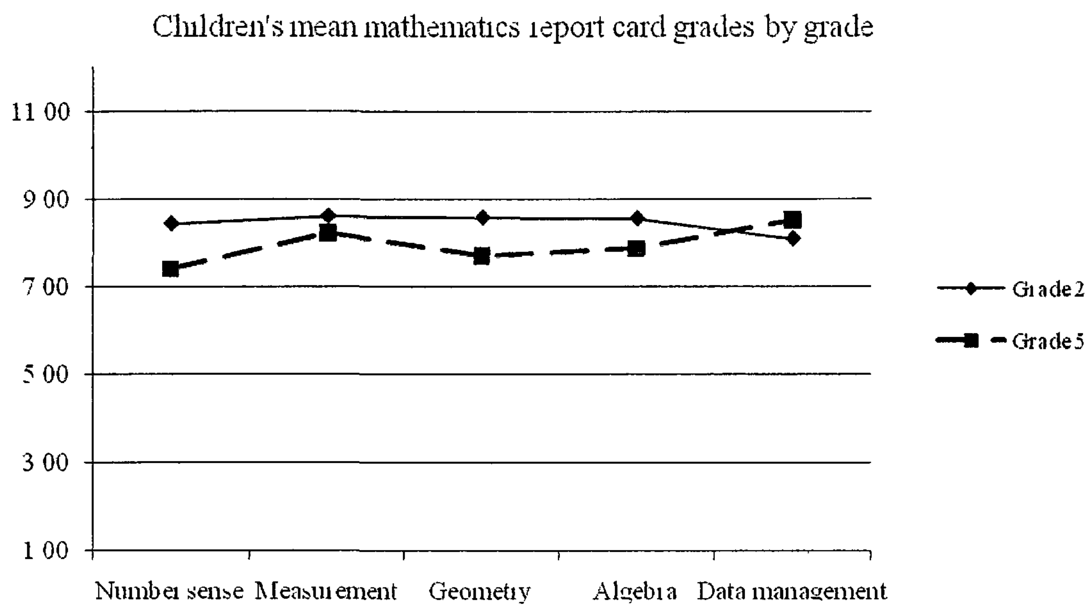
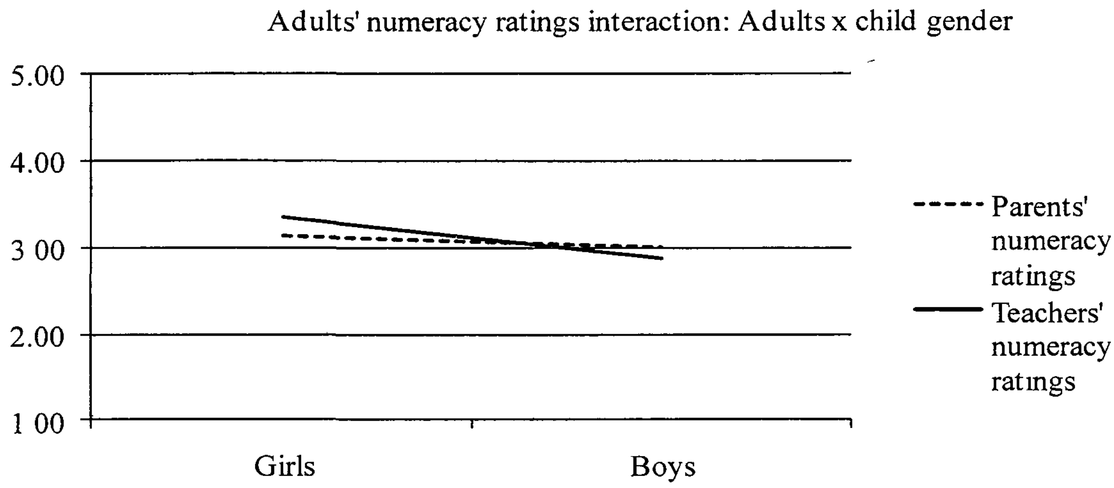


Figure 11

Adults' numeracy ratings interaction Adults x child gender



Appendix A

COAT-AM (occupations - short version)
Who should do these jobs?

