10-24-2023

Amultidimensional examination of children's endorsement of gender stereotypes

Cindy Faith Miller

Lorey A. Wheeler

Bobbi Woods

Received: 6 March 2023 | Revised: 23 October 2023 | Accepted: 24 October 2023

DOI: 10.1111/sode.12725

ORIGINAL ARTICLE

WILEY

A multidimensional examination of children's endorsement of gender stereotypes

Cindy Faith Miller[1] | Lorey A. Wheeler[2] | Bobbi Woods[3]

[1]Department of Counseling, Leadership, Adult Education, & School Psychology, College of Education, Texas State University, San Marcos, Texas, USA

[2]The Nebraska Center for Research on Children, Youth, Families and Schools, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

[3]T. Denny Sanford School of Social and Family Dynamics, Arizona State University, Tempe, Arizona, USA

Correspondence
Cindy Faith Miller, Department of Counseling, Leadership, Adult Education, & School Psychology, College of Education, Texas State University, 601 University Drive, San Marcos, TX 78666-4684, USA.
Email: cfmiller@txstate.edu

Funding information
National Science Foundation, Grant/Award Number: DRL-1561424

Abstract

The present research applied a multidimensional framework to the study of gender stereotypes by investigating whether elementary school children display different levels of endorsement when considering distinct gender stereotype constructs (ability, category, and interest) and feminine versus masculine stereotypes. Study 1 ($N = 403$) compared children's ability and category beliefs using a set of gender-neutral skill items. Study 2 ($N = 539$) extended this research by examining whether children showed different patterns of ability and category decisions for feminine versus masculine occupational items. Study 3 ($N = 974$) furthered our understanding of the construct dimension by comparing children's interest and ability decisions within the STEM domain. Findings revealed that older elementary school children endorsed ability stereotypes more strongly than category stereotypes and, across all age groups, children endorsed interest stereotypes more strongly than ability stereotypes. Findings also revealed age differences in how children think about masculine versus feminine stereotypes. For masculine stereotypes, younger children showed stronger endorsement than older children; however, for feminine stereotypes, the reverse pattern was found such that older children showed more stereotyped thinking than younger children.

The present study illustrates the benefits of employing a multidimensional framework to gain a more nuanced understanding of how children apply their increasing knowledge of gender stereotypes.

KEYWORDS
elementary school, feminine and masculine stereotypes, gender stereotypes, multidimensional framework, stereotype constructs

1 | INTRODUCTION

Children's knowledge of gender stereotypes begins prior to age 2 (Martin et al., 2002; Ruble et al., 2006) and, by the time they are 3–4 years old, most children can sort toys and activities reliably into girl and boy categories (e.g., Pillow et al., 2022; Ruble et al., 2006). Children's rapid growth of gender knowledge is due not only to the highly functional use of gender in our society (Bem, 1981; Bigler & Liben, 2007; Hilliard & Liben, 2010), but also to children's active, self-initiated role in their own gender development (e.g., Martin & Ruble, 2004). As children initially begin to accrue their knowledge of gender stereotypes, they often apply this information quite rigidly (e.g., Dolls are *only* for girls—not boys!) before demonstrating a relative phase of flexibility after age six (Martin & Ruble, 2004; Miller et al., 2006; Ruble et al., 2006). Understanding developmental differences in children's knowledge and endorsement are important as gender stereotypes are related to children's behaviors and views of themselves (e.g., Liben & Bigler, 2002; Miller et al., 2006; Patterson, 2012; Wolff, 2021). Importantly, there is ample evidence that endorsement of gender stereotypes is associated with children's self-perceptions and interest in gender-typed occupations and STEM (Master, 2021; Master et al., 2021; Wolff, 2021), which, in turn, relates to future career aspirations (Eccles, 2011; Garriott et al., 2017; Steffens et al., 2010). Given that gender disparities in male-dominated fields endanger innovation and contribute to the gender-wage gap (Levanon et al., 2009; Shapiro & Sax, 2011), it is critical to investigate and address children's gender stereotype cognitions within academic and career domains.

Despite the extensive literature documenting children's early conceptions of gender stereotypes and the role these beliefs have in children's academic and career aspirations, little research has examined whether age differences depend on the multidimensional nature of stereotypes. Applying a multidimensional framework raises interesting questions that include examining whether children display different levels of endorsement when thinking about distinct gender stereotype constructs. Within the gender stereotype experimental literature, there have been three main stereotype constructs that have been studied (Master et al., 2021; Miller et al., 2006; Weisgram, 2016): ability (e.g., Boys are *better* at math than girls), interest (e.g., Boys *like* math more than girls), and category (e.g., This is a game *for* boys); however, little research has compared whether children show different levels of endorsement for these stereotype constructs. At the same time, it is unclear whether children's endorsement of gender stereotypes depends on the gender role of the item (e.g., feminine or masculine stereotypes). Young children demonstrate increased flexibility with age (Bigler & Liben, 1993; Martin & Ruble, 2004; Miller et al., 2006; Signorella et al., 1993), yet it is unknown if this cognitive shift manifests across constructs and gender role groups. It is critical to compare the components (e.g., ability versus category) within stereotype dimensions (e.g., construct) as different components might show unique impacts on behavior.

By investigating gender stereotype development through a multidimensional lens, the present research takes an initial step toward addressing key gaps in the literature. Across three studies, we: a) compared children's endorsement of stereotype constructs (ability, category, and interest) in domains relevant to career aspirations (i.e., skills, occupations, and STEM); b) compared children's endorsement of masculine and feminine occupational items; and c) examined gender and age differences. We define ability and interest gender stereotypes as beliefs that one gender group has more competence (ability) or interest/enjoyment (interest) in a particular area than another gender group. We define category gender stereotypes as the belief that an object, activity, social role, or skill is associated with or more appropriate for a particular gender group. Category beliefs represent the basic cognition that something is linked with or placed in a specific gender category (e.g., math = boys; Cvenek et al., 2011).

1.1 | Guiding theoretical approaches

The present research was guided by theories focused on the role of cognitions in children's choices and behaviors. According to cognitive theories of gender development, children function as "gender detectives" who are motivated intrinsically to seek out information about what it means to be a boy and girl in this world (Martin & Ruble, 2004; Miller et al., 2013). Children's search for gender cues leads them to amass a wide knowledge base of gender stereotypes that they use to guide their behaviors and views of themselves and others (e.g., Martin & Ruble, 2004). Given developmental changes in children's cognitive abilities, children initially apply stereotype knowledge in a rigid manner (e.g., Only girls play with dolls!) before they begin to think more flexibly after age six (Martin & Ruble, 2004; Trautner et al., 2005). Cognitive theories recognize that there are moderating factors that influence whether children match their behavior to their gender stereotype knowledge; however, little research has explored these moderating factors (e.g., Martin et al., 2002; Miller et al., 2006).

Situated expectancy-value theory (SEVT) is another model that guides the present research (Eccles & Wigfield, 2020). SEVT is a motivational theory that outlines psychological constructs and cultural and social influences that influence achievement-related choices and performance. SEVT distinguishes between different types of psychological constructs that uniquely and jointly impact behavior such as ability/expectancy beliefs (e.g., I am good at math) and intrinsic value/interest (e.g., I like math; Guo et al., 2015; Meece et al., 1990). Guided by SEVT, the present research introduces the study of individual psychological constructs to the examination of gender stereotypes. We examined stereotypes in the context of ability and interest beliefs; however, we also were motivated to investigate the category construct as these beliefs commonly are applied in experimental studies that examine the effect of gender stereotypes on children's behavior (i.e., gender labeling studies; Miller et al., 2006).

1.2 | Multidimensional framework of gender stereotypes

Research examining gender stereotype development has spanned across different constructs, domains/content areas, and gender role groups. For instance, some studies focused on understanding children's gender knowledge or attitudes associated with ability stereotypes (e.g., Who do you think should be good at math?; McGuire et al., 2020), whereas other studies examined interest stereotypes (e.g., Who likes to do math more?; Cvencek et al., 2011) or have applied category stereotypes to novel toys (e.g., This is a toy for boys; Weisgram, 2016). Research methodology also can vary according to domain by asking about toys (e.g., Who likes trucks?; King et al., 2020), occupations (e.g., Who should be a nurse?; Patterson, 2012), academic subjects (e.g., Who do you think should be good at engineering; McGuire et al., 2020), and personality characteristics (e.g., Who is gentle?; Powlishta, 2000). Another stereotype dimension includes gender role because measures often are divided into feminine and masculine items (e.g., Liben & Bigler, 2002). Table 1 provides an illustration of the multidimensional nature of stereotypes that includes the dimensions of construct, domain, and gender role. Although there are more components worthy of study, this table provides an initial example of how a multidimensional framework can be applied to the study of gender stereotypes.

Given the multidimensional nature of gender stereotypes, it is essential that research explores age differences within and across dimensions. Although much attention has been devoted to tracking children's developing

TABLE 1 A multidimensional illustration of gender stereotypes.

