Received: 5 February 2018 | Revised: 25 September 2018 | Accepted: 2 October 2018

DOI: 10.1111/sode.12342

WILEY

ORIGINAL ARTICLE

Children's reasoning about unequal gender-based distributions

Clare Conry-Murray

Department of Psychology, Saint Joseph's University, Philadelphia, Pennsylvania

Correspondence
Clare Conry-Murray, Department of Psychology, Saint Joseph's University, 226 Post Hall, 5600 City Avenue, Philadelphia, PA 19131-1395.
Email: cconrymu@sju.edu

Abstract

Children ages 7 (N = 56, M_{age} = 7.24, SD = 0.83), 9 (N = 55, M_{age} = 9.25, SD = 0.52), and 11 (N = 52, M_{age} = 11.60, SD = 0.79), and emerging adults (N = 50, M_{age} = 20.76, SD = 0.87) judged distributions of different items to boys and girls, when the items distributed varied by gender valence (related or unrelated to gender norms) and equivalency (equivalent or unequal). Distributions were judged to be acceptable most when the items were consistent with gender norms, especially for participants at ages 7 and 9, indicating that in middle childhood, children judge unequal treatment as fair when it is related to gender norms. Items were judged to be more likable when they were gendered, and likability judgments predicted more positive evaluations of the distributions, even when controlling for age.

KEYWORDS

distributive justice, gender development, middle childhood, moral development

1 | INTRODUCTION

Children observe boys and girls being treated differently every day. Boys and girls are provided with different restrooms, different sections of the toy store, different clothing. Some of the differences may be relatively equivalent, but at times, one sex or the other receives resources worth more than the other. Some research suggests that children around the ages of 6–8 years do not judge differential treatment to be unfair when the resources provided are consistent with gender norms, even when the resources favor one sex over the other (Conry-Murray, 2015). However, other research (Conry-Murray & Turiel, 2012; Killen & Stangor, 2001) shows that children as young as 4 reject exclusion of one sex from an activity. The aim of the current research is to examine how children come to the conclusion that differential treatment based on sex is sometimes acceptable.

Several lines of research show that from a young age, children are sensitive to moral issues, including issues of fairness (e.g., Geraci & Surian, 2011; Smetana, 2013; Turiel, 2015). These studies show that children reject unfairness when it is clear and straightforward. However, issues of gender equity require children to coordinate moral issues of fairness with concerns in other domains related to conventional gender norms and personal preferences. Social domain theory (Smetana, 2013; Turiel, 2015) provides a framework for examining reasoning about social issues from multiple domains, including the moral, conventional, and personal domains. Moral issues include issues of welfare or justice. Conventional issues include social rules based on authority or consensus that differ in different situations, including gender norms. Personal issues are those that are not regulated and are left to the individual to decide. Research from a social domain theory perspective (reviewed in Smetana, 2013 and Turiel, 2015) shows that children reason about moral situations differently than conventional or personal issues. Children as young as 3 years old judge moral issues to be less authority-dependent and less context-dependent than conventional issues (Smetana, Jambon, Conry-Murray, & Sturge-Apple, 2012), while personal issues are left to the individual to decide (Lagattuta, Nucci, & Bosacki, 2010). However, reasoning about fairness within the context of gender norms requires children to coordinate moral concerns with fairness and conventional concerns with gender norms. Personal preferences may also be relevant if children judge that gender-norm-related choices are a matter of personal choice. How do children coordinate conventions and personal preferences when they reason about fairness? Previous research on children's ability to consider fairness in light of other types of relevant information has shown mixed results.

Early research found that children have difficulty considering any type of contextual information when making moral judgments (Damon, 1977). However, more recent research indicates that children are able to make social judgments about fairness that include contextual information such as merit by around preschool (Baumard, Mascaro, & Chevallier, 2012) and, as children age, they are increasingly capable of considering other relevant information that affects the fairness of distributions (Schmidt, Svetlova, Johe, & Tomasello, 2016), including social group membership related to race and status (Charafeddine et al., 2016; Elenbaas & Killen, 2016), and this information can lead to acceptance of unequal treatment (Olson, Dweck, Spelke, & Banaji, 2011).