Here is a list of jobs that people can do. We want you to tell us if you think each job should be done by men, by women, or by both men and women. There are no right or wrong answers. We just want to know who you think should do these jobs. If you think it should be done by only men, say 1; if you think it should be done by only women, say 2; if you think it should be done by both men and women, say 3.
Who should be a(n):

	Only men	Only women	Both men & women
1. Dishwasher in a restaurant	1	2	3
2. Supermarket check-out clerk	1	2	3
3. Artist	1	2	3
4. House cleaner	1	2	3
5. Telephone operator	1	2	3
6. School principal	1	2	3
7. Librarian	1	2	3
8. Cook in a restaurant	1	2	3
9. Baby-sitter	1	2	3
10. Secretary	1	2	3
11. Plumber	1	2	3
12. Nurse	1	2	3
13. Factory owner	1	2	3
14. Hair stylist	1	2	3
15. Scientist	1	2	3
16. Baker	1	2	3
17. Police officer	1	2	3
18. Computer builder	1	2	3
19. Architect	1	2	3
20. Dentist	1	2	3
21. Comedian	1	2	3
22. Dental assistant	1	2	3
23. Ship captain	1	2	3
24. Spy	1	2	3
25. Florist (arrange & sell flowers)	1	2	3

Appendix B

COAT-AM (activities - short version)

Who should do these activities?

Here is a list of activities that people can do. We want you to tell us if you think each activity should be done by boys, by girls, or by both boys and girls. There are no right or wrong answers. We just want to know who you think should do these activities. If you think it should be done by only boys, say 1; if you think it should be done by only girls, say 2; if you think it should be done by both boys and girls, say 3.

Who should:

	Only boys	Only girls	Both boys & girls
1. Fly a model plane	1	2	3
2. Iron clothes	1	2	3
3. Sew from a pattern	1	2	3
4. Vacuum a house	1	2	3
5. Go to a beach	1	2	3
6. Go horseback riding	1	2	3
7. Wash clothes	1	2	3
8. Build with tools	1	2	3
9. Play cards	1	2	3
10. Shoot pool	1	2	3
11. Set the table for dinner	1	2	3
12. Fix bicycles	1	2	3
13. Play darts	1	2	3
14. Do gymnastics	1	2	3
15. Play hide and seek	1	2	3
16. Baby-sit	1	2	3
17. Play video games	1	2	3
18. Draw (or design) buildings	1	2	3
19. Bake cookies	1	2	3
20. Sketch (or design) clothes	1	2	3
21. Grocery shop	1	2	3
22. Draw (or design) cars/rockets	1	2	3
23. Play basketball	1	2	3
24. Build model airplanes	1	2	3
25. Do crossword puzzles	1	2	3

Appendix C

COAT-AM (traits - short version)

Who should be this way?

Here is a list of words that describe people. Please circle the number that shows who think should be this way. There are no right or wrong answers. We want to know who you think should be this way. If you think only boys should be this way, say 1; if you think only girls should be this way, say 2; if you think both boys and girls should be this way, say 3; and if you think neither boys nor girls should be this way, say N.

Who should:	Only boys	Only girls	Both boys & girls	Neither boys nor girls
1. Be affectionate	1	2	3	N
2. Misbehave	1	2	3	N
3. Be confident (sure of themselves)	1	2	3	N
4. Be logical	1	2	3	N
5. Be gentle	1	2	3	N
6. Enjoy geography	1	2	3	N
7. Complain	1	2	3	N
8. Be dominant	1	2	3	N
9. Be charming	1	2	3	N
10. Brag a lot	1	2	3	N
11. Be loud	1	2	3	N
12. Be loving	1	2	3	N
13. Have good manners	1	2	3	N
14. Be neat	1	2	3	N
15. Be good at art	1	2	3	N
16. Enjoy art	1	2	3	N
17. Act as a leader	1	2	3	N
18. Try to look good	1	2	3	N
19. Be helpful	1	2	3	N
20. Be competitive	1	2	3	N
21. Be creative	1	2	3	N
22. Enjoy music	1	2	3	N
23. Study hard	1	2	3	N
24. Follow directions	1	2	3	N
25. Be smart	1	2	3	N

Appendix D

COAT-PM (occupations - short version)

What I want to be

Here is a list of jobs that people can do. Please circle the number that shows how much you would want to do each of these jobs.