Constructs	Stereotype domain/Content: Activities/Toys – Feminine	Activities/Toys – Masculine	Occupations – Feminine	Occupations – Masculine	Academic subjects – Feminine	Academic subjects – Masculine
Ability	*Girls are good at hopscotch*	*Boys are good at football*	*Girls would make a good nurse*	*Boys would make a good car mechanic*	*Girls are good at reading*	*Boys are good at math*
Category	*Hopscotch is for girls*	*Football is for boys*	*Nursing is for girls*	*Car mechanics is for boys*	*Reading is for girls*	*Math is for boys*
Interest	*Girls like hopscotch*	*Boys like football*	*Girls are interested in nursing*	*Boys are interested in car mechanics*	*Girls like reading*	*Boys like math*

knowledge within the domain dimension (Miller et al., 2009; Ruble et al., 2006), little research has compared components within the construct and gender role dimensions. The important role of stereotype construct was raised initially by Miller et al. (2006) when discussing the conflicting findings in the research literature concerning the effect of gender stereotypes on children's behavior (i.e., gender labeling studies). For instance, some research found that children's behavior (e.g., performance) was affected by applying gender stereotype labels to novel toys/activities, but other research showed null findings. In their review of these experimental studies, the authors discerned that the reason for the conflicting findings was because studies differed in the type of stereotype label that they manipulated (i.e., category versus ability). When the researchers organized the studies by age of participant and type of label, they concluded that children seven and younger were more influenced by category stereotypes and that children older than seven were more influenced by ability stereotypes (Miller et al., 2006). Yet, we are not aware of any research that has tested this prediction or even directly probed children's cognitions for category stereotypes. Although experimental studies have tested the effect of applying category stereotypes on children's behavior (e.g., this is a game for boys; Weisgram, 2016), research has not examined children's cognitions associated with gender category stereotypes (e.g., who is math for…girls, boys, or both girls and boys?). This is unfortunate because understanding children's cognitions about category stereotypes may have specific implications for promoting a sense of belonging in stereotyped contexts (e.g., math is for me; Cvencek et al., 2011). There has been one paper, however, that included multiple studies comparing children's gender cognitions for interest and ability stereotypes in engineering and computer science that found interest stereotypes were endorsed more strongly than ability stereotypes (Master et al., 2021). The methodology used in this research involved asking one item each about boys and girls (e.g., How good are most boys/girls at computer science?) and creating a difference score to represent endorsement of the gender stereotype in that domain (e.g., computer science; positive scores reflecting boys are better than girls). The present research builds on this research by comparing directly, in three studies, whether children's endorsement of gender stereotypes varies according to construct.

Like stereotype construct, there has been little attention paid to the role of the gender role dimension. Although the commonly used gender stereotype measures include both masculine and feminine items (e.g., Liben & Bigler, 2002), scores tend to be calculated across these items so that the responses index children's overall endorsement with gender stereotypes (Liben & Bigler, 2002; Patterson, 2012). Thus, conclusions drawn from gender stereotype studies tend to be broad (e.g., children's gender stereotyped attitudes become more flexible with age; Leaper, 2015) and have not considered whether children's gender cognitions vary according to the gender role characterization of the stereotype. Given that children tend to think differently about gender norm violations for girls and boys (e.g., boys evaluated more harshly than girls for violating gender appearance norms; e.g., Blakemore, 2003; Masters et al., 2021; Mulvey & Irvin, 2018), and there is much attention devoted to encouraging girls to enter male-dominated fields (e.g., STEM), but minimal efforts focused on boys entering female-dominated fields (e.g., early childhood educators), it is essential that research examine children's developing attitudes for masculine and feminine characteristics separately. Of the few studies that have compared masculine and feminine items directly, the findings are mixed and vary by age. In particular, research examined this issue with 9- to 16-year-olds and found that there were no differences in children's endorsement of masculine and feminine occupational items (Wood et al., 2021). Yet, they found that rejection of masculine items (i.e., higher stereotype knowledge compared to endorsement) was higher than rejection of feminine items in Study 1, but the opposite pattern was found in Study 2 (Wood et al., 2021). Another recent study examining children ages 6-to-11 found that older children showed more flexible responding than younger children for masculine, but not feminine occupations (Canessa-Pollard et al., 2022). The present research builds on this knowledge by evaluating whether children's stereotype endorsement varies for masculine and feminine occupational items.

1.3 | Gender and age differences in stereotype endorsement

The gender development literature includes a large body of research examining gender and age differences in children's gender stereotype endorsement. These studies generally find that children's gender attitudes become more flexible with age (e.g., McGuire et al., 2020; Miller at al., 2006; Signorella et al., 1993; Trautner et al., 2005). This conclusion is based on studies that found, with age, children are more likely to choose that "both girls and boys" can engage in stereotyped activities/occupations. In the present research, we were interested in examining whether this developmental pattern holds across different stereotype constructs and gender role groups. Thus, age (measured as grade levels) was included as an independent variable in the present studies.

Research examining gender differences in stereotype endorsement generally finds that girls are more flexible than boys (Signorella et al., 1993; Wood et al., 2021). However, some studies have not found gender differences in stereotype attitudes (e.g., Patterson, 2012). It is possible that conflicting findings are related to the age of participants, the constructs measured, and the domain and gender role of the items. The present research includes gender as an independent variable to test whether boys show higher stereotype endorsement than girls across the different stereotype dimensions.

1.4 | The present study

This research fills key gaps in the literature by applying a multidimensional framework to the study of gender stereotypes. Across three studies, we tested whether elementary school children's endorsement of gender stereotypes varied by *stereotype construct* (ability, category, and interest). We also examined the dimension of *gender role* by analyzing feminine and masculine items separately. We included children from kindergarten through fifth grade to allow us to test developmental patterns found in previous research (e.g., increased flexibility after age six). In Study 1, we compared children's category and ability beliefs using a set of gender-neutral skill items. For Study 2, we extended this research by studying children's occupational stereotypes. Using a set of masculine- and feminine-typed occupational items, we compared children's category and ability decisions and examined whether children showed different patterns of endorsement for feminine versus masculine items. In Study 3, we were interested in furthering our understanding of stereotype constructs by comparing children's interest and ability beliefs within the STEM domain. Taken together, these studies allowed us to compare stereotype constructs across three different content domains that are relevant to career aspirations. It is important to note that the present study used a binary conceptualization of gender. Although gender development researchers and advocates are calling for more inclusive frameworks for gender identity (e.g., Hyde et al., 2019), young children's gender cognitions still are influenced primarily by the gender binary and, therefore, it is valuable to understand their stereotyped thinking within this framework.

2 | STUDY 1

Our first aim was to determine whether children show different levels of endorsement for category versus ability stereotypes. We also were interested in understanding whether gender or age moderate any construct differences. We initially examined these novel research questions with a set of gender-neutral skill items to see whether children would make gender stereotyped decisions without the knowledge of cultural stereotypes. Given the lack of previous research on comparing ability and category stereotypes, our research questions specifically involving construct differences were exploratory. A second aim was to examine overall gender and age differences in stereotype endorsement. Based on previous research, we expected boys to endorse more stereotypes than girls, and for younger children to endorse more stereotypes than older children. Further, we expected children to show a gender in-group bias by selecting responses that favored their gender (e.g., McGuire et al, 2020).

2.1 | Method

2.1.1 | Participants

Participants included 403 (52% female) elementary school students (42.2% White; 30.8% Latinx; 5.2% Black/African American; 4.2% Asian; 0.2% Native American; 8.4% multiracial/multiethnic; 8.9% non-respondents) who were recruited from two elementary schools (one charter; one public) in the Southwestern U.S. The schools reported rates of free and reduced lunch as 85% and 24%. Children were divided into three grade groups (as a proxy for age): early (kindergarten and 1st grade: $n = 140$); middle (2nd and 3rd grade; $n = 116$); and upper (4th and 5th grade; $n = 147$).

2.1.2 | Procedure

In Spring 2017, parental consent forms were sent home with all children; 40% of parents provided consent for participation. Research staff administered surveys to kindergarten and 1st graders individually and to 2nd–5th graders in groups. Before responding to survey questions, all children answered practice questions to ensure that they understood how to use the response options. For the individual interviews, survey questions and response choices were read aloud, and children selected their answer from a set of pictorial stimuli with response options provided to them. For the group surveys, survey questions and response choices were read aloud, and children selected their answers on a paper survey, while research assistants monitored survey administration and provided individual assistance to children, as necessary. All procedures were approved by the University and School District IRBs.

2.1.3 | Measures

Children's gender stereotyped responses were measured using children's reports on two, 8-item scales; one scale measured children's gender-based ability beliefs and the other scale measured children's gender-based category beliefs. The items used in both measures were comprised of eight identical, gender-neutral skills. These items were developed by a panel of experts (i.e., parents, teachers, and engineers) as part of a larger study designed to measure engineering-related activities and skills (ERAS; Wheeler et al., 2022). The ERAS items were selected for the present research

TABLE 2 Study 1: Intercorrelations by study variables disaggregated by gender and grade.

