Social Development

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Preschoolers' Gender Identification Rigidity Relates to Their Gender-Typed Predictions for Others

Christine Fawcett[1,2] | Kahl Hellmer[2]

[1]Department of Psychology, Stockholm University, Stockholm, Sweden | [2]Department of Psychology, Uppsala University, Uppsala, Sweden

Correspondence: Christine Fawcett (Christine.fawcett@psychology.su.se)

Funding: Funding for the current project was provided by Skandia Ideas for Life.

Keywords: gender development | gender identity | gender stereotypes | preschool age | social cognition | toy preferences

ABSTRACT

Children begin to reason about gender and others' gender-typed preferences from early in life, yet not enough is known about whether their reasoning reflects only binary categorization or a more nuanced way understanding of variation in gender. Further, little is known about how children's conception of their own gender affects how they think about others. In the current study, 3–6-year-old preschool children ($n = 56$) were asked to predict preferred and nonpreferred toys that included feminine, neutral, and masculine options for other children who were described using their binary gender identity (i.e., a gender-typical hairstyle, name, pronouns) and their home environment (i.e., room decoration and toys which were provided by their parents and were either feminine, neutral, or masculine). Children also participated in tasks to assess their own gender-typed toy preferences and their own gender identity using two continuous scales for identification with girls and with boys and their gender-typed behavior was rated by their preschool teachers. Results indicated that children considered both binary gender and information about the home environment when predicting others' toy preferences. Further, children's greater reliance on an individual's binary gender identity when allocating toys was related to the rigidity of their own gender identity and their own gender-typed toy preferences. Together, these findings add nuance to our understanding of children's developing reasoning about gender identity and preferences and reinforce that children can conceptualize gender as more than a binary category.

1 | Introduction

The traditional view of gender as a discrete, binary category has in recent years increasingly been shown to be an oversimplification of the true variation that exists. That is, gender not only incorporates different domains (e.g., identity, social roles, appearance), but each of these domains can also include continuous variation beyond the binary (Hyde et al. 2019; Martin, Cook, et al. 2017). Yet a binary conceptualization of gender is still present in much of the research on child development, particularly when it comes to children's reasoning about gender. Stimuli used in developmental research often only include contrasting masculine and feminine options (e.g., Gelman et al. 1986; King et al. 2020; Martin 1989; Shutts et al. 2017), or focus on whether children categorize novel toys as "for" their gender or not (e.g., Lam 2023; Martin et al. 1995), preventing children from showing nuance in their responses. To better understand how children conceptualize and reason about gender, it is critical to use methods that allow them to respond with more than just binary options. Here we examine whether preschool-aged children (3–6 years of age) combine different information sources to make more nuanced predictions about others' gender-typed preferences. We also explore whether individual differences in aspects of children's own gender (identity, expression,

preferences, etc.) or their age relate to how they make their predictions.

1.1 | Children's Inferences About Gender Typing in Others

By early preschool age, children tend to make gender-typed predictions about others based on identity cues (i.e., gender-typed appearance, name, and pronouns). For example, when presented with unfamiliar toys that lack a gender association, 2–6-year-old children tend to expect that others of their same binary gender will like the toys they do, and their own interest in such toys is influenced by adults' labeling of them as being "for" girls or boys (Lam 2023; Martin et al. 1995). Young children also prioritize binary gender identity over other conflicting cues to a child's preferences, such as their physical appearance (4–7-year-olds; Gelman et al. 1986), or even being described as having preferences atypical for their gender (4–6-year-olds; Martin 1989). However, in some cases, children can take gender *counter*-stereotypical characteristics into account, such as 4–8-year-olds trusting information from a source with expertise in a counter-stereotypical domain (e.g., a boy who is an expert in ballet; Boseovski et al. 2016). Further, more recent research has shown that young children do not see a nonbinary gender identity (i.e., someone described as "neither a girl or a boy") as informative for making predictions about others' traits or preferences (Yuly-Youngblood and Boseovski 2022). That is, when asked to generalize information about a person who "doesn't look like a boy or a girl" or "is not a boy or a girl" from either a person with binary gender or another nonbinary person, 5–6-year-old children did not consistently generalize from one nonbinary person to another, as they typically do for people of the same binary gender. This suggests that it may only be binary gender categories that children rely on for social inferences and that these inferences could be based on children's developing ideas about girls and boys as social categories.

Fewer studies have examined whether children are sensitive to variation in a child's environment as a potential source for gender-typed preferences. That is, while studies examining contrasts between gender identity and appearance or preferences can shed light on whether children factor in different aspects of an individual's gender expression or roles, they can say less about where children think gender-typed preferences originate. By contrasting a hypothetical child's identity with their environment, we can determine whether children believe that both internal and external influences impact preferences. This approach is more in line with the classic "switched-at-birth" tasks in which a child is described as growing up with only members of a different category (e.g., a boy growing up with only girls and women) and participants are asked to predict a variety of characteristics such as their behaviors, preferences, and physical features (Taylor 1996). Such tasks help to determine whether children believe such characteristics are based on internal factors that are present from birth, or whether they are based on the environment one is raised in. For example, 5- and 6-year-olds tend to state that gender-typical physical features are based on one's gender at birth while behaviors are somewhat more influenced by the environment and more flexible (Taylor et al. 2009). A downside of this method, however, is that it requires an unrealistic, rather extreme hypothetical, which children likely find difficult to relate to.

A shortcoming of much of the earlier research on children's inferences about gender-typed preferences is that they typically pit gender identity against another cue (e.g., appearance, activity preference, or environment), forcing children to choose either one cue or the other as the source of their inference. In order to examine whether children are able to combine different information sources when making predictions, we need to allow them to respond in a more nuanced manner by allowing a broader range of options, rather than forcing them into binary choices that are either in line with gender stereotyping or not. For example, a recent study allowed 3–5-year-old children to allocate a toy from an array of gender-typed and neutral toy options to hypothetical boy and girl peers who were described as having masculine, feminine, or neutral activity preferences (Wang et al. 2023). Children's toy allocations reflected primarily the gender typing of the peers' activity preferences, but also to some degree the peers' binary gender (as indicated by name, pronouns, and a stick-figure illustration with gender-typical hair and clothing), such that toy selections at a group level were shifted toward gender-typical choices. This suggests that children can take into account factors beyond binary gender identity when predicting others' preferences, particularly when they are able to respond with more nuanced options.

1.2 | Individual Differences in Children's Inferences About Gender

Another underexamined topic regarding children's reasoning about gender is individual differences, which can give critical insight into underlying mechanisms. Age differences can give some insight into the changes in children's reasoning over time. It is thought that age changes can be seen as a proxy for shifts in children's growing understanding of what gender encompasses as well as their ability to think flexibly about gender. Rigid belief in gender stereotypes for others' preferences (Halim et al. 2014; Martin and Ruble 2009; Trautner et al. 2005) and rejection of gender norm violations (Conry-Murray et al. 2015; Riggs et al. 2023) seem to peak at around 5–6 years of age, with increasing flexibility afterward. This shift toward flexibility can also be seen in 3–10-year-old children's transition from selecting idealized, extreme, and narrow gender prototypes to more flexible and broad gender prototypes as they get older (Foster-Hanson and Rhodes 2023). Evidence further suggests that children's social cognitive development, such as their Theory of Mind, can underlie shifts in flexible thinking about gender. For example, with a more developed Theory of Mind, 4–6-year-old children are more likely to allocate resources based on merit instead of gender stereotypic expectations, such as rewarding a boy who is good at making dolls (Rizzo and Killen 2018).