One explanation for these findings is that children judge different groups by different standards; in other words that the children showed evidence of bias (Olson et al., 2011). If children are showing bias, it is important to understand the basis of their bias. Reasoning regarding preferences may be the possible mechanism to explain approval of differential treatment. Indeed, in judgments of distributions of different items to boys and girls, children are sensitive to preference, even atypical preferences. Conry-Murray and Turiel (2012) found that children as young as age 4 advocated for a child to receive a gender atypical resource if that was their preference. In another study, children distributed resources and provided more gender atypical resources to a child who expressed a preference for those resources (Conry-Murray, 2017). However, when preferences are not specified, children may rely on gender norms to help them infer preferences. Conry-Murray and Turiel (2012) found that most children at ages 4–8 judged that a gendered item should be given to the gender-norm consistent sex, when they knew nothing of the child's preferences. Thus, it appears that assumptions about preferences can be based on gender norms, and this may explain children's apparent bias.

However, none of these studies examine how children judge already-established distributions, which is necessary in order to understand how children approach inequalities already existing in their environment. One study, Conry-Murray (2015), found that children accepted inequalities based on gender norms and found that children at ages 6 and 8 often judged that everyone in a class would agree to receiving gender-norm consistent gifts, even when the gifts were of different values. The current study aims to examine the role of assumptions about preferences in judgments and reasoning about distributions based on gender.

The period of middle childhood may be a time when children's beliefs about gender are changing with development. Children become increasingly inflexible or essentialist about gender around age 4 (Levy, Taylor, & Gelman, 1995; Taylor, 1996; Taylor, Rhodes, & Gelman, 2009). During this period, children judge gender to be an inherent quality of people that is not influenced by parents or the environment where they are raised (Taylor, 1996). Around age 8, children become more aware of exceptions to the rule in regard to gender and their thinking about gender becomes more flexible (Conry-Murray & Turiel, 2012; Gelman, Taylor, & Nguyen, 2004).

If young children infer that gender is a reliable predictor of preferences, it may lead them to accept gender-norm-related distributions more than older children. The current study examines children at ages 7, 9, and 11 and their judgments of distributions of gender-oriented items (gendered or neutral) of equivalent or unequal value, as well as whether children's judgments of the likability of the resources predict evaluations of the distributions. These ages were chosen because they include children within the period of gender essentialism but also include children old enough to provide justifications so their reasoning could also be examined. An additional group of emerging adults was assessed to examine how children's thinking might differ from mature thinking.

Thus, the current study examines whether assumptions of preferences based on group norms predicts evaluations of differential treatment, and whether those assumptions about gender are stronger during the period of gender essentialism, similar to Conry-Murray (2015). Given that young children may have been responding to features of the Conry-Murray (2015) assessment other than gender norms, the current research examines two additional explanations for approval of differential treatment: (a) a lack of sensitivity to the inequality and (b) authority demands. These are described in more detail below.

Past research on distributions of gendered items has used unequal stimuli that differ qualitatively. In Conry-Murray (2015), children judged distributions of toys and classes with different values, such as a Robotics kit for the boys and an Old Maid card game for the girls. The current study expands on the findings of Conry-Murray (2015) to examine whether children may have had difficulty considering the different values of the items in their judgments of a completed distribution. Recent research confirms that children focus on quantitative equality, without considering quality differences until about age 10 (Sheskin et al., 2016). Instead, young children may focus on the numerical comparison, (i.e., if each sex got one toy, value differences in the toys do not matter). In the current study, children judged both qualitative and quantitative inequalities, and we examined whether there were age differences in judgments of qualitative versus quantitative inequalities.

Given that gender norms are conventions and conventions can be constructed by authorities, it may be possible that children would judge gendered distributions by an authority differently than those distributed by a peer. Adults often call attention to gender in ways that can reinforce rather than challenge gender norms. For example, they enforce rules about bathrooms being single-sex, they provide toys and clothing that are gender-norm consistent and they also treat boys and girls differently in some areas like chores (Lytton & Romney, 1991). Experimental research shows that children pick up on messages about gender and this affects their attitudes about gender (Hilliard & Liben, 2010). Other research by Laupa (1991) has found that children assume that teachers have knowledge that legitimatizes their authority. In either case, it is possible that an authority's distribution could be judged more favorably than a child's distribution, especially by young children. On the other hand, since conventions can also be developed through consensus (which can include peers), either distributor could be seen as having a legitimate opinion on gender norms. In the current study, resources are distributed either by a teacher or a child to examine the effect of authority on judgments of distributions.