	K/1st				2nd/3rd				4th/5th			
Variable	1	2	3	4	1	2	3	4	1	2	3	4
1. G_ability	_	.27*	.58**	.48**	_	.47**	.58**	.31*	_	.44**	.54**	.34**
2. B_ability	.09	_	.33**	.48**	.26	_	.49**	.72**	.46**	_	.18	.82*
3. G_category	.44**	.12	_	.45**	.47**	.49**	_	.38**	.50**	.45**	_	.28*
4. B_category	.15	.57**	.12	_	.21	.70**	.31*	_	.49**	.71**	.63**	_

Note: G_ability and G_category refer to the proportion of "only girls" responses; B_category and B_ability refers to the proportion of "only boys" responses. The results for the female sample are shown above the diagonal. The results for the male sample are shown below the diagonal.
*$p < .05$. **$p < .01$.

because they have implications for academic and career success and were believed to have minimal gender stereotypes associated with them. Sample items for the measures include: "Learning from their mistakes and failures" and "Trying out their ideas".

Ability stereotypes were assessed by asking children to report on who they think is good at the eight gender-neutral skills (e.g., Who do you think is good at learning from their mistakes and failures?). To match the format of gender labeling studies (e.g., Miller et al., 2006; Weisgram, 2016), *category stereotypes* were assessed by asking children to report on who they think the eight gender-neutral skills are for (e.g., Who do you think learning from their mistakes and failures is for?). Responses for both measures were rated as 0 (*only boys*), 1 (*both boys and girls*), or 2 (*only girls*). Based on prior research (e.g., Liben & Bigler, 2002; Trautner et al., 2005), children's ability and category stereotype scores were measured as the proportion of "*only boys/girls*" responses they provided to each scale. Because the items were gender-neutral, two scores were created for each measure: the proportion of "only girls" (female-bias) and the proportion of "only boys" (male-bias) responses, which allowed us to examine the possibility of ingroup bias. To account for possible order effects, the ability and category measures were counterbalanced.

2.2 | Results

2.2.1 | Preliminary analyses

To assess whether findings depended on counterbalancing order, a mixed-design analysis of variance (ANOVA) with gender and grade was conducted. Results revealed that the order X stereotype (ability or category; $p = .418$), order X gender ($p = .905$), order X stereotype X grade ($p = .518$), and order X stereotype X gender ($p = .088$) interactions were all not significant. The order X grade interaction, however, was significant, F (2, 385) = 3.55, $p = .03$, partial $\eta^2 = .02$. Tests of simple effects of grade within each level of measure order revealed that grade differences in stereotyped responses were similar across both versions. In contrast, examining order differences within each grade level indicated differences for early elementary students ($p = .038$), but not middle ($p = .146$) and upper ($p = .426$) elementary students. Namely, kindergarten and first grade students who answered the category measure first ($M = .44$, $SE = .03$) provided more stereotyped responses, averaged across both measures, than students who answered the ability measure first ($M = .37$, $SE = .02$). To control for the possible influence of measure order on findings, order was included as a covariate in the main analyses.

A key premise of our research is that children draw distinctions between stereotype constructs. Initially to examine construct distinctions, we conducted correlations between category and ability proportions. As presented in Table 2, the same-bias/cross-construct correlations (e.g., "only girls" ability proportions with "only girls" category proportions) ranged from .44–.82, which shows a moderate level of discriminant validity (Rönkkö & Cho, 2022).

TABLE 3 Study 1: Stereotype proportions by gender and grade.

Variables	Means (*SD*)		Means (*SD*)		
	Boys	Girls	K-1st grade	2nd–3rd grade	4th–5th grade
Gender ability stereotype					
Only girls	0.11 (.14)	0.15 (.19)	0.22 (.19)	0.07 (.12)	0.09 (.14)
Only boys	0.14 (.20)	0.08 (.12)	0.18 (.20)	0.09 (.14)	0.06 (.11)
Gender category stereotype					
Only girls	0.09 (.15)	0.12 (.18)	0.21 (.19)	0.06 (.12)	0.06 (.12)
Only boys	0.13 (.18)	0.06 (.10)	0.18 (.18)	0.06 (.13)	0.03 (.08)

2.2.2 | Overview of primary analyses

The main analyses involved a 2 (gender: girl or boy) × 3 (grade: K/1, 2nd/3rd, or 4th/5th) × 2 (stereotype: ability or category) × 2 (gender bias: female-bias or male-bias) × 2 (order: category first or ability first) mixed-design ANCOVA with gender and grade as between-subjects factors, stereotype and gender bias as within-subjects factors, and order as a covariate. In these analyses, between-subjects effects were relevant for examining overall social group (gender and grade) differences on proportion of stereotyped responses. The within-subject effects for stereotype tested whether children responded differently to the ability and category measures and whether any differences depended on grade and gender. The within-subjects effects for gender bias examined differences in the proportion of "only girls" and "only boys" responses and if any differences depended on grade and gender. When describing the findings below, the between-subjects and stereotype effects are presented first, followed by a description of the target's gender analyses. Follow-up tests were conducted using the Ŝidàk equation for multiple comparisons (Midway et al., 2020). Table 3 includes raw means and standard deviations for stereotype proportions. When describing the results below, estimated marginal means and standard errors are presented.

2.2.3 | Ability versus category

Overall, children provided stereotyped responses to .22 (*SE* = .01) of the items, which indicates that children responded "*only girls*" *or* "*only boys*" to approximately two out of the eight neutral items. The mixed-design ANCOVA results indicated that there was a *between-subjects* main effect for grade, F (2, 390) = 70.16, $p < .001$, partial $\eta^2 = .27$. Follow-up tests revealed that early elementary students (M = .20, SE = .01) provided a higher proportion of stereotyped responses overall when compared to middle (M = .07, SE = .01; $p < .001$) and upper (M = .06, SE = .01; $p < .001$) students; there were no differences between middle and upper students (p = .359). There was no main effect for gender (p = .084). For *within-subjects effects*, there was a main effect of stereotype, F (1, 390) = 4.56, p = .033, partial η^2 = .012, which indicated that, overall, participants provided more stereotyped responses when making ability (M = .12, SE = .01) than category (M = .10, SE = .01) decisions. There were not stereotype X grade (p = .192), stereotype X gender (p = .479) or 3-way (p = .525) interactions.

2.2.4 | Female-bias versus male-bias

Although there was not a main effect for gender bias (p = .377), the gender bias × grade (p = .019), gender bias × participant gender ($p < .001$) and 3-way (p = .001) interactions were all significant. Tests of simple effects revealed a gender ingroup bias; girls provided more "only girls" responses than boys (girls: $M = .13$, SE = .01; boys: $M = .10$, SE = .01; $p = .017$), whereas boys provided more "only boys" responses than girls (girls: $M = .07$, SE = .01; boys: $M = .14$, SE = .01; $p < .001$). However, tests of simple effects examining the 3-way interaction revealed that the gender ingroup bias depended on gender and grade. For girls, the ingroup bias was evident in the K/1st and 4th/5th grade groups; for boys, the ingroup bias emerged in the K/1st and 2nd/3rd grade groups.

2.3 | Discussion

The primary goal of Study 1 was to examine whether children think differently about ability and category stereotypes and to see whether any differences depended on gender and grade. We found that, overall, children endorsed more ability than category stereotypes. Correlations also showed a moderate to large positive relation between ability and category beliefs, which suggests that these two constructs show a moderate level of distinction. Although the overall level of stereotyping was low for the neutral items, differences between ability and category stereotypes still emerged with a small effect size. Although Miller et al. (2006) speculated that older children might be more influenced by ability stereotypes whereas younger children might be more influenced by category stereotypes, we did not find significant interactions with gender and grade. Yet, an examination of means (see Table 3) indicated that the significant findings may have been driven by the older age groups as the middle and upper elementary students had larger mean differences than the younger students who, overall, showed negligible differences in their ability and category decisions. It is possible that older children are more sensitive to differences between ability and category stereotypes due to their increasing sensitivity to ability evaluations (Dweck, 2002).

A secondary goal was to replicate age and gender patterns in the literature. As expected, children's stereotype decisions were more flexible with age. Namely, middle and upper elementary students were more likely to select "both girls and boys" when compared to early elementary students. In fact, even though children were presented with a set of neutral items, early elementary students (kindergarten and first graders), on average, made gender stereotyped decisions (either "only girls" or "only boys") forty percent of the time. This finding provides further evidence of young children's rigid cognitions about gender and active role in constructing their gender knowledge (Miller et al., 2013). In contrast to predictions, analyses did not reveal a gender difference in stereotype endorsement, and findings for ingroup bias varied by gender and grade. Given the low levels of stereotype endorsement for the neutral items, there may not have been enough variability for consistent gender differences to emerge.

3 | STUDY 2

The first aim was to replicate results from Study 1 and to generalize the findings to the occupational domain with a set of feminine and masculine stereotyped items. Based on the results from Study 1, we expected children to show higher levels of endorsement when making ability versus category decisions, and this pattern would be consistent across gender and age. The second aim was to determine whether children show age differences when responding to feminine and masculine occupational items. Although research has repeatedly found that children's attitudes generally become more flexible with age, we were interested in understanding whether this difference holds for both feminine and masculine occupations. The limited and conflicting findings in the literature did not allow us to make a strong hypothesis for this research question. Like Study 1, we also examined gender and age differences in endorsement, and predicted boys and younger children to endorse more stereotypes than girls and older children. Although we did not find the expected gender difference in Study 1, we believe that this likely was due to the gender-neutral nature of the items and the low overall levels of endorsement.