Group-level differences between how boys and girls respond can also be informative when it comes to sources of children's reasoning about gender. However, research findings have shown mixed results when it comes to these effects. While many studies show no significant differences between boys and girls (e.g., 3–7 years: Ruble et al. 2007; 4–7 years: Spinner et al. 2018), some studies suggest that girls are more likely to hold rigid stereotypes

or view counter-stereotypical behaviors negatively (e.g., 2 years: Poulin-Dubois et al. 2002; 3–6 years: Shutts et al. 2017; 3–5 years: Wang et al. 2023), and others suggest that boys do so (4–6 years: Banerjee and Lintern 2000; Ebert et al. 2024; 6–8 years: Riggs et al. 2023; Skočajić et al. 2020). For example, in one study, 3–5-year-old girls were more likely than boys to assume that others of the same gender shared biological and behavioral properties (Pillow et al. 2015). Three- to 10-year-old girls spontaneously named more gender stereotypes than boys, particularly in the domain of appearance (Miller et al. 2009). It could be that contextual or cultural factors partly contribute to these differences—but importantly, more nuanced assessments of children's own gender are needed for them to be reliable predictors of gender reasoning across ages.

A recent review on the development of gender identity describes how children's identity has been assessed in previous research (Fisher et al. 2024). One common method is asking children to make a categorical statement about their gender, which they can do by around 2 years of age. Other studies make use of behaviors and preferences as a proxy for gender identity, for example, asking adults or children themselves to rate preferred toys, activities, or peers to determine how gender-typed their preferences are and thus infer how much the child relates to boys or girls. Such measures show good consistency within individuals over time, and at a group level, a roughly similar trajectory with increasing gender-typing rigidity early in childhood until a peak sometime in the preschool years, followed by increasing flexibility (with the exception of peer preferences which remain rigid in their gender segregation throughout childhood) (Fisher et al. 2024). Further, when children's gender identity is assessed broadly using a composite score including categorical gender identity (boy, girl, or something else), more continuous similarity to boys and girls in different domains, peer preferences, toy preferences, and clothing preferences, children show moderate stability in their responses from 7 to 9 years of age, though also decreasing rigidity in their gender typing in this age span (Hässler et al. 2022).

Together, this research on gender identity development suggests that while there are developmental changes in the rigidity of children's gender identity and gender-typed behaviors and preferences, there is also meaningful variation between children in how rigidly they identify their gender which could potentially be an important predictor for how children think about others' gender as well. That is, those with very rigid views of their own gender may have rigid, stereotyped views of others' gender as well and vice versa. As evidence for this suggestion, when 4–8-year-old children's gender-typed preferences are rated by their parents, those whose preferences are rigidly in line with gender typing are less accepting of gender norm violations by other children (Riggs et al. 2023). Six- to 10-year-old children are also less accepting of gender norm violations when they have strong beliefs in gender essentialism (Gross et al. 2024). Girls from 5 to 13 years of age who describe themselves as tomboys (i.e., their gender identity and interests are more masculine than average for girls) have more inclusive stereotypes than girls who are not tomboys (Ahlqvist et al. 2013; Martin and Dinella 2012). Finally, 6–8-year-old transgender children and their siblings are more accepting of violations of gender norms than other children (Olson and Enright 2018). These findings suggest that children with less rigid gender identities themselves might have less gender prejudice against others, which further suggests that they may be sensitive to more nuance in others' preferences as well.

By examining individual differences, such as in gender rigidity, we can gain a greater understanding of the social and cognitive mechanisms behind children's gender reasoning as well as test theoretical predictions. A relevant theory for considering children's reasoning about gender and how it relates to their own identity is The Gender Self-Socialization Model (GSSM; Tobin et al. 2010), which is based on a broader model of social cognition (Greenwald et al. 2002) connecting the self, social categories, and attributes in a dynamic, developmental process. GSSM focuses this broader model on gender development and proposes that children's ideas about gender are inherently connected to their view of their own gender, such that, for example, gender identity, together with attribute self-perceptions influence gender stereotypes (Tobin et al. 2010). For example, if a child identifies as a girl and has a preference for a certain toy, she may form a gender stereotype that the toy is "for" girls. Previous research has shown that children do indeed project their own preferences onto others of the same gender at 4–5 years of age (Martin et al. 1995) and prefer toys that they categorize as being for their own gender at 3–5 years of age (Cherney and Dempsey 2010). A further prediction would be that flexibility or variation beyond the binary in a child's gender identity or how strongly they embody gender-related attributes and preferences will affect the strength of their gender stereotypes for others. Most previous research, however, has focused on gender as a binary category and thus has not been sensitive to more nuanced variations in identities, attributes, or stereotypes. By exploring gender identity, gender typing, and gender stereotypes as more continuous concepts, and by examining the impact of individual differences between children in factors such as their gender identity rigidity, it is possible to examine whether this variation does indeed spread across the GSSM. Specifically, a further prediction one could make from the model is that children with less rigid gender identities themselves will have a more flexible and nuanced view of others' gender and how closely gender is tied to traits and preferences.

In the current study, we asked 3–6-year-old children to predict the toy preferences of unfamiliar children with typical male or female appearance/name/pronouns but varying in home upbringing environment as indicated by photos of their bedrooms and toys they were given to play with at home. Previous research has on the one hand indicated that 3–6-year-old children's gender stereotypes are more idealized and extreme than those of 7–10-year-olds (Foster-Hanson and Rhodes 2023), and on the other hand that 5- and 6-year-olds' conceptualization and reasoning about nonbinary unfamiliar others is unclear (Yuly-Youngblood and Boseovski 2022). Here, we, therefore, presented unfamiliar target children as gender-typical boys and girls but included gender-neutral upbringing environments alongside gender-typical and counter-gender-typical environments. This allowed for a narrower focus on children's thinking about how unfamiliar others' preferences relate to their varied upbringing environments and their binary gender. Children's pattern of responses when allocating toys can give insight into whether children rely more on individuals' binary gender identity or more on their social environment when determining the source of gender-typed preferences. Allocations of toys are a common way

to assess preschool-aged children's ideas about others' gender typing (e.g., Hässler et al. 2022; King et al. 2020; Shutts et al. 2017; Wang et al. 2023) as toys are concrete and relatable for them, though often only binary feminine and masculine options are available to select. In the current study, we not only offer a range of gender typing in the presented toys, as in Wang et al. (2023), we also allow children to select several toys that the child might like as well as several toys that they think the child would not like, thus giving a more detailed picture of their intuitions about the child's preference with less trial to trial random variation.

To explore individual differences in these prediction tendencies, we asked children themselves to identify their gender on two continuous scales (identification with girls and identification with boys); asked them about their own preferred and nonpreferred toys; and asked their preschool teachers for a rating of how gender-typed their behavior and preferences are. These individual difference assessments allow us to test whether children's own gender identity and preference rigidity relate to their expectations for others. A similar gender identity task which allowed children to identify along one bipolar boy-to-girl spectrum, indicated initial validity and test–retest reliability in children aged 3–14 years (Gülgöz et al. 2022). Here, we chose to examine preschool-age children given that at this age, children's knowledge of and rigidity in gender stereotypes is increasing (Cherney and Dempsey 2010; Serbin et al. 1993), making it a formative age for ideas about gender.

We hypothesize that when given the opportunity for more nuanced responses, children will be able to incorporate both identity and environment as information sources when predicting others' toy preferences. Further, and in line with GSSM, children with less rigid views of their own gender identity and less rigidly gender-typed own preferences will be more sensitive to potential variation in others' preferences, giving more nuanced, less binary responses when allocating toys.