2 | THE PRESENT STUDY

The current study examined reasoning in children at ages 7, 9, and 11 and in emerging adults, who provided an adult comparison group. Participants judged several distributions of resources to boys and girls. Distributions varied the gendered nature of the resources, the relative value of the resources, the type of value difference (qualitative or quantitative), and whether the distributor was an authority figure or a child. Distributions of gendered items were expected to be judged more positively by younger children at age 7, who were also expected to accept unequal gendered distributions at the highest rates. However, for older children and emerging adults, distributions of equivalent items were expected to be judged more positively than distributions of unequal items.

Older children were expected to be increasingly critical of the gendered distributions, with emerging adults being the most critical of gendered distributions.

We also hypothesized that *distribution likability*, a judgment that those receiving the items would like them, would predict evaluations of the distributions, over and above the effect of age.

In addition, differences between qualitatively different items and quantitatively different items were also examined. Quantitative differences in items (with one sex receiving numerically more items) were expected to be judged more unacceptable than distributions of qualitatively different items. However, the expected patterns related to age and gender valance of the items described above were expected in both conditions.

Finally, to examine whether participants' judgments reflect a respect for a teacher's authority, distributions by teachers were compared to distributions by children.

3 | METHOD

3.1 | Participants

Children ages 7 (N = 56, M_{age} = 7.24, SD = 0.83), 9 (N = 55, M_{age} = 9.25, SD = 0.52), and 11 (N = 52, M_{age} = 11.60, SD = 0.79), and emerging adults (N = 50, M_{age} = 20.76, SD = 0.87) were interviewed to assess their judgments of distributions of different resources to boys and girls. Participants were approximately evenly divided between males and females. Participants were recruited from a suburban school district and an urban Catholic school near a large east coast city in the United States. The sample included participants who identified as white (53.3%), African American (38.7%), Asian (4.2%), Latino (0.9%), or other identifiers (2.8%).

3.2 | Procedures

Saint Joseph's University board for the ethical treatment of human subjects approved the study. Parental consent was obtained for all children and children provided assent. Adult participants provided consent. Children who participated received a scented pencil as an inducement, and adults received credit through the university's psychology subject pool. All participants were interviewed individually, and interviews were audio recorded for later transcription.

In a between-subjects design, approximately half of participants were told that a teacher was providing gifts for their class ("Some teachers are choosing gifts for their classes. They are trying to decide between different gifts to hand out") and the other half of participants were told that a child was providing gifts as part of a birthday party that was attended by the entire class ("At some schools, kids invite all the kids in their class to their birthday parties. The kid whose birthday it is chooses items for the gift bags. They are trying to decide between different gifts to hand out"). Thus, each giver had the authority to give out the items, but one was an adult with additional authority over the children.

All other variables were within-subject. The items varied by condition but were the same across the two gift-giver conditions. In each case, the girls in the class got one item and the boys in the class got a different item. Both items were pictured on a page shown to participants. The items varied in terms of the relative value of the items (equivalent items, boys got a more valuable item, or girls got a more valuable item), and the gendered nature of the items (gendered or neutral items). Gendered items were chosen because they were consistent with American stereotypes and this was confirmed through pilot testing or the item showed only one sex on the packaging (e.g., a sewing machine kit showed two girls using the item on the packaging). Although some items were targeted to younger children (e.g., the Frozen doll) than others (e.g., the Bionicle), all items were targeted to children in middle childhood.

Participants also judged distributions with both qualitative and quantitative differences in the items. In the qualitatively unequal conditions, the toys were worth different amounts of money (one item was worth 3–10 times as much as the smaller item but this information was not presented to participants) and the more expensive toys were illustrated as about two times larger to emphasize the inequality. In the quantitatively unequal conditions, books were used, where one sex received a set of books and the other sex received a single book. The set of books was always larger than the single book. See Table 1 for the items that were distributed across the 12 stories in the different conditions. Interviewers shuffled the pages (which each showed one pairs of gifts) before each interview to prevent order effects.

After describing the situation where the giver decided on different items to be given to the boys and girls of an entire class (e.g., "One teacher wants to get princess coloring books for each girl and a superhero coloring books for each boy"), participants were asked several questions. *Evaluations* were assessed with a question about whether the distribution was OK or not OK, (e.g., "Is it OK to give the boys each a Bionicle and the girls each a Frozen doll?"). Participants were asked to point to one of five faces from smiling to frowning, in order to indicate how OK or not OK the distribution was. Each face was labeled, from Definitely OK (coded as 4), A little OK (coded as 3), Neither OK or not OK (coded as 2), A little not OK (coded as 1), and Not OK (coded as 0). Participants were asked to point to the evaluation they selected, and the interviewer noted this on a response sheet. The evaluation scale was followed with a request for a justification ("Why?" or "Why not?"). Coding of the justifications is described below.