How much would you want to be a(n):

	Not at all	Not much	Some	Very much
1. Supermarket check-out clerk	1	2	3	4
2. Artist	1	2	3	4
3. Perfume salesperson	1	2	3	4
4. Elevator operator	1	2	3	4
5. Jockey (ride a horse in a race)	1	2	3	4
6. Librarian	1	2	3	4
7. Cheerleader	1	2	3	4
8. Cook in a restaurant	1	2	3	4
9. Secretary	1	2	3	4
10. Nurse	1	2	3	4
11. Banker	1	2	3	4
12. Writer	1	2	3	4
13. Geographer	1	2	3	4
14. Lawyer	1	2	3	4
15. Hair stylist	1	2	3	4
16. Construction worker	1	2	3	4
17. Scientist	1	2	3	4
18. Baker	1	2	3	4
19. Computer builder	1	2	3	4
20. Architect	1	2	3	4
21. Dental assistant	1	2	3	4
22. Ship captain	1	2	3	4
23. Spy	1	2	3	4
24. Jewellery maker	1	2	3	4
25. Florist (arrange and seller flowers)	1	2	3	4

Appendix E

COAT-PM (activities – short version)
What I do in my free time

Here is a list of activities that people do. Please circle the number that shows how often you do each of these activities
How often do you:

	Never	Rarely	Sometimes	Often or Very often
1. Wash the dishes	1	2	3	4
2. Iron clothes	1	2	3	4
3. Build forts	1	2	3	4
4. Paint pictures	1	2	3	4
5. Vacuum a house	1	2	3	4
6. Go fishing	1	2	3	4
7. Wash clothes	1	2	3	4
8. Fix a car	1	2	3	4
9. Practise cheerleading	1	2	3	4
10. Build with tools	1	2	3	4
11. Cook dinner	1	2	3	4
12. Shoot pool	1	2	3	4
13. Jump rope	1	2	3	4
14. Play tag	1	2	3	4
15 Play darts	1	2	3	4
16. Do gymnastics	1	2	3	4
17. Play dodgeball	1	2	3	4
18. Ride a bicycle	1	2	3	4
19. Play hide and seek	1	2	3	4
20. Watch game/quiz shows	1	2	3	4
21. Baby-sit	1	2	3	4
22. Hunt	1	2	3	4
23. Shoot a bow and arrow	1	2	3	4
24. Bake cookies	1	2	3	4
25. Draw (or design) cars/rockets	1	2	3	4

Appendix F

COAT-PM (traits – short version)
What I am like

Here is a list of words and phrases that describe people. Please circle the number that shows how much each of the words or phrases describes you.

Is this like you?

	Not at all like me	Not much like me	Somewhat like me	Very much like me
1. Emotional (express feelings)	1	2	3	4
2. Aggressive	1	2	3	4
3. Excitable	1	2	3	4
4. Dependent	1	2	3	4
5. Ambitious	1	2	3	4
6. Affectionate	1	2	3	4
7. Adventurous	1	2	3	4
8. Enjoys geography	1	2	3	4
9. Good at geography	1	2	3	4
10. Confident (Sure of yourself)	1	2	3	4
11. Enjoys physical education (gym)	1	2	3	4
12. Logical	1	2	3	4
13. Good at math	1	2	3	4
14. Dominant	1	2	3	4
15. Charming	1	2	3	4
16. Good at foreign languages	1	2	3	4
17. Has good manners	1	2	3	4
18. Creative	1	2	3	4
19. Tries to look good	1	2	3	4
20. Appreciative (thankful)	1	2	3	4
21. Gentle	1	2	3	4
22. Good at social studies	1	2	3	4
23. Loving	1	2	3	4
24. Helpful	1	2	3	4
25. Good at music	1	2	3	4

Appendix G

Questionnaire for Students - Parents

The parent (guardian) who spends the most time with your child should answer the questions below. Your answers are completely anonymous and will be used for research purposes only. You can choose not to answer any question. It will take approximately 5-10 minutes to complete these questions.