2 | Method

2.1 | Participants and Recruitment

Fifty-six children (26 girls and 30 boys) between 3 years, 3 months and 6 years, and 5 months (M_{age} = 5 years, 1 month, SD_{age} = 8.75 months) were included in the analysis. Of these, three were 3-year-olds, 21 were 4-year-olds, 25 were 5-year-olds, and 7 were 6-year-olds. An a priori power analysis indicated that this sample size should be sufficient to detect medium to large effects with a power of 0.8 and 5 predictors in a linear regression model (details on the power analysis are available on the Open Science Framework, OSF, https://osf.io/qu4yz/). One additional child began the study but did not want to participate in the tasks, so they were excluded. Children were recruited from five different preschools in a medium-sized city in Sweden. The study was reviewed and approved by the local ethical review authority (protocol number 2021-06099-02). Information and consent forms were left at the preschool for parents of 3–6-year-olds. Children whose parents had given informed consent were asked if they would like to participate and those who verbally assented were brought to a quiet area in the preschool to participate in the tasks. Participating children were primarily from White, educated families.

2.2 | Materials and Procedure

2.2.1 | Questionnaire for Preschool Teachers

Preschool teachers of children who completed the tasks were asked to fill out a brief questionnaire about each child. Preschool teachers rated children's clothing, play behavior (as typically feminine, masculine, and gender neutral), play partner preferences (plays with boys, plays with girls), and popularity (among boys, among girls, among all children) each on a 5-point scale (see supplementary material for full list of items). Teachers also provided the child's gender as boy, girl, or nonbinary and their age.

2.2.2 | Toy Allocation Task

Children were shown 12 laminated photographs of toys (see Figure 1) that had previously been rated by adults as typically feminine (a baby doll, a fashion doll, a pink horse, and a tea set), typically masculine (a tool belt, a construction crane, a motorcycle, and a toy rifle[1]), and gender-neutral (a pizza set, a puzzle, a board game, and a set of colorful character toys popular in Sweden: "Babblarna"). The photographs were presented in random order in a 3 by 4 array. Then children were asked to choose three toys that they would like to play with and three toys that they would not like to play with as an assessment of their own gender-typed toy preferences.

Next, they were introduced to six characters, one at a time on a laptop computer screen in one of six different presentation orders. For each character, an avatar-style face in black and white was shown alongside a photograph of a child's room and two toys that were either typically masculine, typically feminine, or gender-neutral (as rated by a group of adults prior to the study; see Figure 2 for an example trial). The six characters (three girls and three boys) allowed for each possible gender and room combination to be presented (e.g., girl in a masculine room, girl in a neutral room, and so on). The experimenter presented each character by saying, for example, "This is Maryam and this is her room and her toys. Her parents gave her the toys and decorated her room like this. Maryam has never been to preschool before but she is starting now. Which of these toys do you think she will choose to play with when she is at preschool?" After the child made three choices, the experimenter asked, "Which of these toys do you think she *won't* want to play with at preschool?" and the child was able to choose three nonpreferred toys. The full stimuli set can be viewed on OSF (https://osf.io/qu4yz/).

2.2.3 | Gender Identity Task

Children rated their gender identity using two continuous scales, one for identification with boys and one for identification with girls. To introduce and ensure that the participants understood the task, each test began with establishing the paradigm using different types of birds (see supplementary materials on OSF for

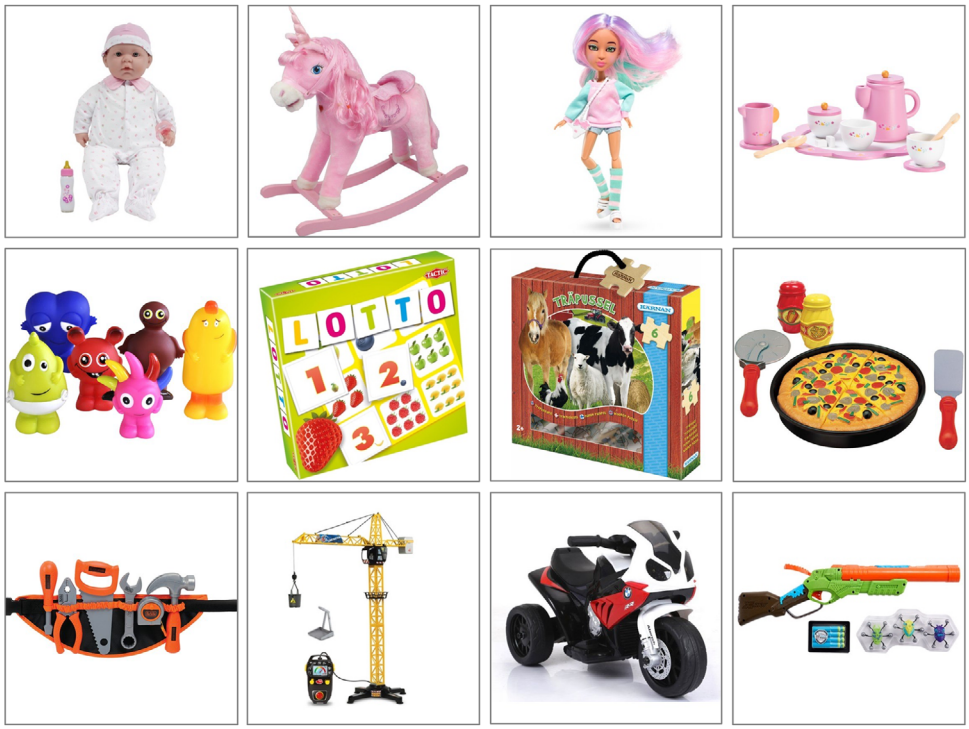

FIGURE 1 | Toys for the allocation tasks and own toy preferences. *Note:* There were four toys per category: feminine (top row), neutral (middle row), and masculine (bottom row), which were arranged randomly when presented to participating children.

FIGURE 2 | Example trial for a girl character with a masculine room.

a detailed description of the training, https://osf.io/qu4yz/). A yellow laminated circle with a diameter of 12 cm was placed on the table in front of the child and the experimenter explained "imagine that this circle represents all birds." Adjacent to the circle was a 50-cm long scale marked subtly with numbers ranging from 0 to 10, and with the yellow circle placed at 10. The experimenter introduced three different birds (crow, penguin, and flamingo) one at a time, describing them and placing them on the scale to show how they identify (e.g., the crow identifies strongly as a bird, while the penguin does not). The experimenter emphasized that there are no right or wrong answers, it is how one feels that is important. As a test of the participants' understanding of the paradigm, the participants were given pictures of an owl and a pigeon and asked to reflect on how these two birds might identify. All participants were able to reason hypothetically from the birds' point-of-view and motivate their individual placing of the two pictures along the scale. The experimenter then removed all the bird pictures and said that

TABLE 1 | Spearman's rho correlations between main study variables for all participants.

Variable	1	2	3	4	5	6	7	8
1. Age in months	—							
2. Own gender-typed toy preference	0.10	—						
3. Allocation by gender identity	0.38**	0.36**	—					
4. Allocation by environment	0.05	0.08	−0.11	—				
5. Identification with girls	−0.15	−0.09	−0.03	−0.09	—			
6. Identification with boys	−0.01	0.25	−0.03	−0.10	−0.05	—		
7. Teacher-rated gender typing	0.08	−0.03	−0.13	−0.09	0.24	0.13	—	
8. Rigidity	0.36**	−0.05	0.45**	−0.19	−0.16	−0.11	0.29*	—

*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

the circle now represents "all children," and presented the small green circle saying "and this is you. How would you place yourself if the circle is all other children?" This question was used to familiarize children with placing themselves on the scale and to deemphasize binary gender as being the target of the task. The experimenter repeated the task for the two questions of interest: "all boys" and "all girls." The child's response for each gender was coded from 0 to 10, depending on where they placed the green circle, with 0 representing low identification and 10 representing high identification with the corresponding group.

2.3 | Data Coding and Analysis

Teacher ratings were examined both as individual items and as collapsed into an overall gender typing score in which the total of the other gender items was subtracted from the total of the own gender items (see Results).