In the *distribution likability* measure ("How many of the kids do you think will like these rewards/gifts?"), participants indicated their response on a scale to indicate how many people like the pair of items. The scale was visually demonstrated with pictures of groups of children (all balanced by sex), from no children, to smaller groups, to a larger group of 10 children pictured over the response labels. The response choices were: None like it (coded as 0), Only a few people like it (coded as 1), Some people like it (coded as 2), Most people like it (coded as 3), or All people like it (coded as 4). Participants were asked to point to the gift popularity level they selected, and the interviewer noted this on a response sheet.

Before the assessment began, participants were trained on the scales using simple questions (e.g., "Is it OK to eat ice cream for dinner?" and "How many people have salad for dinner?"). All children used the scales appropriately during the training so none were excluded.

TABLE 1 Assessment stimuli by condition

	Equivalent		Boys more		Girls more	
	Books	**Toys**	**Books**	**Toys**	**Books**	**Toys**
Gendered	Princess coloring book	Elsa doll	*Pippi Longstocking*	Nail polish	*The Babysitter's Club* book set (5 books)	Sewing machine kit
	Superhero coloring book	Bionicle figure	*Fablehaven* book set (5 books)	Lego set Thunder Wings	*How to Train your Dragon*	Pokemon cards
Neutral	Flags coloring book	Bear figure	*Magic Tree House* book set (4 books)	Small stuffed bear	*Magic School Bus* book set (6 books)	The Game of Life
	States coloring book	Elephant figure	*The Littles*	Large stuffed tiger	*The Velveteen Rabbit*	Small Uno card game

Note. The first item listed was described as given to each girl in the class while the second item was given to each boy.

3.3 | Justification coding

Evaluations of the distributions were followed by a request for a justification ("Why is this OK/not OK?"). Interviews were transcribed for coding of justifications. Justifications coding categories were based on previous research (Conry-Murray, 2015) but adapted using a portion of interviews from the current study. The justification categories included references to *Gender Norms* (references to traditional gender norms as legitimate, 37.2%), *Moral Fairness* (comparisons between groups indicating that distributions were fair or equivalent, 18.5%), references to *Moral Unfairness* (comparisons between groups indicating that distributions were not equivalent or were unfair, 16%), and *Personal Choice* (references to the personal desires of those involved, 14.7%). An additional category for responses that were uncodable or for responses with missing responses was labeled *Unelaborated* (9.7%). Other justification categories were included in the coding, but they were used so rarely that they were not included in analysis. These included references to the authority figure having a bias against one sex (1.8%), references to the teacher's authority (0.2%), and the cultural context (0.2%). Coding for justifications was dichotomous. When a justification was used, it was coded as 1; when it was not used, it was coded as 0. Up to two justifications were coded for each response, and when participants used multiple justifications, proportional coding was used (i.e., each of the two justifications was coded as 0.5).

Trained research assistants coded 15% of interviews to assess reliability. Cohen's kappa was 0.71, indicating acceptable reliability.

4 | RESULTS

4.1 | Evaluations of distributions

Since participants came from two very different school districts, initial analyses were conducted to examine whether this was a factor but we found no significant main effects or interactions for ethnicity or district.

Hypothesized differences related to sex, authority condition and item type were examined. No effects for these variables were found. There were also no significant sex differences (Males = 2.87, SD = 0.76; Females = 2.83, SD = 0.78; p = 0.849). The lack of effect for authority condition (child condition: M = 2.86, SD = 0.75; teacher condition: M = 2.85, SD = 0.78; p = 0.492) indicated that participants did not judge distributions by a child significantly differently than distributions by a teacher.

There was also no significant effect for the type of item (books with clear numerical inequality M = 2.76, SD = 0.91 vs. toys with qualitative inequalities M = 2.78, SD = 0.88, p = 0.541). To be sure that item type had no effect, even within each age group and even when the difference would be most noticeable (when the items were unequal), a repeated measures ANOVA with books versus toys was conducted within each age group in the unequal conditions. Within each age group, books versus toys was again not significant.