Demographics

Please indicate:

a) Child's age: _______ Years and ________ Months

b) Child's gender: _____ Boy _____ Girl

c) Your relationship to the child: _____ Mother _____ Father

_____ Other, please specify: ______________

d) The ages of any brother(s) or sister(s) of your child:

Brothers ______________________ Sisters ______________________

e) The highest level of education reached by:

Yourself: _____________________ Your spouse (if applicable): ________________

f) The occupation of:

Yourself: _____________________ Your spouse (if applicable): ________________

Please circle your child's ability in

1.) Vocabulary

-------- 1 ----------------------- 2 ---------------------3 ------------------- 4 ------------------5 ---------

definitely not as good as children this age | as good as children this age | well above children this age

2.) Pronouncing unfamiliar words

-------- 1 ----------------------- 2 ---------------------3 ------------------- 4 ------------------5 ---------

definitely not as good as children this age | as good as children this age | well above children this age

3) Reading comprehension

-------- 1 ----------------------- 2 ---------------------3 ------------------- 4 ------------------5 ---------

definitely not as good as children this age | as good as children this age | well above children this age

4) Performing mathematical calculations (in addition, subtraction, multiplication and division)

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

5) Performing mathematical calculations with fractions and decimals

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

6.) Measuring dimensions (e.g., length, distance, area)

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

7) Understanding geometry (e g , rotation)

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

8.) Solving practical/applied problems

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

9.) Mathematical reasoning and analysis

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

Place your questionnaire and consent form in the enclosed envelope and seal it.
Return the envelope to your child's teacher. He or she will pass it on to us.
Thank you very much for participating in this research project

Appendix H

Questionnaire for Students - Teachers

Your answers are completely anonymous and will be used for research purposes only. You can choose not to answer any question. It will take approximately 10 minutes to complete these questions.

Child ID# ____________

Demographics

Please indicate:

a) Your gender: _____ Male _____ Female ______ Other

b) Your highest level of education obtained: ________________________

c) Additional professional trainings received:

____

d) Years as a teacher: _________________

e) If applicable, years as a teacher of this grade: _________________

Please circle your student's ability in

1.) Vocabulary

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------

definitely not as good as children this age | as good as children this age | well above children this age

2) Pronouncing unfamiliar words

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------

definitely not as good as children this age | as good as children this age | well above children this age

3) Reading comprehension

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------

definitely not as good as children this age | as good as children this age | well above children this age

4.) Performing mathematical calculations (in addition, subtraction, multiplication and division)

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------

definitely not as good as children this age | as good as children this age | well above children this age

5.) Performing mathematical calculations with fractions and decimals

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

6.) Measuring dimensions (e.g., length, distance, area)

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

7.) Understanding geometry (e.g., rotation)

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

8) Solving practical/applied problems

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

9.) Mathematical reasoning and analysis

-------- 1 ------------------------ 2 ----------------------3 -------------------- 4 -------------------5 ---------
definitely not as good as children this age | as good as children this age | well above children this age

Please place your questionnaires and consent form in the enclosed envelope and seal it.
Thank you very much for participating in this research project

References

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

Arnett, J. J. (2007). *Adolescence and emerging adulthood (3rd Ed)* New Jersey: Pearson Prentice Hall.

Association of Universities and Colleges of Canada. (2010). Advocacy and research: Quick facts. Retrieved March 10, 2011, from http://www.aucc.ca/policy/quick-facts_e.html.

Banse, R., Gawronski, B., Rebetez, C., Gutt, H., & Morton, B. J. (2010). The development of spontaneous gender stereotyping in childhood: Relations to stereotype knowledge and stereotype flexibility. *Developmental Science, 13,* 298-306.

Blakemore, J. E., Berenbaum, S. A., & Liben, L. S. (2009). *Gender Development.* New York: Psychology Press.

Bleeker, M. M., & Jacobs, J. E. (2004). Achievement in math and science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology, 96,* 97-109.

Below, J. L., Skinner, C. H., Fearrington, J. Y., & Sorrell, C. A. (2010) Gender differences in early literacy: Analysis of kindergarten through fifth grade dynamic indicators of basic early literacy skills probe. *School Psychology Review, 39,* 240-257.

Bouffard, T., Marcoux, M., Vezeau, C. & Bordeleau, L. (2003). Changes in self-perceptions of competence and intrinsic motivation among elementary school children. *British Journal of Educational Psychology, 73,* 171-186.