3.1 | Method

3.1.1 | Participants

The sample included 539 (53% female) elementary school students (37.5% White; 18.1% Latinx; 12.3% Black/African American; 4.0% Asian; 2.9% Native American; 15.2% multiracial/multiethnic; 9.9% non-respondents) who were recruited from two public elementary schools in the Southwestern U.S. The schools reported rates of free and reduced lunch as 27% and 48%. Children were divided into three grade groups: early (kindergarten and 1st grade: $n = 179$); middle (2nd and 3rd grade; $n = 168$); and upper (4th and 5th grade; $n = 192$).

3.1.2 | Procedure

In Fall 2017, parental consent forms were sent home with all children; 50% of parents provided consent for participation. Research staff administered surveys to kindergarten and 1st graders individually and to 2nd–5th graders in groups. Before responding to survey questions, all children answered practice questions to ensure that they understood how to use the response options. For the individual interviews, questions and response choices were read aloud and children selected their answer from a set of stimuli with response options provided to them. For the group surveys, survey questions and response choices were read aloud, and children selected their answers on an online survey. Research assistants monitored survey administration and provided individual assistance to children, as necessary. All procedures were approved by the University and School District IRBs.

3.1.3 | Measures

Children's occupational gender stereotypes were measured using children's reports on two, 14-item scales, which were counterbalanced. The items used were selected from the list of gender-typed occupations on the COAT scales (Liben & Bigler, 2002). The scales included nine masculine (e.g., soldier, car mechanic, police officer) and five feminine (e.g., nurse, babysitter, hair stylist) occupations.

Ability stereotypes were assessed by asking children to report on who they think is good at the occupations (e.g., Who do you think would be a good babysitter?). *Category stereotypes* were assessed by asking children to report on who they think the occupations are for (e.g., Who do you think being a police officer is for?). Responses were rated as 0 (*only boys*), 1 (*both boys and girls*), or 2 (*only girls*). To examine endorsement of cultural gender stereotypes (Liben & Bigler, 2002), the dependent measures were the proportion of "*only boys*" to the masculine occupations and *"only girls*" to the feminine occupations.

3.2 | Results

3.2.1 | Preliminary analyses

Preliminary analyses, using a mixed-design ANOVA (with gender and grade), were conducted to assess whether counterbalancing order affected rigidity proportions. There was not a main effect for order ($p = .728$) or order X grade ($p = .364$) and order X gender ($p = .106$) interactions. Yet, results suggested stereotype X order, F (1, 527) = 23.06, $p < .001$, partial $\eta^2 = .04$ and stereotype X order X grade, F (2, 527) = 4.42, $p = .013$, partial $\eta^2 = .02$, interactions. Tests of simple effects revealed that there were order differences for the category stereotypes ($p = .039$), but not for the

TABLE 4 Study 2: Intercorrelations by study variables disaggregated by gender and grade.

Variable	K/1st				2nd/3rd				4th/5th			
	1	2	3	4	1	2	3	4	1	2	3	4
1. Fem ability	_	.48**	.60**	.35**	_	.43**	.64**	.24*	_	.64**	.78**	.52**
2. Masc ability	.23*	_	.28**	.53**	.28*	_	.24*	.59*	.76**	_	.58**	.78**
3. Fem category	.73**	.32**	_	.47**	.73**	.24*	_	.35**	.68**	.61**	_	.63**
4. Masc category	.33**	.71**	.43**	_	.19	.72**	.34**	_	.49**	.76**	.67**	_

Note: The results for the female sample are shown above the diagonal. The results for the male sample are shown below the diagonal.
*$p < .05$. **$p < .01$.

TABLE 5 Study 2: Stereotype proportions by gender and grade.

Variables	Means (SD)		Means (SD)		
	Boys	Girls	K-1st grade	2nd–3rd grade	4^{th}–5th grade
Gender ability stereotype					
Feminine occupation	0.60 (.30)	0.61 (.29)	0.51 (.26)	0.62 (.28)	0.67 (.32)
Masculine occupation	0.50 (.28)	0.35 (.21)	0.45 (.25)	0.40 (.26)	0.40 (.26)
Gender category stereotype					
Feminine occupation	0.56 (.31)	0.57 (.30)	0.51 (.28)	0.59 (.30)	0.60 (.32)
Masculine occupation	0.42 (.28)	0.32 (.21)	0.42 (.24)	0.36 (.25)	0.32 (.25)

ability stereotypes ($p = .155$). Analyses indicated that students who received the category measure after the ability measure endorsed a higher number of category stereotypes ($M = .46$, $SE = .01$) than students who received the category measure first ($M = .42$. $SE = .01$). Test of simple effects for the stereotype X order X grade interaction indicated that the differences in category stereotypes based on counterbalancing order were evident in the middle age group ($p = .006$), but not the other age groups (early: $p = .916$; upper: $p = .386$). To account for these counterbalancing effects, order was included as a covariate in all analyses.

The same-gender role/cross-construct correlations (e.g., feminine ability with feminine category) ranged from .53–.78, suggesting that children's thinking about category and ability decisions were related strongly (see Table 4). At the same time, the pattern of correlations showed that children make distinctions between these constructs. For instance, the cross-gender role correlations within the same construct (e.g., feminine ability with masculine ability) were generally higher than the cross-gender correlations between constructs (e.g., feminine ability with masculine category).

3.2.2 | Overview of primary analyses

The data analytic approach was the same as Study 1 except that the two within-subjects variables in Study 2 included stereotype (ability or category) and gender-role (feminine or masculine). Table 5 includes raw means and standard deviations for stereotype proportions.

3.2.3 | Ability versus category

Participants' overall stereotype proportion was .49 (*SE* = .01), which indicates that, on average, students responded '*only boys/girls*' to approximately half of the occupation items across both measures. There was a main effect for gender, *F* (1, 532) = 10.89, *p* = .001, partial η^2 = .02. Pairwise comparisons revealed that boys (*M* = .52, *SE* = .01) provided a higher percentage of stereotyped responses compared to girls (*M* = .46, *SE* = .01). There was not a main effect for grade (*p* = .522)

Controlling for order, there was a main effect for stereotype, *F* (1, 532) = 44.93, *p* < .001, partial η^2 = .08. As in Study 1, students reported a higher percentage of stereotyped responses when making ability (*M* = .51, *SE* = .01) than category (*M* = .47, *SE* = .01) decisions; however, analyses revealed a stereotype X grade interaction, *F* (2, 532) = 6.06, *p* = .002, partial η^2 = .02. Tests of simple effects examining stereotype within each level of grade indicated that there was an observed stereotype difference (i.e., ability higher than category) for the middle (*p* = .002) and upper (*p* < .001) elementary students, but not the early elementary students (*p* = .117). The stereotype X gender interaction (*p* = .112) and 3-way interaction involving stereotype, gender, and grade (*p* = .177) were not significant.

3.2.4 | Feminine versus masculine

Tests of within-subjects effects involving gender role were examined to determine whether children responded differently to the feminine and masculine items. There was a main effect for gender role, *F* (1, 532) = 25.00, *p* < .001, partial η^2 = .05, indicating that children provided more stereotyped responses to feminine (*M* = .58, *SE* = .01) than masculine (*M* = .40, *SE* = .01) items. There were also two-way interactions involving gender, *F* (1, 532) = 42.93, *p* < .001, partial η^2 = .08, and grade, *F* (2, 532) = 35.06, *p* < .001, partial η^2 = .12. Tests of simple effects examining *gender* within each level of gender role indicated that when responding to masculine (*p* < .001), but not feminine (*p* = .705) items, boys (feminine: *M* = .58, *SE* = .02; masculine: *M* = .46, *SE* = .01) provided more stereotyped responses than girls (feminine: *M* = .59, *SE* = .02; masculine: *M* = .33, *SE* = .01). Tests of simple effects for *grade* revealed different developmental patterns for the feminine and masculine stereotypes (Figure 1). For the feminine stereotypes (averaged across ability and category), the early elementary students (*M* = .51, *SE* = .02) endorsed *less* stereotyped responses when compared to the middle (*M* = .61, *SE* = .02; *p* = .002) and upper (*M* = .64, *SE* = .02; *p* < .001) elementary students; there was not a difference between the middle and upper students (*p* = .649). For masculine stereotypes (averaged across ability and category), early elementary students (*M* = .44, *SE* = .02) provided *more* stereotyped responses than middle (*M* = .38, *SE* = .02; *p* = .032) and upper (*M* = .36, SE = .02; *p* = .002) elementary students, and there also was not a difference between the middle and upper students (*p* = .826). The 3-way interaction of grade, role, and stereotype was not significant (*p* = .639), indicating that patterns were consistent across ability and category responses.