Although evaluations of distributions when boys received the more valuable item (M = 2.67, SD = 1.13) compared to when girls received the higher value item (M = 2.55, SD = 1.10) were judged significantly differently in the omnibus ANOVA (p = 0.029), they were no longer significantly different (p = 0.062) when sex, condition, and item type were removed from the analysis. There was also no evidence of in-group bias, as would be seen in an interaction between participant sex and boys or girls receiving a better item (p = 0.574). Females judged the distributions where boys received more (M = 2.61, SD = 1.16) and girls received more (M = 2.54, SD = 1.16) similarly, and boys also judged the distributions where boys received more (M = 2.72, SD = 1.11) and girls received more (M = 2.60, SD = 1.06) similarly. Therefore, these evaluations were averaged for a new category called unequal distributions.

The final analysis was conducted with a 2 (Equivalent status: Equivalent, Unequal) × 2 (Gendered/Neutral status) × 4 age group (7, 9, 11 and adults) repeated measures ANOVA with Equivalent status and Gendered/Neutral status, as repeated measures. Interactions were followed up with Bonferroni corrected alpha levels.

TABLE 2 Judgments of the acceptability of the distribution by age, equivalency status, and gendered/neutral status of the rewards

Equivalency status	Gendered items				Neutral items				Both				
	7	9	11	Adults	7	9	11	Adults	7	9	11	Adults	Total
Equivalent	3.66 (0.74)	3.71 (0.69)	3.54 (0.67)	2.67 (1.33)	2.66 (1.32)	2.81 (0.91)	2.89 (0.98)	2.79 (1.09)	3.16_a (0.79)	3.26_a (0.60)	3.22_a (0.60)	2.73_b (0.94)	3.10 (0.77)
Unequal	3.30 (1.11)	3.11 (0.91)	2.85 (1.03)	1.80 (1.29)	2.69 (1.22)	2.68 (0.98)	2.41 (0.98)	1.88 (1.29)	3.00_a (1.05)	2.90_a (0.84)	2.63_a (0.87)	1.84_b (1.10)	2.61 (1.06)
Both	3.48_a (0.78)	3.41_a (0.58)	3.20_a (0.68)	2.24_b (1.17)	$2.68_{a,b}$ (1.08)	2.75_b (0.69)	$2.65_{a,b}$ (0.70)	2.33_a (1.03)	3.08 (0.80)	3.08 (0.51)	2.92 (0.52)	2.28 (0.87)	2.86 (0.76)

Note. Acceptability judgments were scored from 0 (definitely not OK) to 4 (definitely OK). Means in the same row within a condition with subscripts that differ, differ at $p < 0.025$ (Bonferroni corrected level).

Three main effects were found. A main effect for gendered versus neutral items, $F(1, 209) = 47.83$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.19$, confirmed that gendered items ($M = 3.10$, $SD = 0.96$) were judged to be more acceptable in distributions to boys and girls compared to gender-neutral items ($M = 2.61$, $SD = 0.90$). The main effect for equivalent status, $F(1, 209) = 50.97$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.20$, indicated that distributions of equivalent items were judged to be more acceptable than distributions of unequal items (see Table 2 for means and *SD*s.) Finally a main effect for age, $F(3, 209) = 15.26$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.18$, indicated that adults were more critical of all distributions than each of the younger age groups, who did not differ significantly, (all *p*s < 0.001 between adults and each child age-group), as shown in Table 2.

All the main effects were qualified by interactions. First, a gendered/neutral status × age effect, $F(3, 209) = 8.01$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.10$, was followed up. As expected, children at all ages judged distributions of gendered items more positively than neutral items (all $p < 0.001$), but adults did not distinguish between the two ($p = 0.612$). See Table 2 for means and Figure 1 for a visual display which shows that the size of the effect grew smaller after age 9 (for ages 7, 9, 11 and adults, $\eta^2_{\text{partial}} = 0.41$, 0.43, 0.28 and 0.01, respectively).

A second interaction was between equivalent status and age, $F(3, 209) = 4.94$, $p = 0.002$, $\eta^2_{\text{partial}} = 0.07$. Follow-up tests indicated only 7 year olds did not distinguish between distributions of unequal and equivalent items. The 9 and 11 year olds and the adults distinguished between equivalent distributions and unequal distributions ($p < 0.012$, $p < 0.001$, and $p < 0.001$ respectively), evaluating unequal distributions as less acceptable than equivalent distributions, although the older children and adults made a larger distinction ($\eta^2_{\text{partial}} = 0.11$, 0.24, and 0.42, for ages 9, 11, and adults, respectively) as Table 2 shows.