Bowlby, J. (1982). Attachment. (2nd Ed.). (p. 265-378). New York: Basic Books

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.

Burnett, P. C. (1996). Gender and grade differences in elementary school children's descriptive and evaluative self-statements and self-esteem. *School Psychology International, 17,* 159-170.

Dick, T. P., & Rallis, S. F. (1991). Factors and influences on high school students' career choices. *Journal of Research in Mathematics Education, 22,* 281-292.

Dickhauser, D., & Meyer, W. U. (2006). Gender differences in young children's math ability attributions. *Psychology Science, 28,* 3-16.

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

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis. *Psychological Bulletin, 136,* 103-127.

Fordham, S., & Ogbu, J. U. (1986). Black student's school success: Coping with the "Burden of 'acting White'" *The Urban Review, 18,* 176-206.

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

Freedman-Doan, C., Wigfield, A., Eccles, J. S., Blumenfeld, P., Arbreton, A., & Harold, R. D. (2000). What am I best at? Grade and gender differences in children's beliefs about ability improvement. *Journal of Applied Developmental Psychology, 21,* 379-402.

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal Applied Economics, 2,* 210-240.

Helwig, R., Anderson, L., & Tindal, G. (2001). Influence of elementary student gender on teachers' perceptions of mathematics achievement. *Journal of Educational Research, 95,* 93-102.

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.

Jacobs, J. E. (1991). Influence of gender stereotypes on parent and child mathematics attitude. *Journal of Educational Psychology, 83,* 518-527.

Karkkainen, R., Raty, H., & Kasanen, K. (2010). How are children's perceptions of the malleability of their academic competencies related to their teachers' and parents' views? *Social Psychology of Education, 13,* 557-573.

Kenney-Benson, G. A., Pomerantz, E. M., Ryan, A. M., & Patrick, H. (2006). Sex differences in math performance: The role of children's approach to homework. *Developmental Psychology, 42,* 11-26.

Kiefer, A. K., & Sekaquaptewa, D. (2007a). Implicit stereotypes, gender identification, and math-related outcomes. *Psychological Science, 18,* 13-18.

Kiefer, A. K., & Sekaquaptewa, D. (2007b). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology, 43,* 825-832.

Kurtz-Costes, B., Rowley, S. J., Harris-Britt, A., & Woods, T. A. (2008). Gender stereotypes about mathematics and science and self-perceptions of ability in late childhood and early adolescence. *Merrill-Palmer Quarterly, 54,* 386-409

Lee, J., & Schell, A. (under review). Gender differences in emergent mathematics: Perception vs. reality.

Lee, J. & Schell, A. (2010). *Do young children have gender-stereotyped knowledge of math competence?* Poster session presented at the 22nd Annual Association for Psychological Science Convention, Boston MA.

Leinbach, M. D. & Fagot, B. I. (1993). Categorical habituation to male and female faces: gender schematic processing in infancy. *Infant Behavior & Development, 16,* 317-332.

Liben, L. S. & Bigler, R. S. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development, 67,* i-183.

Lindsay, C., & Almey, M. (2006). Chapter 4: Education. In C. Lindsay & M. Almey (Eds). *Women in Canada (5th Edition). A gender based statistical report* (pp. 89 – 102). Ottawa, ON: Statistics Canada.

Lloyd, J. E. V., Walsh, J., & Yailagh, M. S. (2005). Sex differences in performance attributions, self-efficacy, and achievement in mathematics: If I'm so smart, why don't I know it? *Canadian Journal of Education, 28,* 384-408.

Marinak, B. A. & Gambell, L. B. (2010). Reading motivation: Exploring the elementary gender gap. *Literacy Research and Instruction, 49,* 129-141.

Mather, N., Woodcock, R. W. (2001). Examiner's Manual. Woodcock-Johnson III Tests of Achievement. Rolling Meadows, IL: Riverside Publishing

Martin, C. L., & Ruble, D. N. (2009). Patterns of gender development. *Annual Review of Psychology, 61,* 353-381.

Martinot, D., & Desert, M. (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,* 455-471.

Mıller, C. F., Lurye, L. E., Zosuls, K. M., & Ruble, D. N. (2009) Accessibility of gender stereotype domains: Developmental and gender differences in children. *Sex Roles, 60,* 870-881.