3.3 | Discussion

An examination of correlations and mean differences suggest that children distinguish between their category and ability beliefs. As expected, the analyses revealed that children endorsed a higher proportion of ability than category stereotypes (medium effect size); however, this finding was only significant for the two older age groups. Consistent with Study 1, the results suggested that children in 2nd–5th grade endorsed more rigid stereotypes when making ability versus category gender comparisons. These findings mirror research within SEVT demonstrating that children distinguish between unique construct beliefs (e.g., Eccles & Wigfield, 1995, 2020), and that some beliefs (e.g., ability self-concepts) become more differentiated with age (Wan et al.,

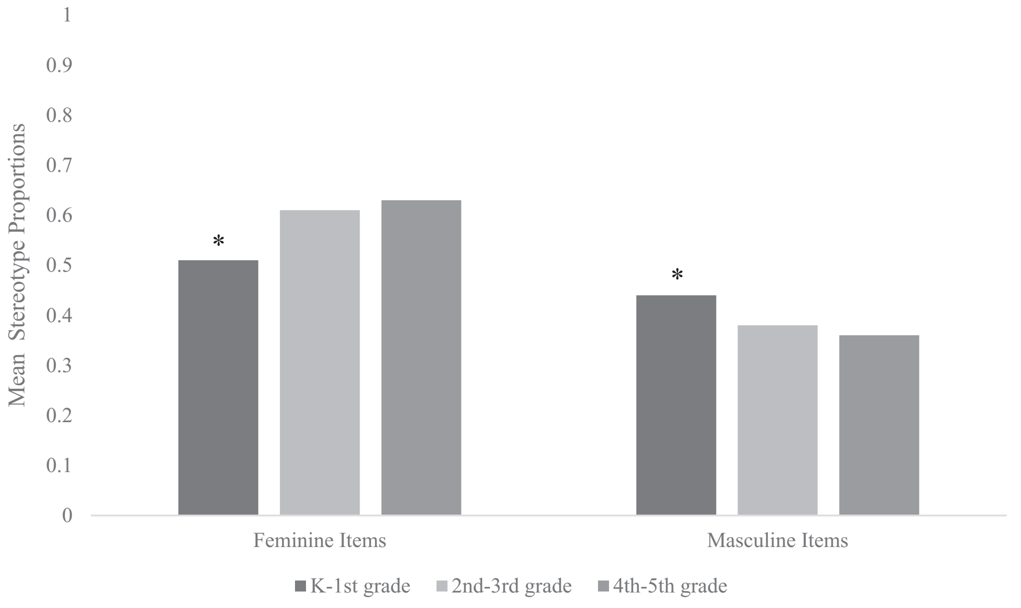

FIGURE 1 Children's mean stereotype proportions by grade level and gender role.
Note: *indicates a significant difference in stereotype proportions compared to the other grade levels.

2023). The increased pressure on performance and grades in upper elementary school students (Butler, 2005; Cimpian, 2017) might explain why older children, in particular, think more rigidly about ability than category stereotypes. As children age, they may shift their focus from who is associated with specific activities to who is good at specific activities (Miller et al., 2006) due to changes in their educational contexts.

Study 2 also examined whether children's level of stereotyping depended on the gender role of the items. Overall, children endorsed more gender stereotypes for feminine than masculine occupations, which suggests that children believe that it is more acceptable for women to be good at/associated with "masculine" occupations than it is for men to be good at/associated with "feminine" occupations. Further, although boys were more rigid than girls when responding to masculine items, there were no gender differences for the feminine items; both girls and boys reported that "only girls" are good at/associated with about 60% of the feminine occupations (Table 3). A striking pattern emerged for the grade analyses (Figure 1). Consistent with the literature, younger children showed more rigid responses than older children when responding to the masculine occupations; however, the opposite pattern emerged when the children were presented with the feminine occupations. These findings suggest that, in contrast to general conclusions in the literature, children's gendered thinking does not become consistently more flexible with age; the older children were more rigid than the younger children when deciding if men are good at/associated with feminine occupations.

The gender role results are consistent with the long-standing finding in the literature documenting that children find it more acceptable for females than males to engage in counter-stereotypical behaviors (Antill et al., 1996; Blakemore, 2003; Masters et al., 2021; Mulvey & Irvin, 2018), which has been linked to the higher status and desirability of masculine than feminine items (Canessa-Pollard et al., 2022; Liben & Bigler, 2002; Liben et al., 2001; Mulvey & Irvin, 2018). When considering females' participation in masculine occupations, increasing flexibility with age largely might account for grade differences; however, older children's increasing awareness of the lower status of feminine occupations and the harsher social sanctions of boys' cross-gender behaviors might transcend flexible thinking when making decisions about boys' participation in feminine occupations. For gender differences, the interaction of in-group bias, girls' higher flexibility, and the higher status of masculine occupations might explain why boys showed more rigidity than girls for masculine occupations, but not feminine items. Given that masculine occupations are inherently higher in status and desirability, and lower in stereotypicality than feminine occupations, it is difficult and perhaps ecologically invalid to create measures that would allow researchers to disentangle these factors (Liben & Bigler, 2002).

4 | STUDY 3

The aim was to further our understanding of the construct dimension by comparing children's interest and ability decisions within the STEM domain. Research has shown repeatedly that STEM stereotypes are associated with children's engagement and interest in STEM (Master, 2021) and that younger children endorse STEM stereotypes more than children in middle childhood and adolescence (McGuire et al., 2020). As such, we are expanding our knowledge of the multidimensional nature of gender stereotypes by examining another construct component (i.e., interest) within a domain that is highly relevant to gender equity in academic and occupational settings (i.e., STEM). Based on one study that compared children's endorsement of interest and ability stereotypes (Master et al., 2021), we expected children to endorse interest stereotypes more than ability stereotypes. We also expected that, overall, boys would endorse more stereotypes than girls, and that younger children would endorse more stereotypes than older children (McGuire et al., 2020). Given that Study 3 focused only on a masculine domain (i.e., STEM), the gender and grade hypotheses are consistent with the results of Study 2.

4.1 | Method

4.1.1 | Participants

The sample included 974 (50% female) elementary school students (43% White; 24.6% Latinx; 3.3% Black/African American; 2.5% Asian; 1.9% Native American; 13.9% multiracial/multiethnic; 10.5% non-respondents) who were recruited from four public elementary schools in the Southwestern U.S. The proportion of children on free and reduced lunch at the participating schools ranged from 37% to 52%. Children were divided into three grade groups: early (kindergarten and 1st grade: $n = 161$); middle (2nd and 3rd grade; $n = 344$); and upper (4th and 5th grade; $n = 469$).

4.1.2 | Procedure

Study 3 utilized the same procedures as Study 2; 61% of parents provided consent for participation. Data was collected in Spring 2018.

4.1.3 | Measures

In the present study, children's stereotype beliefs focused on the STEM domain (i.e., Science, Computers, Engineering and Math). For ability, gender stereotypes were assessed by asking children to report on who they think is *good at* each STEM domain (e.g., Who is good at math?). Interest gender stereotypes were assessed by asking children to report on who they think *likes* each STEM domain (e.g., Who do you think likes math?). Responses were rated as 0 (*only boys*), 1 (*more boys than girls*), 2 (*both boys and girls*), 3 (*more girls than boys*), or 4 (*only girls*). To represent children's endorsement of cultural gender stereotypes, scores were based on the proportion of responses that included "*only boys*" or "*more boys than girls*".

TABLE 6 Study 3: Intercorrelations by study variables disaggregated by gender and grade.

	K/1st		2nd/3rd		4th/5th	
Variable	1	2	1	2	1	2
1. Ability	_	.40**	_	.70**	_	.50**
2. Category	.54**	_	.60**	_	.64**	_

Note: The results for the female sample are shown above the diagonal. The results for the male sample are shown below the diagonal.
**$p < .01$.

TABLE 7 Study 3: Stereotype proportions by gender and grade.

Variables	Means (*SD*)		Means (*SD*)		
	Boys	Girls	K-1st grade	2^{nd}–3rd grade	4th–5th grade
Gender ability stereotype	0.35 (.29)	0.22 (.23)	0.41 (.26)	0.31 (.29)	0.22 (.25)
Gender interest stereotype	0.37 (.29)	0.25 (.23)	0.43 (.25)	0.33 (.29)	0.25 (.25)

4.2 | Results

4.2.1 | Preliminary analyses

We first conducted a mixed-design ANOVA (with gender and grade) to assess whether counterbalancing order affected stereotype proportions. All interaction effects involving order were not significant: stereotype X order ($p = .269$), grade X order ($p = .054$), gender X order ($p = .307$), and grade X gender X order ($p = .714$). Thus, order was not included as a covariate.

The correlations between interest and ability ranged from .40 to .70 (see Table 6). Like Studies 1 and 2, the correlations show that children show moderate to strong consistency in their gender stereotype decisions across different constructs. At the same time, the strength of the correlations suggests a moderate level of discriminant validity between interest and ability.

4.2.2 | Overview of primary analyses

The analytic approach was the same as Studies 1 and 2 except that Study 3 included one within-subjects variable: stereotype (interest or ability). Table 7 includes raw means and standard deviations for stereotype proportions.