An equivalency status × gendered/neutral interaction, $F(1, 209) = 8.39$, $p = 0.004$, $\eta^2_{\text{partial}} = 0.04$, indicated that distributions of gendered items were judged to be more acceptable than neutral distributions in both equivalent, $F(1, 209) = 45.13$, $p < 0.001$, $\eta^2 = 0.18$ and unequal distributions, $F(1, 209) = 22.77$, $p < 0.001$, $\eta^2 = 0.10$, but the effect was stronger for equivalent items. Means and *SD*s can be found in Table 3.

4.2 | Evaluation justifications

Justifications with 10% or more of responses were included in analyses. Because of concerns with the lack of independence of the justifications, each justification was analyzed separately. Each was analyzed with a 2 (Equivalent status: Equivalent, Unequal) × 2 (Gendered/Neutral status) × 4 (Age group: 7, 9, 11 and adults) repeated measures ANOVA with Equivalent status and Gendered/Neutral status as repeated measures. Interactions were followed up with Bonferroni-corrected alpha levels. Means and standard deviations for each justification by Age group, Equivalent Status and Gendered/Neutral Status are located in Table 4.

4.2.1 | Gender-related justifications

With 37% of responses, gender-related justifications were used the most frequently, and there were several main effects and one interaction. First a main effect for gendered versus neutral status, $F(1, 166) = 308.70$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.65$, indicated that gender-related justifications were used more when items distributed were related to gender norms, as expected. A main effect of equivalent status, $F(1, 166) = 9.31$, $p = 0.003$, $\eta^2_{\text{partial}} = 0.05$, showed that the gender-related justifications were used more when items were equivalent than when they were unequal. A gendered versus neutral × equivalent status interaction $F(1, 166) = 9.31$, $p = 0.003$, $\eta^2_{\text{partial}} = 0.05$, showed that this effect was especially true in gendered distributions where gender justifications were used more when items were equivalent ($M = 0.65$, $SD = 0.40$) than when items were unequal ($M = 0.52$, $SD = 0.35$).

A main effect for age, $F(1, 166) = 10.11$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.16$, showed that younger aged participants (age 7 and 9) used the gender-related justification more than adults ($p < 0.001$).