Mitra, A. (2002). Mathematics skill and male-female wages. *The Journal of Socio-Economics, 31,* 443-456.

Muzzatti, B., & Agnoli, F. (2007). Gender and mathematics: Attitudes and stereotype threat susceptibility in Italian children. *Developmental Psychology, 43,* 747-759.

Parsons, J. E., Adler, T. F., & Kaczala, C. (1982). Socialization of achievement attitudes and beliefs: Parental influences. *Child Development, 53,* 310-321.

Paulsen, K., & Johnson, M. (1983). Sex role attitudes and mathematical ability in 4th-, 8th-, and 11th-grade students from a high socioeconomic area. *Developmental Psychology, 19,* 210-214.

Raty, H., & Kasanen, K. (2007). Gendered views of ability in parents' perceptions of their children's academic competencies. *Sex Roles, 56,* 117-124.

Schell, A. & Lee, J. (2009) *Gender stereotypes in emergent mathematics Exploring children's perspectives* Poster session presented at the 2009 Biennial Society for the Study of Human Development meeting, Ann Arbor, MI.

Schmader. T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38,* 194-201.

Shapka, J. D., Domene, J. F., & Keating, D. P. (2008). Gender, mathematics achievement, and the educational and occupational aspirations of Canadian youth. In H. M. G Watt & J. S. Eccles (Eds). *Gender and occupational outcomes Longitudinal assessments of individual, social, and cultural influences* (p 27-54). Washington, DC. American Psychological Association.

Sherman, J. (1980). Mathematics, spatial visualization, and related factors: Changes in girls and boys, grades 8-11. *Journal of Educational Psychology, 72,* 476-482.

Signorella, M. L. & Liben, L. S. (1985). Assessing children's gender stereotype attitudes. *Psychological Documents, 15,* 7. Retrieved March 29, 2011, from http://www.personal.psu.edu/sdq/articles/gasc.html

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2005). Parents' socializing behaviour and children's participation in out-of-school activities. *Applied Developmental Science, 9,* 14-30.

Six, B., & Eckes, T. (1991). A closer look at the complex structure of gender stereotypes. *Sex Roles, 24,* 57-71.

Skaalvik, E. (1990). Gender differences in general academic self-esteem and in success expectations on defined academic problems. *Journal of Educational Psychology, 82,* 593-598.

Statistics Canada (2010). Table 2: University qualifications awarded by field of study and gender. Retrieved March 15, 2011, from http://www.statcan.gc.ca/daily-quotidien/100714/t100714b2-eng.htm

Steele, J. (2003). Children's gender stereotypes about math: The role of gender stratification. *Journal of Applied Social Psychology, 33,* 2587-2606.

Stevenson, H. W., & Newman, R. S. (1986). Long-term prediction of achievement and attitudes in mathematics and reading, *Child Development, 57,* 646-659.

Stipek, D. J., & Gralinski, J. H. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology, 83,* 361-371.

Strambler, M. J., & Weinstein, R. S. (2010). Psychological disengagement in elementary school among ethnic minority students. *Journal of Applied Developmental Psychology, 31,* 155-165.

Tenenbaum, H. R. (2009). "You'd be good at that": Gender patterns in parent-child talk about courses. *Social Development, 18,* 447-463.

Tenenbaum, H. R., & Leaper, C. (2003). Parent-child conversations about science: The socialization of gender inequalities? *Developmental Psychology, 39,* 34-47.

Thompson, S. K. (1975). Gender labels and early sex role development. *Child Development, 46,* 339-347.

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology, 92,* 144-151.

Trautner, H. M., Ruble, D. N., Cyphers, L., Kirsten, B., Behrendt, R., & Hartmann, P. (2005). Rigidity and flexibility of gender stereotypes in childhood: Developmental or differential? *Infant and Child Development, 14,* 365-381.

Wood, E., Groves, A., Bruce, S., Willoughby, T., & Desmarais, S. (2003). Can gender stereotypes facilitate memory when elaborative strategies are used? *Educational Psychology, 23,* 169-180.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson III Tests of Achievement. Rolling Meadows, IL: Riverside Publishing.