4.2.3 | Primary analyses

Participants' overall stereotype proportion was .32 ($SE = .01$), which indicates that, on average, students responded "*only boys*" or "*more boys than girls*" to less than half of the STEM items across both measures. It is noteworthy that students' responses varied across the four STEM domains. For interest, the percentage of students who chose either "only boys" or "more boys than girls" was: science (21.5), computers (31.0), math (22.1) and engineering (49.1). For ability, the percentages were: science (19.9), computers (27.5), math (18.4) and engineering (47.5).

The mixed-design ANOVA revealed a main effect for gender, F (1, 968) = 73.28, $p < .001$, partial $\eta^2 = .07$, indicating that boys ($M = .39$, $SE = .01$) provided more stereotyped responses compared to girls ($M = .26$, $SE = .01$). There also was a main effect for grade, F (2, 968) = 38.28, $p < .001$, partial $\eta^2 = .07$, indicating early elementary students ($M = .41$, $SE = .02$) reported a higher proportion of stereotyped responses compared to both middle ($M = .32$, $SE = .01$; $p < .001$) and late elementary ($M = .24$, $SE = .01$; $p < .001$) students; the difference between middle and late students also was significant ($p < .001$). The *within-subjects effect* for stereotype was significant, F (1, 968) = 9.17, $p = .003$, partial $\eta^2 = .009$, which revealed that, overall, children were more likely to choose "only boys" or "more boys than girls" when making interest ($M = .34$, $SE = .01$) than ability ($M = .31$, $SE = .01$) decisions. The stereotype X grade ($p = .594$), stereotype X gender ($p = .323$), and stereotype X grade X gender ($p = .736$) interactions were not significant.

4.3 | Discussion

As expected, analyses revealed construct differences such that children endorsed a higher percentage of interest stereotypes compared to ability stereotypes. Although the effect size was small, these results provide further evidence that children make distinctions between constructs when thinking about gender stereotypes. In this case, children, across gender and age categories, were more likely to indicate that girls are good at STEM subjects than like STEM subjects. Both girls and boys generally are expected to show competence in school and meet academic standards, yet they do not have the same pressure to "like" all subjects. Thus, children might be more likely to believe that even if girls are required to be "good at" STEM, they do not necessarily have to "like" STEM.

The gender and grade hypotheses also were confirmed. Consistent with Study 2 and the STEM literature (Master et al., 2021; McGuire et al., 2020), boys and younger children endorsed more masculine stereotypes than girls and older children. Because our primary interest was comparing stereotype endorsement for ability and interest STEM stereotypes, we calculated the proportion of "only boys" responses instead of also examining "only girls" and "both girls and boys" responses. In this case, we have limited data to interpret patterns of ingroup bias. A recent study that compared equitable, male-bias, and female-bias endorsement for STEM ability found that, in contrast to males, females in middle childhood did not continue to show the in-group bias that was apparent in early childhood (McGuire et al., 2020). It would be interesting for future research to investigate whether gender and developmental patterns of ingroup bias differ for unique stereotype constructs.

5 | GENERAL DISCUSSION

The present research advances our understanding of children's gender stereotype beliefs in several ways. Across three studies relevant to children's academic and career beliefs, we moved beyond a broad understanding of gender stereotype development to examine whether elementary school children's endorsement of stereotypes varied by construct and gender role. In Studies 1 and 2, we found that older elementary children were more likely to endorse ability than category stereotypes. In Study 3, the results indicated that, across gender and age, children endorsed STEM interest stereotypes more than STEM ability stereotypes. In Study 2, findings also revealed developmental differences in how children think about masculine versus feminine stereotypes. For masculine stereotypes, our findings confirmed existing conclusions in the literature by documenting that children's stereotyped thinking becomes more flexible with age; however, for feminine stereotypes, the reverse pattern was found such that older children showed more stereotyped thinking than younger children.

5.1 | Stereotype construct differences

Overall, the findings from the present research suggest that there are construct distinctions in how children think about gender stereotypes. This study is the first to examine children's developing beliefs for category stereotypes and to compare children's beliefs for ability and category stereotypes. Findings revealed that older children (2nd–5th grade) were more flexible when making category than ability decisions and, consistent with the findings from Master et al. (2021), children across all grade groups, endorsed more interest than ability stereotypes. These results suggest that even when children believe that all children can be associated with a stereotyped activity, they still might hold gender stereotyped views regarding who likes and who is good at that activity.

Children's conceptions of fairness might explain why they are more flexible considering category versus ability gender stereotypes. Research has found that even when children are aware of gender stereotypes, they tend to believe that it is morally wrong to exclude a child from an activity due to their gender (e.g., Killen & Stangor, 2001). Thus, children may link category stereotypes with exclusion decisions (e.g., who is or is not allowed to participate in this activity). To illustrate, children may think that girls should not be excluded from pursuing a career in math (e.g., math is for both girls and boys), but still endorse the belief that boys are better at math when compared to girls. Consistent with this idea, a recent study asked highschoolers to make decisions about a hypothetical robotics project and found that participants were less likely to make male-bias choices when choosing who they would pick to *join* their group compared to choosing who they think is *good* at "building robots" (Joy et al., 2023). It will be important for future research to examine whether children's conception of category stereotypes overlaps with their beliefs about who should engage in or be included in specific activities.

It is noteworthy that kindergarten and first grade children did not make a distinction between category and ability stereotypes. Cognitive changes in children's ability conceptions may underlie these age differences. Research indicates that starting around age 7, children become more interested in ability comparisons (Dweck, 2002), report that ability is more likely to be stable (e.g., Spinath & Steinmayr, 2008) and recognize domain differences in ability (e.g., math versus reading skill; Wigfield et al., 1997). Further, children younger than 7 tend to indicate that effort is likely to predict performance (Heyman et al., 2003; Muenks & Miele, 2017). Taken together, early elementary children may not yet distinguish between a child being associated with an activity and being good at an activity. It is important to note that these age differences may be due to changes in children's contexts (e.g., increased pressure on school achievement) rather than deficits in younger children's cognitions (Butler, 2005; Cimpian, 2017).

The findings from Study 3 support the view that elementary children endorse STEM interest stereotypes more strongly than STEM ability stereotypes (Master et al., 2021). This finding suggests that even if children believe that both girls and boys are good at a STEM subject, they may still think that boys like that subject more than girls. Children might be more comfortable endorsing interest stereotypes compared to ability stereotypes because interest statements are less evaluative and less linked to school success. The belief that boys *like* math more than girls is less disparaging than the belief that boys are *better* than girls at math. Further, interest stereotypes might be one of the first gender stereotypes that children are exposed to because activity/toy stereotypes have been shown to emerge before trait stereotypes (e.g., Miller et al., 2006; boys like trucks versus boys are good at math). Thus, for elementary school students, interest stereotypes might be more consolidated and engrained when compared to ability stereotypes. Further, research has shown that interest gender stereotypes predict children's interest and motivation in STEM more than ability gender stereotypes (Master et al., 2021). Interest gender stereotypes might be especially salient for elementary school children and have a strong influence on their motivation to engage in STEM activities. At the same time, ability stereotypes are related to important STEM outcomes as demonstrated in a recent study that found endorsing a science ability stereotype was positively related to selecting a boy for help with a science question (McGuire et al., 2022). Thus, future research should investigate if distinct stereotype constructs uniquely and jointly predict STEM outcomes.

5.2 | Gender role (Feminine versus masculine stereotypes)

The gender development literature includes decades of evidence documenting that children's gender stereotype attitudes become more flexible with age (e.g., Leaper, 2015; Ruble et al., 2006). However, the present research shows that these developmental distinctions depend on whether children are considering feminine or masculine stereotypes. Consistent with recent research (Canessa-Pollard et al., 2022), older children showed more flexibility than younger children when asked about masculine occupations, but the reverse pattern emerged for feminine occupations. These findings illustrate the importance of considering the dimension of gender role when examining and drawing conclusions about gender stereotype development.

As mentioned earlier, older elementary school children's increasing awareness of status differences might explain why they become more rigid when considering who is linked with or good at feminine occupations. During middle childhood, children begin to recognize that men have higher status and power than women (Liben et al., 2001; Neff et al., 2007), and that boys are more likely than girls to be excluded for engaging in gender nonconforming activities (Mulvey & Killen, 2015). Thus, older children appear to recognize that females increase their status by engaging in masculine activities, but males decrease their status by engaging in feminine activities (Mulvey & Irvin, 2018). Children's exposure to the inequitable status of gender roles likely is exemplified further by the widespread outreach programs encouraging females to participate in male-dominated fields (e.g., STEM) and the absence of programs encouraging males to participate in female-dominated fields (e.g., early childhood education). Thus, there is a need for interventions to focus on increasing the status of feminine occupations while also fostering acceptance of males who engage in gender nonconforming activities.