Identification with girls and with boys was scored from 0 to 10, based on children's placement of the circle on the scale. Six children declined to respond to the task, but their data were included in the analyses where possible.

Children's own gender-typed toy preferences were scored with 1 point per feminine toy choice, 0 for a gender-neutral toy choice, and −1 for a masculine toy choice. The average of their nonpreferred toys was subtracted from the average for their preferred toys, resulting in a score between 2 (most feminine preference) and −2 (most masculine preference), then scores were reversed for boys, such that 2 indicated a high own-gender preference for both boy and girl participants.

For toy allocations to the characters, the same initial score calculation was used, resulting in a feminine (2) to masculine (−2) toy allocation score. However, since it is difficult to identify the allocation of neutral toys—a key aspect of this study—with such a score, we also separately examined the number of neutral toys allocated to each character as a preferred toy. In addition, in order to more easily examine individual differences in children's patterns of toy allocations, we calculated a score for allocation based on character binary gender by subtracting allocation scores for boy characters from allocation scores for girl characters (identity-based allocation score); and a score for allocation by room type by subtracting allocations to characters with masculine rooms from those to characters with feminine rooms (environment-based allocation score). All children participated in their own toy preferences task, though five children did not select six toys and their score was based on the toys they did select. When allocating toys to others, 39/56 children completed all six trials with six toy allocations, and all children completed at least two trials. If a child made no allocations for a given trial (e.g., they chose to end participation before reaching that trial), it was excluded from the analyses (28 trials, 8.3%), if at least two toy allocations were made, the trial was included with the partial data used to calculate the allocation score (19 trials, 5.7%).

Analyses were carried out in Jamovi (Lenth 2020; R Core Team 2021; Revelle 2019; Singmann 2018; The jamovi project 2022). The full data file and detailed output for each analysis are available on OSF (https://osf.io/qu4yz/).

3 | Results

3.1 | Preliminary Analyses and Descriptive Statistics

Spearman's rho (ρ) correlations between the study variables are presented in Table 1 for all children and in Table 2 for girls and boys separately. Correlation tables that also include individual trial types (e.g., boys in masculine rooms) can be found in the supplementary results. Correlations between age and other variables were examined for the full sample to determine if age should be included as a factor in later analyses. Age was found to correlate positively with children's gender identity rigidity ($\rho = 0.36$, $p = 0.010$), identity-based allocation of toys ($\rho = 0.38$, $p = 0.004$), and to correlate negatively with allocation of feminine toys to boys in masculine rooms ($\rho = -0.35$, $p = 0.010$). Age did not correlate with allocations in any of the other trials, with environment-based allocations, own toy preferences, teacher-rated gender typing score, or identification with either girls or boys (all p's > 0.05; see supplementary materials for detailed results).

Independent samples tests (either t-tests or Mann–Whitney U tests, depending on whether the assumption of normality was violated) to examine binary gender differences indicated that boys and girls differed on teacher-rated femininity ($t(54) = -10.83$,

TABLE 2 | Spearman's rho correlations between main study variables for girls and boys separately.

Variable	1	2	3	4	5	6	7	8
1. Age in months	—	0.11	0.47**	0.07	−0.33	0.14	0.10	0.43*
2. Own gender-typed toy preference	0.06	—	0.46*	0.10	−0.09	0.08	−0.11	0.08
3. Allocation by gender identity	0.20	0.29	—	−0.06	−0.26	0.18	−0.20	0.47*
4. Allocation by environment	0.06	0.08	−0.28	—	−0.13	−0.34	−0.21	−0.21
5. Identification with girls	0.05	−0.12	0.13	−0.12	—	0.12	0.14	−0.61***
6. Identification with boys	−0.19	0.08	−0.02	0.21	0.11	—	0.38	0.34
7. Teacher-rated gender typing	0.02	0.05	−0.06	0.06	0.53**	−0.21	—	0.11
8. Rigidity	0.29	−0.19	0.34	−0.17	0.44*	−0.62**	0.54**	—

Note: Correlations for girls are displayed below the diagonal and those for boys are displayed above the diagonal.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

$p < 0.001$), identification with girls ($U = 186.0$, $p = 0.015$), and identification with boys ($U = 103.0$, $p < 0.001$). They did not differ on gender identity rigidity, own gender-typed toy preferences, age, or any of the toy allocation variables (all p's > 0.05). Within and across binary genders, individual toys within the same category (masculine, feminine, gender neutral) were not preferred or allocated significantly differently from other toys within the same category.

Analyses indicate that there were no significant differences across preschools in the study's main variables, including teacher ratings of children, children's gender rigidity, and children's toy allocation by identity or by environment (all p's > 0.05; see supplementary materials for detailed results).

Based on the zero-order correlations between teacher rating items, an exploratory factor analysis was used to determine if the items could be aggregated into one factor. A maximum likelihood analysis including all teacher ratings was performed with direct oblimin rotation and extraction based on minimum residuals. The assumption of sphericity was confirmed by Bartlett's test ($\chi^2(45) = 227.51$, $p < 0.001$) and sampling adequacy was adequate overall (Kaiser–Meyer–Olkin = 0.74). The factor analysis indicated that teacher's ratings loaded onto one factor with good fit ($X^2(35) = 117.88$, $p < 0.001$), with the exception of the two gender-neutral items (i.e., "wears gender neutral clothes" and "is popular with children"; see supplementary materials for detailed results). The masculine items loaded negatively and the feminine items loaded positively. Given that we are interested in children's gender typing (i.e., girls' femininity and boys' masculinity, we calculated an overall *teacher-rated gender typing* score from teacher ratings in which the total score on the masculine items was subtracted from the total score on the feminine items for girls and the total score on the feminine items was subtracted from the total score on the masculine items for boys, and the neutral items were excluded.

3.2 | Children's Gender Identification

Children's gender identification ratings were examined using a 2 (scale gender) by 2 (child binary gender) ANOVA. A significant interaction ($F(1, 48) = 39.99$, $p < 0.001$, $\eta^2_p = 0.45$; see Figure 3) was followed up with Bonferroni-adjusted pairwise comparisons, revealing that boys identified more with boys than girls did ($t(48) = 5.11$, $p_{bonferroni} < 0.001$) and more than they identified with girls ($t(48) = 3.95$, $p_{bonferroni} = 0.002$). Girls identified with girls more than they identified with boys ($t(48) = 4.96$, $p_{bonferroni} < 0.001$) and more than boys did ($t(48) = 3.04$, $p_{bonferroni} = 0.023$). Further, boys' identification with boys was not significantly different from girls' identification with girls ($t(48) = 0.20$, $p_{bonferroni} = 1.000$), and boys' identification with girls was not significantly different from girls' identification with boys ($t(48) = 1.44$, $p_{bonferroni} = 0.942$). No other effects in the ANOVA were significant (see supplementary materials).

Of particular interest in the current study was the rigidity of children's gender identity, that is, the difference between their identification with boys and girls, in which a larger difference suggests a more rigid, binary view of their own gender. The preliminary correlations between rigidity and other variables (Table 1) suggested that it might be related to age ($\rho = 0.36$, $p = 0.010$) and teacher-rated gender typing ($\rho = 0.29$, $p = 0.039$). For a more complete analysis to determine whether children's gender identity rigidity can be predicted by other variables, or whether it captures an independent aspect of a child's gender, we used linear regression to analyze whether rigidity was associated with children's binary gender, age, gender-typed toy preferences, and teacher-rated gender typing. A model with these predictors was not overall significant ($F(4, 45) = 2.39$, $p = 0.064$, adjusted $R^2 = 0.10$); however, age was a significant predictor (Estimate = 0.11, $t = 2.03$, SE = 0.05, $p = 0.049$), suggesting a possible shift toward stronger identity rigidity with increasing age, though overall identity rigidity seems to be largely unrelated to other aspects of a child's gender.