5.3 | Gender differences

Overall, when asked about gender stereotyped occupations or STEM in Studies 2 and 3, boys showed higher stereotype endorsement than girls. These results are consistent with general conclusions that girls tend to be more flexible than boys when thinking about and applying gender stereotypes (e.g., Canessa-Pollard et al., 2022; Wood et al., 2021). Although in Study 1 girls also endorsed fewer items than boys when asked about neutral activities and skills, these differences were not significant. The null gender findings when asked about neutral items suggest that boys' higher rigidity than girls may be specific to gender stereotypes rather than a generalized tendency to respond in an inflexible manner. It is noteworthy that gender differences also did not emerge when children were asked about feminine occupations. It seems that both girls and boys similarly recognize the costs associated with males engaging in feminine activities (Masters et al., 2021). It also is possible that girls' higher flexibility compared to boys is overshadowed by an ingroup bias when it comes to thinking about feminine occupations or activities (Kurtz-Costes et al., 2014). Specifically, girls might feel protective of the occupations that are associated with their gender group and feel that it is in their best interest to exclude boys from these settings.

5.4 | Limitations and future directions

The present studies advance our understanding of the multidimensional nature of gender stereotype beliefs and offer exciting avenues for future research. There are several limitations, however, that need to be considered when interpreting the results. First, the effect sizes for the stereotype analyses were small to medium; although children appeared to make distinctions based on stereotype construct, the mean differences were generally small. It is possible that the rating scales used in the measures did not allow for larger differences to emerge. Future research should consider using varied measurement tools to assess construct differences to determine whether findings are dependent on methodology. Second, this research employed a within-subjects design, which showed some order effects in Studies 1 and 2. Although within-subject designs have several benefits (e.g., comparing participants to themselves), our research supports existing conclusions that the order of questions can affect findings (Oldendick, 2008). It would be beneficial for future studies to analyze within-subject and between-subject data simultaneously to disentangle and explain order effects in stereotype research. A third limitation of this study is that data were summarized across items to index children's overall stereotyping. Yet, as shown in Study 3, children show stronger stereotyped beliefs for some items more than others (i.e., engineering and computers were more stereotyped than other STEM domains). These findings are consistent with a recent study that found children endorsed stronger gender ability stereotypes for engineering and technology than science and mathematics (McGuire et al., 2022). Thus, it would be valuable to analyze data by items to identify the activities and occupations that might require more targeted interventions (McGuire et al., 2022). Relatedly, given young children's limited knowledge of the engineering field (Ozogul et al., 2017), it is important to be cautious about interpreting their beliefs in this domain. To address this issue for future research, we developed a new measure for elementary students that asks about developmentally relevant "engineering-related activities and skills" (Wheeler et al., 2022).

We look forward to future research that not only addresses the limitations of the present research, but extends the multidimensional framework of gender stereotypes in new directions. In particular, it is essential to extend this research by implementing longitudinal or experimental designs to examine whether specific constructs predict or influence different types of behaviors like expectancy and task-value beliefs in SEVT (e.g., Guo et al., 2015). For instance, it is important to investigate whether distinct stereotype constructs differentially impact academic and career outcomes, and whether these patterns change across the lifespan. It also would be valuable to include additional constructs (such as attainment value in SEVT; Eccles and Wigfield, 2020) and domains (verbal versus STEM stereotypes; Skinner et al., 2021) to determine whether findings vary for these dimensions. Finally, it is critical to examine race differences in future studies as gender stereotyped attitudes have been shown to vary across racial groups (e.g., Skinner et al., 2021).

6 | CONCLUSIONS

Across three studies, the current research provided evidence that elementary school children consider constructs and gender role when making stereotyped decisions. Older children endorsed ability stereotypes more strongly than category stereotypes and, across all age groups, children endorsed interest stereotypes more strongly than ability stereotypes. Findings also revealed that whereas children show more flexible beliefs with age when presented with masculine items, the reverse pattern emerged for feminine items. The present research illustrates the benefits of employing a multidimensional framework to gain a more nuanced understanding of how children make sense of their increasing knowledge of gender stereotypes. Future research that implements longitudinal, multi-method, and experimental designs has the potential to offer valuable insights on the influence of construct-specific stereotype beliefs across development.

ACKNOWLEDGMENTS

We are grateful to the teachers, parents, and youth who participated in this project, and to the school districts for collaboration. We thank Julie Rotberg, Fabianne Gondim, Donna Chen, and the many undergraduate research assistants for their assistance in conducting this investigation. Portions of the data included in this manuscript were presented at the Eighth Gender Development Research Conference in San Francisco, CA. (October 2018). This work was supported by the National Science Foundation under grant number [DRL-1561424] to Arizona State University. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data is not yet available in a public repository.

ETHICS STATEMENT

The research protocols used in this research were approved by the ethics committee of Arizona State University, Tempe, AZ, United States, [ASU IRB Protocol #STUDY00003903].

ORCID

Cindy Faith Miller https://orcid.org/0000-0003-0003-5011

REFERENCES

Antill, J. K., Cotton, S., Russell, G., & Goodnow, J. J. (1996). Measures of children's sex-typing in middle childhood, II. *Australian Journal of Psychology*, *48*(1), 35–44. https://doi.org/10.1080/00049539608259503

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, *88*(4), 354–364. https://doi.org/10.1037/0033-295X.88.4.354

Bigler, R. S., & Liben, L. S. (1993). A cognitive-developmental approach to racial stereotyping and reconstructive memory in Euro-American children. *Child Development*, *64*(5), 1507–1518. https://doi.org/10.1111/j.1467-8624.1993.tb02967.x

Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science*, *16*(3), 162–166. https://doi.org/10.1111/j.1467-8721.2007.00496.x

Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex Roles*, *48*(9), 411–419. https://doi.org/10.1023/A:1023574427720

Butler, R. (2005). Competence assessment, competence, and motivation between early and middle childhood. In A. J. Elliott & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 202–221). Guilford Press.

Canessa-Pollard, V., Reby, D., Banerjee, R., Oakhill, J., & Garnham, A. (2022). The development of explicit occupational gender stereotypes in children: Comparing perceived gender ratios and competence beliefs. *Journal of Vocational Behavior*, *134*, 103703. https://doi.org/10.1016/j.jvb.2022.103703

Cimpian, A. (2017). Early reasoning about competence is not irrationally optimistic, nor does it stem from inadequate cognitive representations. In (A. J. Elliot, C. S. Dweck, & D. S. Yeager Eds.), *Handbook of competence and motivation: Theory and application* (pp. 387–407). The Guilford Press.

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

Dweck, C. (2002). The development of ability conceptions. In (A. Wigfield & J.S. Eccles Eds.), *Development of Achievement Motivation* (pp. 57–88). Elsevier, Inc. https://doi.org/10.1016/B978-0-12-750053-9.X5000-1

Eccles, J. S. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development*, *35*(3), 195–201. https://doi.org/10.1177/0165025411398185

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, *21*(3), 215–225. https://doi.org/10.1177/0146167295213003

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Garriott, P. O., Hultgren, K. M., & Frazier, J. (2017). STEM stereotypes and high school students' math/science career goals. *Journal of Career Assessment*, *25*(4), 585–600. https://doi.org/10.1177/1069072716665825

Guo, J., Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2015). Directionality of the associations of high school expectancy-value, aspirations, and attainment: A longitudinal study. *American Educational Research Journal*, *52*(2), 371–402. https://doi.org/10.3102/0002831214565786

Heyman, G. D., Gee, C. L., & Giles, J. W. (2003). Preschool children's reasoning about ability. *Child Development*, *74*(2), 516–534. https://www.jstor.org/stable/3696328

Hilliard, L. J., & Liben, L. S. (2010). Differing levels of gender salience in preschool classrooms: Effects on children's gender attitudes and intergroup bias. *Child Development*, *81*(6), 1787–1798. https://doi.org/10.1111/j.1467-8624.2010.01510.x

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist*, *74*(2), 171–193. https://doi.org/10.1037/amp0000307

Joy, A., Mathews, C. J., Cerda-Smith, J., Knox, J. L., & Mulvey, K. L. (2023). The roles of school racial climate and science, technology, engineering, and mathematics classroom climate in adolescents' gender and racial peer inclusion and attribution decisions. *Journal of Experimental Child Psychology*, *226*, 105573. https://doi.org/10.1016/j.jecp.2022.105573

Killen, M., & Stangor, C. (2001). Children's social reasoning about inclusion and exclusion in gender and race peer group contexts. *Child Development*, *72*(1), 174–186. https://www.jstor.org/stable/1132478

King, R. A., Scott, K. E., Renno, M. P., & Shutts, K. (2020). Counterstereotyping can change children's thinking about boys' and girls' toy preferences. *Journal of Experimental Child Psychology*, *191*, 104753. https://doi.org/10.1016/j.jecp.2019.104753

Kurtz-Costes, B., Copping, K. E., Rowley, S. J., & Kinlaw, C. R. (2014). Gender and age differences in awareness and endorsement of gender stereotypes about academic abilities. *European Journal of Psychology of Education*, *29*(4), 603–618. https://doi.org/10.1007/s10212-014-0216-7

Leaper, C. (2015). Gender and social-cognitive development. In (L. S. Liben, U. Müller, & R. M. Lerner Eds.), *Handbook of child psychology and developmental science: Cognitive processes* (pp. 806–853). John Wiley & Sons Inc. https://doi.org/10.1002/9781118963418.childpsy219

Levanon, A., England, P., & Paul Allison, U. O. P. (2009). Occupational feminization and pay: Assessing causal dynamics using 1950–2000 U.S. Census data. *Social Forces*, *88*(2), 865–891. https://doi.org/10.1353/sof.0.0264