3.3 | Children's Allocation of Toys to Characters at a Group Level

To examine how children allocated toys to the characters at a group level, we first analyzed the toy score (feminine minus masculine toys) using a 3 (environment: room gender-type feminine, masculine, or neutral) by 2 (identity: character gender identity girl or boy) ANOVA. There were main effects for both environment ($F(2, 86) = 9.84$, $p < 0.001$, $\eta^2_p = 0.19$) and identity

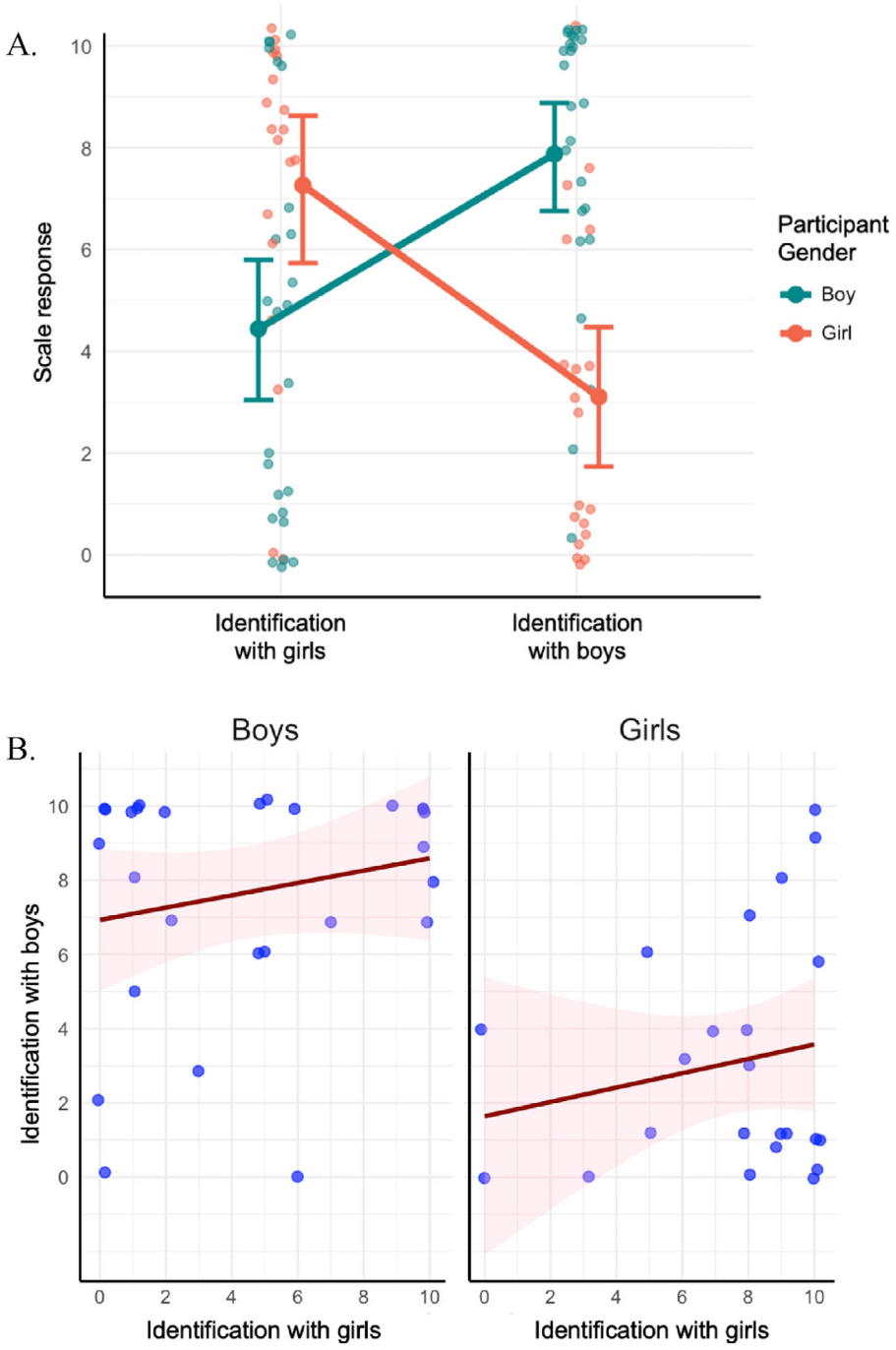

FIGURE 3 | Children's identification with girls and with boys by binary gender. *Note:* Panel (A) shows children's group-level responses on the two scales, including jittered individual data points. Panel (B) shows the relationship between individual boy and girl participants' responses on the two identification scales.

($F(1, 43) = 76.75$, $p < 0.001$, $\eta^2_p = 0.64$), as well as a significant interaction between the two factors ($F(2, 86) = 3.38$, $p = 0.039$, $\eta^2_p = 0.07$; see Figure 4). Post hoc tests with Bonferroni correction revealed that girls were allocated more feminine, less masculine toys than boys, regardless of the environment (feminine: $t(43) = 5.93$, $p_{tukey} < 0.001$; neutral: $t(43) = 9.52$, $p_{bonferroni} < 0.001$; masculine: $t(43) = 5.84$, $p_{bonferroni} < 0.001$). However, toy allocations to girls with masculine environments were not different from those for boys with feminine environments ($t(43) = 2.49$, $p_{bonferroni} = 0.250$). Within girl characters, children allocated similar toys to girls with feminine and neutral environments ($t(43) = 0.13$, $p_{bonferroni} = 1.000$), but less feminine toys to girls with masculine environments than those with either feminine ($t(43) = 3.22$, $p_{bonferroni} = 0.037$) or neutral ($t(43) = 3.33$, $p_{bonferroni} = 0.027$) environments. Within boy characters, toy allocations to boys with masculine and feminine environments differed ($t(43) = 3.18$, $p_{bonferroni} = 0.041$), but allocations to boys with neutral environments did not differ significantly from either of these (masculine: $t(43) = 0.64$, $p_{bonferroni} = 1.000$; feminine: $t(43) = 2.38$, $p_{bonferroni} = 0.33$).

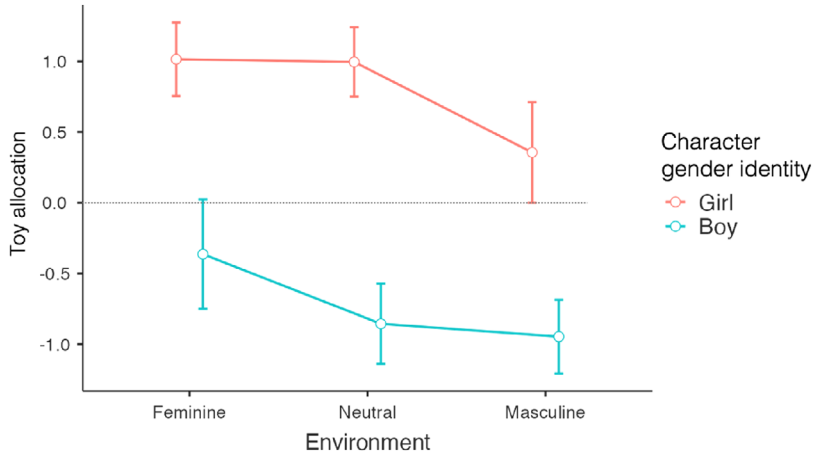

FIGURE 4 | Children's allocation of toys to girl and boy characters with feminine, neutral, and masculine environments. *Note:* Higher toy allocation scores indicate more feminine and less masculine toy allocations.

When examining the allocation of the neutral toys specifically, using the same type of analysis, there were main effects for character gender identity ($F(1, 50) = 6.22$, $p = 0.016$, $\eta^2_p = 0.11$), with boy characters being allocated more neutral toys than girl characters ($M_{boys} = 0.78$, $SE_{boys} = 0.07$; $M_{girls} = 0.56$, $SE_{girls} = 0.08$); and for the environment ($F(2, 100) = 3.98$, $p = 0.02$, $\eta^2_p = 0.07$), with more neutral toys allocated to characters with neutral than with feminine environments ($M_{neutral} = 0.80$, $SE_{neutral} = 0.09$; $M_{feminine} = 0.56$, $SE_{feminine} = 0.08$; $t(50) = 2.86$, $p_{bonferroni} = 0.018$), though not more than those in masculine environments ($M_{masculine} = 0.65$, $SE_{masculine} = 0.07$; $t(50) = 1.72$, $p_{bonferroni} = 0.276$).

3.4 | Individual Differences in Children's Allocation of Toys

We were further interested in how children's toy allocation to the characters might be related to aspects of their own gender. That is, does the rigidity of children's own gender identity, their own gender-typed toy preferences, their teacher-rated gender typing, their binary gender, or their age relate to their preferential reliance on the character's binary gender identity or their gender-typed environment when allocating toys. Thus, we used the composite score for allocation based on character gender identity and on environment and analyzed these using linear regression with the factors listed above.

For allocation of toys by character gender identity, the model was overall significant ($F(5, 44) = 4.37$, $p = 0.003$, adjusted $R^2 = 0.33$; see Table 3), and children's gender identity rigidity (Estimate = 0.15, $t = 3.30$, SE = 0.05, $p = 0.002$) and their own gender-typed preferences (Estimate = 0.41, $t = 2.35$, SE = 0.18, $p = 0.023$) were significant predictors. That is, children with greater rigidity in their own gender identity or in their own toy preferences were more likely to allocate toys based on the character's binary gender identity, with less reliance on their environment. For the allocation of toys by environment, the model ($F(5, 43) = 0.65$, $p = 0.662$, adjusted $R^2 = -0.04$) was not significant overall and neither were any of the predictors.

TABLE 3 | Model results for individual differences in toy allocation by character gender identity.

Predictor	Estimate	SE	*t*	*p*
Intercept	−0.81	1.06	−0.77	0.446
Gender (girl–boy)	−0.05	0.52	−0.09	0.931
Age in months	0.02	0.02	1.40	0.168
Identity rigidity	0.12	0.05	2.68	0.010*
Own gender-typed preference	0.40	0.18	2.20	0.033*
Teacher-rated femininity	0.15	0.18	0.83	0.410

4 | Discussion

In the current study, we aimed to examine whether preschool-aged children could give nuanced predictions of others' gender-typed preferences, as well as whether there are individual differences in children's nuanced predictions based on aspects of their own gender. The results revealed that children do predict others' gender-typed preferences using a combination of information sources: binary identity and the environment in which a child was raised. Overall, more feminine toys were allocated to girls and more masculine toys to boys, but the environment did matter, such that children with a cross-gender room were allocated less gender-typed toys than children in a gender-matched room. Further, there was no significant difference in the toys allocated to girls in masculine rooms and boys in feminine rooms. Together, this suggests that gender identity seems to weigh more heavily in predicting others' toy preferences than environment, but that preschoolers do also show sensitivity to the potential influence of the gender typing of the environment a child grew up in. For neutral toys specifically, which are often not included as options in research on children's toy preferences, we found that these items were allocated more often to children with a gender-neutral room than a feminine room and more often to boys than to girls. Despite gathering ratings in advance from adults to ensure that the gender-neutral toys were not seen as feminine or masculine, this suggests that to children, the neutral toys may have been perceived as somewhat masculine, or that neutral toys

are seen as more appropriate for boys than for girls. This finding for neutral toys is in line with previous research showing that when 3–5-year-old children were presented with toys that adults considered gender-neutral, they were more likely to judge them to be masculine than feminine (Cherney and Dempsey 2010).

We were further interested in individual differences in how children allocated toys to the characters and specifically, their reliance on characters' binary gender identity or the character's home environment. We found no individual difference predictors for children's use of the environment information. However, relying primarily on character's binary identity to predict their preferences was associated with a more rigid own gender identity and more gender-typed own preferences, though not associated with participant age, binary gender, or teacher perceptions of a child's masculine/feminine behavior and gender expression. This suggests that children's perceptions of their own gender identity relate to their thinking about others' gender as well.

The findings that variation in the rigidity of children's own gender identity and gender-typed preferences are related to how rigidly they expect other children's preferences to align with their binary gender aligns with our expanded predictions from the GSSM (Tobin et al. 2010). That is, the relationship between self and gender proposed in the original model presents gender identity as a binary choice, yet there is important variation beyond the binary. How rigid or flexible a child's identity and preferences are appear to predict how rigid or flexible their gender-typed expectations for others' preferences are as well. This suggests that children's gender schemas are not based on simple, binary categories, but rather show more complex and nuanced aspects that vary across individuals.

It is interesting to note that we found no group-level gender differences in the responses of girls and boys in the study for toy allocations, gender identity, or gender-typed preference rigidity. Previous research has indeed been mixed on binary gender differences in reasoning about gender in the preschool age range and our findings support that one's own gender identity as a girl or boy is not a predictor of reasoning about gender, but rather the more nuanced aspects of how rigidly one identifies with gender that matters. The results also indicated very similar patterns of allocation for girl and boy characters. Some research has suggested that boys are often subject to stricter gender norms than girls (4–9 years: Kwan et al. 2020; 4–8 years: Riggs et al. 2023) and this could have resulted in more identity-based allocations to boys than girls overall. However, it could be that such differences appear, not when children are asked to predict others' preferences, but rather when they are asked to evaluate how acceptable those choices are.

In our gender identity task, we found that children can describe their gender identities with more than just binary categories, and that identities may become more rigidly binary with age. Other factors such as child binary gender, teacher perception of child gender-typed behavior and preferences were not related to identity rigidity in our sample. This suggests that gender identity rigidity may increase with age, yet be independent of other aspects of a child's gender, such as their gender-typed behavior or preferences. Recent research has been making more use of such nonbinary assessments of children's felt gender identity (3–14 years: Gülgöz et al. 2022) or their sense of similarity to boys and to girls (5–9 years: Martin, Andrews et al. 2017). Our results confirm that this type of measure is meaningful in the preschool age range and that such self-report ratings relate to other aspects of how children conceptualize their own and others' gender.

A strength of the current study is that we collected information on children's gender in different domains and using different methods. That is, we assessed both binary and continuous gender identity, obtained ratings from preschool teachers regarding children's gender-typed behaviors and expression, and asked children to rate their preferences for gender-typed and neutral toys. However, it is possible that adding more redundant assessments could have given an even more robust assessment (e.g., researchers rating the child's behavior in a classroom observation in addition to using the teacher reports and child explicit preference ratings) as previous research has shown that children sometimes give more gender-typed responses when being asked explicitly for their preferences compared to natural observation data (4–6 years: Dinella et al. 2017). Other important extensions for future research would be to allow children to rate others' gender identities using the continuous scales and to present children with less binary target characters as well. That is, in the current study, we contrasted binary gender identity and more continuous gender-typed environments; in the future, it would be interesting to know more about how nonbinary identities might interact with information about gender typing in the environment and to be able to assess their gender typing through multiple measures.