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*(2), vii–147. https://doi.org/10.1111/1540-5834.t01-1-00187

Liben, L. S., Bigler, R. S., & Krogh, H. R. (2001). Pink and blue collar jobs: Children's judgements of job status and job aspirations in relation to sex of worker. *Journal of Experimental Child Psychology*, *79*(4), 346–363. https://doi.org/10.1006/jecp.2000.2611

Martin, C. L., & Ruble, D. (2004). Children's search for gender cues: Cognitive perspectives on gender development. *Current Directions in Psychological Science*, *13*(2), 67–70. https://doi.org/10.1111/j.0963-7214.2004.00276.x

Martin, C. L., Ruble, D. N., & Szkrybalo, J. (2002). Cognitive theories of early gender development. *Psychological Bulletin*, *128*(6), 903–933. https://doi.org/10.1037/0033-2909.128.6.903

Master, A. (2021). Gender stereotypes influence children's STEM motivation. *Child Development Perspectives*, *15*(3), 203–210. https://doi.org/10.1111/cdep.12424

Master, A., Meltzoff, A. N., & Cheryan, S. (2021). Gender stereotypes about interests start early and cause gender disparities in computer science and engineering. *Proceedings of the National Academy of Sciences*, *118*(48), e2100030118. https://doi.org/10.1073/pnas.2100030118

Masters, S. L., Hixon, K., & Hayes, A. R. (2021). Perceptions of gender norm violations among middle school students: An experimental study of the effects of violation type on exclusion expectations. *Journal of Early Adolescence*, *41*(4), 527–549. https://doi.org/10.1177/0272431620931193

McGuire, L., Hoffman, A. J., Mulvey, K. L., Hartstone-Rose, A., Winterbottom, M., Joy, A., Law, F., Balkwill, F., Burns, K. P., Butler, L., Drews, M., Fields, G., Smith, H., & Rutland, A. (2022). Gender stereotypes and peer selection in STEM domains among children and adolescents. *Sex Roles*, *87*, 455–470. https://doi.org/10.1007/s11199-022-01327-9

McGuire, L., Mulvey, K. L., Goff, E., Irvin, M. J., Winterbottom, M., Fields, G. E., Hartstone-Rose, A., & Rutland, A. (2020). STEM gender stereotypes from early childhood through adolescence at informal science centers. *Journal of Applied Developmental Psychology*, *67*, 101109. https://doi.org/10.1016/j.appdev.2020.101109

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its consequences for young adolescents' course enrollment intentions and performances in mathematics. *Journal of Educational Psychology*, *82*, 60–70. https://doi.org/10.1037/0022-0663.82.1.60

Midway, S., Robertson, M., Flinn, S., & Kaller, M. (2020). Comparing multiple comparisons: practical guidance for choosing the best multiple comparisons test. *PeerJ*, *8*, e10387. https://doi.org/10.7717/peerj.10387

Miller, 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*(11), 870–881. https://doi.org/10.1007/s11199-009-9584-x

Miller, C. F., Martin, C. L., Fabes, R. A., & Hanish, L. D. (2013). Bringing the cognitive and social together: How gender detectives and gender enforcers shape children's gender development. In (M. R. Banaji & S. A. Gelman Eds.), *Navigating the social world: What infants, children, and other species can teach us*, (pp. 306–313). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199890712.003.0056]

Miller, C. F., Trautner, H. M., & Ruble, D. N. (2006). The role of gender stereotypes in children's preferences and behavior. In (L. Balter & C. S. Tamis-LeMonda Eds.), *Child psychology: A handbook of contemporary issues*, (2nd edn.). Psychology Press.

Muenks, K., & Miele, D. B. (2017). Students' Thinking About Effort and Ability: The Role of Developmental, Contextual, and Individual Difference Factors. *Review of Educational Research*, *87*(4), 707–735. https://doi.org/10.3102/0034654316689328

Mulvey, K. L., & Irvin, M. J. (2018). Judgments and reasoning about exclusion from counter-stereotypic STEM career choices in early childhood. *Early Childhood Research Quarterly*, *44*, 220–230. https://doi.org/10.1016/j.ecresq.2018.03.016

Mulvey, K. L., & Killen, M. (2015). Challenging gender stereotypes: Resistance and exclusion. *Child Development*, *86*(3), 681–694. https://doi.org/10.1111/cdev.12317
Neff, K. D., Cooper, C. E., & Woodruff, A. L. (2007). Children's and adolescents' developing perceptions of gender inequality. *Social Development*, *16*(4), 682–699. https://doi.org/10.1111/j.1467-9507.2007.00411.x
Oldendick, R. W. (2008). Question order effects. In (P. J. Lavrakas Ed.), *Encyclopedia of survey research methods* (pp. 663–665). Sage Publications, Inc. https://dx.doi.org/10.4135/9781412963947
Ozogul, G., Miller, C. F., & Reisslein, M. (2017). Latinx and Caucasian elementary school children's knowledge of and interest in engineering activities. *Journal of Pre-College Engineering Education Research (J-PEER)*, *7*(2), 1–12. https://doi.org/10.7771/2157-9288.1122
Patterson, M. M. (2012). Self-perceived gender typicality, gender-typed attributes, and gender stereotype endorsement in elementary-school-aged children. *Sex Roles*, *67*(7), 422–434. https://doi.org/10.1007/s11199-012-0184-9
Pillow, B. H., Vilma, T., & Low, N. (2022). The influence of gender categories and gender stereotypes on young children's generalizations of biological and behavioral characteristics. *Psychological Reports*, *125*(1), 328–343. https://doi.org/10.1177/0033294120973933
Powlishta, K. K. (2000). The effect of target age on the activation of gender stereotypes. *Sex Roles*, *42*(3), 271–282. https://doi.org/10.1023/A:1007095307977
Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods*, *25*(1), 6–14. https://doi.org/10.1177/1094428120968614
Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In (N. Eisenberg, W. Damon, & R. M. Lerner Eds.), *Handbook of child psychology: Vol. 3, social, emotional, and personality development* (6th edn.) (pp. 858–932). John Wiley & Sons Inc. https://doi.org/10.1002/9780470147658.chpsy0314
Shapiro, C. A., & Sax, L. J. (2011). Major selection and persistence for women in STEM. *New Directions for Institutional Research*, *2011*, 5–18. https://doi.org/10.1002/ir.404
Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children′ s gender schemata about others: A meta-analytic review. *Developmental review*, *13*(2), 147–183. https://doi.org/10.1006/drev.1993.1007
Skinner, O. D., Kurtz-Costes, B., Vuletich, H., Copping, K., & Rowley, S. (2021). Race differences in Black and white adolescents' academic gender stereotypes across middle and late adolescence. *Cultural Diversity and Ethnic Minority Psychology*, *27*(3), 537–545. https://doi.org/10.1037/cdp0000384
Spinath, B., & Steinmayr, R. (2008). Longitudinal analysis of intrinsic motivation and competence beliefs: Is there a relation over time? *Child Development*, *79*(5), 1555–1569. https://doi.org/10.1111/j.1467-8624.2008.01205.x
Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology*, *102*(4), 947–963. https://doi.org/10.1037/a0019920
Trautner, H. M., Ruble, D. N., Cyphers, L., Kirsten, B., Behrendt, R., & Hartmann, P. (2005). Rigidity and flexibility of gender stereotypes in children: Developmental or differential? *Infant and Child Development*, *14*(4), 365–381. https://doi.org/10.1002/icd.399
Wan, S., Lauermann, F., Bailey, D. H., & Eccles, J. S. (2023). Developmental changes in students' use of dimensional comparisons to form ability self-concepts in math and verbal domains. *Child Development*, *94*, 272–287. https://doi.org/10.1111/cdev.13856
Weisgram, E. S. (2016). The cognitive construction of gender stereotypes: Evidence for the dual pathways model of gender differentiation. *Sex Roles*, *75*(7), 301–313. https://doi.org/10.1007/s11199-016-0624-z
Wheeler, L. A., Miller, C. F., Chen, D., & Woods, B., & Reisslein, M. (2022). A suite of measures for children's achievement beliefs in engineering-related activities and skills. *European Journal of Engineering Education*, *47*(6), 1335–1355. https://doi.org/10.1080/03043797.2023.2169600
Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, *89*(3), 451–469. https://doi.org/10.1037/0022-0663.89.3.451
Wolff, F. (2021). How classmates' gender stereotypes affect students' math self-concepts: A multilevel analysis. *Frontiers in Psychology*, *12*, 599199. https://doi.org/10.3389/fpsyg.2021.599199
Wood, L. A., Hutchison, J., Aitken, M., & Cunningham, S. J. (2021). Gender stereotypes in UK children and adolescents: Changing patterns of knowledge and endorsement. *British Journal of Social Psychology*, *61*(3), 768–789. https://doi.org/10.1111/bjso.12510

How to cite this article: Miller, C. F., Wheeler, L. A., & Woods, B. (2023). A multidimensional examination of children's endorsement of gender stereotypes. *Social Development*, 1–22. https://doi.org/10.1111/sode.12725