The current study was conducted in Sweden, a country that is known for relatively high gender equality (Global Gender Gap Report 2023 2023) and in which ensuring equal treatment of children regardless of gender is part of the preschool curriculum (Läroplan för förskolan. Lpfö 18, 2018). Thus, it could be that in countries with less gender equality, children might be less sensitive to gender variation and instead tend toward more rigid views of gender categories. Previous research suggests that broader cultural values can influence children's reasoning about gender. For example, Canadian and US American children are more accepting of gender nonconforming behavior than Chinese (4–9 years; Kwan et al. 2020) and Korean (5–9 years: Conry-Murray et al. 2015) children. It could be informative to examine whether children in cultures with less progressive values around gender show similar responses as those in the current study.

Another potential direction for future research is to expand the age range of participating children. We selected 3–6-year-olds as it is a time in which ideas about gender are becoming more elaborate and more rigid, though it could be interesting to see whether children who are past the most rigid phase of gender stereotyping show similar patterns of responses when it comes to incorporating multiple sources of information for gender typing and for the relation between the rigidity of their own identity and the rigidity of their expectations for others. Interestingly, a recent study indicated that a brief intervention aimed at decreasing how rigidly 6–10-year-old children view gender categories was not successful in decreasing prejudice against children who defy gender norms (Gross et al. 2024). Yet it is possible that younger children who are in an earlier stage of forming their ideas about gender would be more impacted by such an intervention. Finally, replicating the results with larger samples (higher powered

studies) would be beneficial for ensuring that the results are reliable and generalizable.

Together the results of the current study show that children from 3 to 6 years of age show nuance in their understanding of gender as more than simple binary categories. That is, they predict variation in others' gender-typed preferences by weighing multiple sources of information. Further, the rigidity of children's own gender identity and their gender-typed preferences is associated with how rigidly gender-typed they expect others' preferences to be. This is in line with previous research suggesting that children with more flexible gender identities are more accepting of others' gender norm violations (Ahlqvist et al. 2013; Martin and Dinella 2012; Olson and Enright 2018) and supports our proposed extension of the GSSM (Tobin et al. 2010) to predict that variation beyond the binary in any one aspect of gender (identity, stereotypes, preferences) will lead to flexibility in other aspects as well. The findings have potential implications not only methodologically for how researchers can present and assess gender beyond the binary, but also from a wider perspective for how gender can be discussed and represented by parents or preschool educators given that preschool-aged children do appear to be able to comprehend gender as more than a discrete, binary category. It is critical that research on children's reasoning about gender and identity continues to offer more nuanced response options if we are to be able to gain a full picture of the complexity of their gender concepts.

Ethics Statement

The research conducted ethically in line with The Declaration of Helsinki and was reviewed and approved by the local ethical review board (protocol number 2021-06099-02).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Anonymized data are available on the Open Science Framework (https://osf.io/qu4yz/).

Endnotes

[1]Note that in Sweden, firearms do not carry the same negative connotation as in, for example, the United States, and which might cause children to select the rifle at different rates than the other toys. In fact, the rifle was chosen equally as often as the other typically masculine toys in our sample.

References

Ahlqvist, S., M. L. Halim, F. K. Greulich, L. E. Lurye, and D. Ruble. 2013. "The Potential Benefits and Risks of Identifying as a Tomboy: A Social Identity Perspective." *Self and Identity* 12, no. 5: 563–581. https://doi.org/10.1080/15298868.2012.717709.

Banerjee, R., and V. Lintern. 2000. "Boys Will Be Boys: The Effect of Social Evaluation Concerns on Gender-Typing." *Social Development* 9, no. 3: 397–408. https://doi.org/10.1111/1467-9507.00133.

Boseovski, J. J., C. Hughes, and S. E. Miller. 2016. "Expertise in Unexpected Places: Children's Acceptance of Information From Gender Counter-Stereotypical Experts." *Journal of Experimental Child Psychology* 141: 161–176. https://doi.org/10.1016/j.jecp.2015.09.002.

Cherney, I. D., and J. Dempsey. 2010. "Young Children's Classification, Stereotyping and Play Behaviour for Gender Neutral and Ambiguous Toys." *Educational Psychology* 30, no. 6: 651–669. https://doi.org/10.1080/01443410.2010.498416.

Conry-Murray, C., J. M. Kim, and E. Turiel. 2015. "Judgments of Gender Norm Violations in Children From the United States and Korea." *Cognitive Development* 35: 122–136. https://doi.org/10.1016/j.cogdev.2015.04.002.

Dinella, L. M., E. S. Weisgram, and M. Fulcher. 2017. "Children's Gender-Typed Toy Interests: Does Propulsion Matter?" *Archives of Sexual Behavior* 46, no. 5: 1295–1305. https://doi.org/10.1007/s10508-016-0901-5.

Ebert, W. M., L. Jost, and P. Jansen. 2024. "Gender Stereotypes in Preschoolers' Mental Rotation." *Frontiers in Psychology* 15: 1284314. https://doi.org/10.3389/fpsyg.2024.1284314.

Fisher, E., S. Wright, and C. Sargeant. 2024. "What Do Measures of Gender Identity Tell Us About Gender Identity Over Time?" *British Journal of Developmental Psychology* 1–23. https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1111/bjdp.12491.

Foster-Hanson, E., and M. Rhodes. 2023. "Stereotypes as Prototypes in Children's Gender Concepts." *Developmental Science* 26, no. 4: e13345. https://doi.org/10.1111/desc.13345.

Gelman, S. A., P. Collman, and E. E. Maccoby. 1986. "Inferring Properties From Categories Versus Inferring Categories From Properties: The Case of Gender." *Child Development* 57, no. 2: 396–404. https://doi.org/10.2307/1130595.

Global Gender Gap Report 2023. 2023. World Economic Forum. http://reports.weforum.org/globalgender-gap-report-2023.

Greenwald, A. G., M. R. Banaji, L. A. Rudman, S. D. Farnham, B. A. Nosek, and D. S. Mellott. 2002. "A Unified Theory of Implicit Attitudes, Stereotypes, Self-Esteem, and Self-Concept." *Psychological Review* 109, no. 1: 3–25. https://doi.org/10.1037/0033-295X.109.1.3.

Gross, E. B., R. D. Fine, S. Gülgöz, K. R. Olson, and S. A. Gelman. 2024. "Children's Gender Essentialism and Prejudice: Testing Causal Links via an Experimental Manipulation." *Developmental Science* 27, no. 5: e13532. https://doi.org/10.1111/desc.13532.

Gülgöz, S., D. L. Edwards, and K. R. Olson. 2022. "Between a Boy and a Girl: Measuring Gender Identity on a Continuum." *Social Development* 31, no. 3: 916–929. https://doi.org/10.1111/sode.12587.

Halim, M. L., D. N. Ruble, C. S. Tamis-LeMonda, K. M. Zosuls, L. E. Lurye, and F. K. Greulich. 2014. "Pink Frilly Dresses and the Avoidance of all Things "Girly": Children's Appearance Rigidity and Cognitive Theories of Gender Development." *Developmental Psychology* 50, no. 4: 1091–1101. https://doi.org/10.1037/a0034906.

Hässler, T., J. J. Glazier, and K. R. Olson. 2022. "Consistency of Gender Identity and Preferences Across Time: An Exploration Among Cisgender and Transgender Children." *Developmental Psychology* 58, no. 11: 2184.

Hyde, J. S., R. S. Bigler, D. Joel, C. C. Tate, and S. M. van Anders. 2019. "The Future of Sex and Gender in Psychology: Five Challenges to the Gender Binary." *American Psychologist* 74, no. 2: 171–193. https://doi.org/10.1037/amp0000307.

King, R. A., K. E. Scott, M. P. Renno, and K. Shutts. 2020. "Counter-stereotyping Can Change Children's Thinking About Boys' and Girls' Toy Preferences." *Journal of Experimental Child Psychology* 191: 1–19. https://doi.org/10.1016/j.jecp.2019.104753.

Kwan, K. M. W., S. Y. Shi, A. N. Nabbijohn, L. N. MacMullin, D. P. VanderLaan, and W. I. Wong. 2020. "Children's Appraisals of Gender Non-conformity: Developmental Pattern and Intervention." *Child Development* 91, no. 4: e780–e798. https://doi.org/10.1111/cdev.13316.

Lam, V. L. 2023. "Gender-Based Reasoning About Novel Toys: The Role of Child and Parental Factors." *Infant and Child Development* 32, no. 4: e2423. https://doi.org/10.1002/icd.2423.

Läroplan för förskolan. Lpfö 18 (with Skolverket). 2018. Norstedts Juridik AB.

Lenth, R. 2020. "emmeans: Estimated Marginal Means, aka Least-Squares Means [R Package]." https://cran.r-project.org/package=emmeans.

Martin, C. L. 1989. "Children's Use of Gender-Related Information in Making Social Judgments." *Developmental Psychology* 25, no. 1: 80–88. https://doi.org/10.1037/0012-1649.25.1.80.

Martin, C. L., N. C. Z. Andrews, D. E. England, K. Zosuls, and D. N. Ruble. 2017. "A Dual Identity Approach for Conceptualizing and Measuring Children's Gender Identity." *Child Development* 88, no. 1: 167–182. https://doi.org/10.1111/cdev.12568.

Martin, C. L., R. E. Cook, and N. C. Z. Andrews. 2017. "Reviving Androgyny: A Modern Day Perspective on Flexibility of Gender Identity and Behavior." *Sex Roles* 76, no. 9–10: 592–603. https://doi.org/10.1007/s11199-016-0602-5.

Martin, C. L., and L. M. Dinella. 2012. "Congruence Between Gender Stereotypes and Activity Preference in Self-Identified Tomboys and Non-Tomboys." *Archives of Sexual Behavior* 41, no. 3: 599–610. https://doi.org/10.1007/s10508-011-9786-5.

Martin, C. L., L. Eisenbud, and H. Rose. 1995. "Children's Gender-Based Reasoning About Toys." *Child Development* 66, no. 5: 1453–1471. https://doi.org/10.2307/1131657.

Martin, C. L., and D. N. Ruble. 2009. "Patterns of Gender Development." *Annual Review of Psychology* 61: 353–381. https://doi.org/10.1146/annurev.psych.093008.100511.

Miller, C. F., L. E. Lurye, K. M. Zosuls, and D. N. Ruble. 2009. "Accessibility of Gender Stereotype Domains: Developmental and Gender Differences in Children." *Sex Roles* 60, no. 11: 870–881. https://doi.org/10.1007/s11199-009-9584-x.

Olson, K. R., and E. A. Enright. 2018. "Do Transgender Children (gender) Stereotype Less Than Their Peers and Siblings?" *Developmental Science* 21, no. 4: e12606. https://doi.org/10.1111/desc.12606.

Pillow, B. H., R. M. Pearson, and C. Allen. 2015. "Young Children's Inductive Generalizations About Social Categories: When Is Gender Essential?" *Merrill-Palmer Quarterly* 61, no. 4: 441–467.

Poulin-Dubois, D., L. A. Serbin, J. A. Eichstedt, M. G. Sen, and C. F. Beissel. 2002. "Men Don't Put on Make-Up: Toddlers' Knowledge of the Gender Stereotyping of Household Activities." *Social Development* 11, no. 2: 166–181. https://doi.org/10.1111/1467-9507.00193.

R Core Team. 2021. "R: A Language and Environment for Statistical Computing (Version 4.1) [Computer Software]." https://cran.r-project.org.

Revelle, W. 2019. "psych: Procedures for Psychological, Psychometric, and Personality Research [R Package]." https://cran.r-project.org/package=psych.

Riggs, A. E., D. Kinard, and M. Long. 2023. "Children's Evaluations of Gender Non-Conforming Peers." *Sex Roles* 88, no. 1: 17–34. https://doi.org/10.1007/s11199-022-01322-0.

Rizzo, M. T., and M. Killen. 2018. "Theory of Mind Is Related to Children's Resource Allocations in Gender Stereotypic Contexts." *Developmental Psychology* 54, no. 3: 510.

Ruble, D. N., L. J. Taylor, L. Cyphers, F. K. Greulich, L. E. Lurye, and P. E. Shrout. 2007. "The Role of Gender Constancy in Early Gender Development." *Child Development* 78, no. 4: 1121–1136. https://doi.org/10.1111/j.1467-8624.2007.01056.x.

Serbin, L. A., K. K. Powlishta, J. Gulko, C. L. Martin, and M. E. Lockheed. 1993. "The Development of Sex Typing in Middle Childhood." *Monographs of the Society for Research in Child Development* 58, no. 2: 1–95. https://doi.org/10.2307/1166118.

Shutts, K., B. Kenward, H. Falk, A. Ivegran, and C. Fawcett. 2017. "Early Preschool Environments and Gender: Effects of Gender Pedagogy in Sweden." *Journal of Experimental Child Psychology* 162: 1–17. https://doi.org/10.1016/j.jecp.2017.04.014.

Singmann, H. 2018. "afex: Analysis of Factorial Experiments [R Package]." https://cran.r-project.org/package=afex.

Skočajić, M. M., J. G. Radosavljević, M. G. Okičić, I. O. Janković, and I. L. Žeželj. 2020. "Boys Just Don't! Gender Stereotyping and Sanctioning of Counter-Stereotypical Behavior in Preschoolers." *Sex Roles* 82, no. 3: 163–172. https://doi.org/10.1007/s11199-019-01051-x.

Spinner, L., L. Cameron, and R. Calogero. 2018. "Peer Toy Play as a Gateway to Children's Gender Flexibility: The Effect of (Counter)Stereotypic Portrayals of Peers in Children's Magazines." *Sex Roles* 79, no. 5–6: 314–328. https://doi.org/10.1007/s11199-017-0883-3.

Taylor, M. G. 1996. "The Development of Children's Beliefs About Social and Biological Aspects of Gender Differences." *Child Development* 67, no. 4: 1555–1571. https://doi.org/10.1111/j.1467-8624.1996.tb01814.x.

Taylor, M. G., M. Rhodes, and S. A. Gelman. 2009. "Boys Will Be Boys; Cows Will be Cows: Children's Essentialist Reasoning About Gender Categories and Animal Species." *Child Development* 80, no. 2: 461–481. https://doi.org/10.1111/j.1467-8624.2009.01272.x.

The jamovi project. 2022. "Jamovi (Version 2.3) [Computer Software]." https://www.jamovi.org.

Tobin, D. D., M. Menon, M. Menon, B. C. Spatta, E. V. E. Hodges, and D. G. Perry. 2010. "The Intrapsychics of Gender: A Model of Self-Socialization." *Psychological Review* 117, no. 2: 601–622. https://doi.org/10.1037/a0018936.

Trautner, H. M., D. N. Ruble, L. Cyphers, B. Kirsten, R. Behrendt, and P. Hartmann. 2005. "Rigidity and Flexibility of Gender Stereotypes in Childhood: Developmental or Differential?" *Infant and Child Development* 14, no. 4: 365–381. https://doi.org/10.1002/(ISSN)1522-7219.

Wang, M. Z., V. Ng, and T. R. Gleason. 2023. "Toy Stories: Children's Use of Gender Stereotypes in Making Social Judgments." *Acta Psychologica* 235: 103879. https://doi.org/10.1016/j.actpsy.2023.103879.

Yuly-Youngblood, A. C., and J. J. Boseovski. 2022. "Children's Inductive Inferences About Individuals With Gender Category Uncertainty." *Social Development* 31, no. 4: 1255–1279. https://doi.org/10.1111/sode.12609.