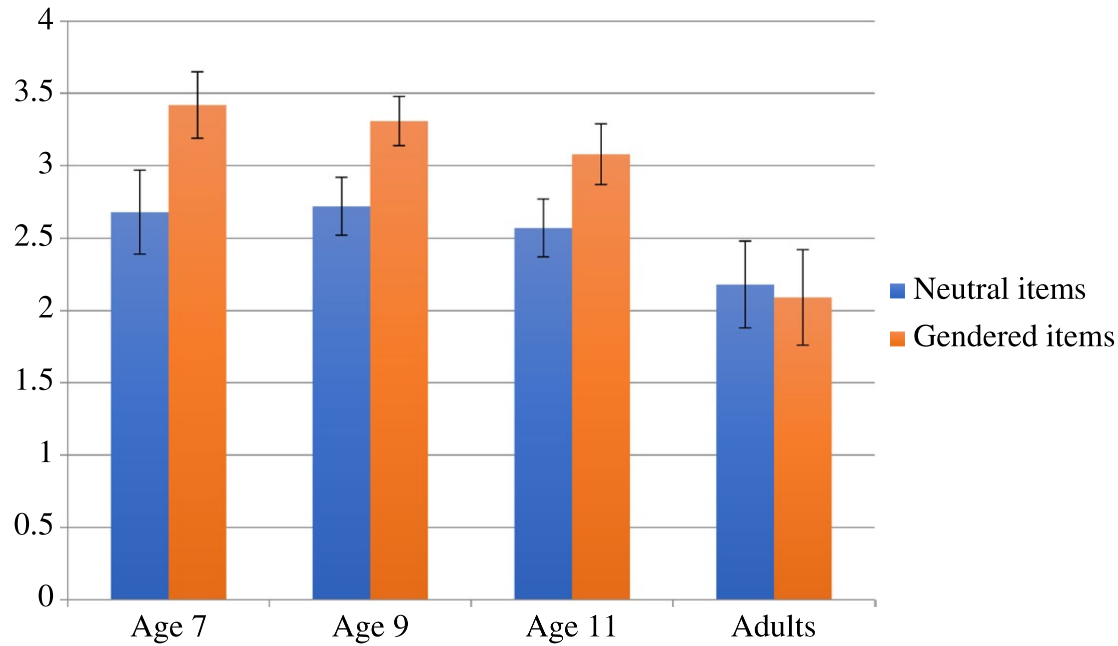

FIGURE 1 Evaluations of distribution on a scale from 0 (Definitely not OK) to 4 (Definitely OK). Bars are 95% confidence intervals

TABLE 3 Judgments of the acceptability of the distribution by gendered/neutral and equivalency status of the rewards

	Gendered	Neutral
Equivalent	3.41_a (0.98)	2.78_b (1.09)
Unequal	2.79_a (1.22)	2.43_b (1.16)

Note. Acceptability judgments were scored from 0 (definitely not OK) to 4 (definitely OK). Means in the same row with subscripts that differ were statistically significantly different at $p < 0.05$.

4.2.2 | Fairness justifications

The second most common justification referred to the fairness of the distributions (18% of responses). A main effect for gendered versus neutral status, $F(1, 166) = 102.18$, $p < 0.001$, $\eta^2_{partial} = 0.38$, indicated that fairness justifications were used more when items distributed were not related to gender norms.

A main effect of equivalent status, $F(1, 166) = 12.78$, $p = 0.003$, $\eta^2_{partial} = 0.07$, showed that distributions were justified as fair more when items were equivalent than when they were unequal.

Finally a main effect for age, $F(3, 166) = 11.40$, $p < 0.001$, $\eta^2_{partial} = 0.17$, showed that younger aged participants at age 7 used the fairness justification significantly less than adults ($p < 0.001$), but no other age group comparisons were significantly different.

All of the main effects were qualified by interactions. A gendered versus neutral status × equivalent status interaction, $F(1, 166) = 14.32$, $p < 0.001$, $\eta^2_{partial} = 0.08$, indicated that participants used references to the fairness of the distribution rarely in gendered distributions, which did not differ by equivalent status (Equivalent: $M = 0.05$, $SD = 0.12$; Unequal: $M = 0.09$, $SD = 0.20$). In neutral distributions, fairness justifications were used more when distributions were equivalent ($M = 0.37$, $SD = 0.40$) than when they were unequal ($M = 0.23$, $SD = 0.30$, $p < 0.001$).

A equivalent status × age interaction, $F(3, 166) = 6.55$, $p < 0.001$, $\eta^2_{partial} = 0.11$, showed that adults used the fairness justification more than all other ages, but the difference was only significant within the unequal distributions (adults: $M = 0.35$, $SD = 0.32$; 11 year olds: $M = 0.13$, $SD = 0.15$; 9 year olds: $M = 0.09$, $SD = 0.11$; 7 year olds: $M = 0.10$, $SD = 0.17$, $p < 0.001$ for all comparisons with adults) .

4.2.3 | Unfair justifications

Sixteen percent of justifications referred to concerns that the distributions were unfair. A main effect of equivalent status, $F(1, 166) = 61.81$, $p < 0.001$, $\eta^2_{partial} = 0.27$, showed that justifications referring to the unfairness of the distribution were used more when items were unequal than when they were equivalent.

A main effect for gendered versus neutral status, $F(1, 166) = 27.95$, $p < 0.001$, $\eta^2_{partial} = 0.14$, indicated that references to unfairness were more common in distributions of neutral items than gendered items.

4.2.4 | Personal choice justifications

Fifteen percent of responses referred to personal choice. A main effect for gendered versus neutral status, $F(1, 166) = 17.48$, $p < 0.001$, $\eta^2_{partial} = 0.10$, indicated that personal choice justifications were used more when items distributed were not related to gender norms. This effect was qualified by a gendered versus neutral status × age interaction, $F(3, 166) = 8.34$, $p = 0.004$, $\eta^2_{partial} = 0.05$, which showed that distinctions between gendered and neutral conditions were only made by 7 (gendered: $M = 0.06$, $SD = 0.13$, Neutral: $M = 0.18$, $SD = 0.27$, $p = 0.003$),) and 9 year olds (gendered: $M = 0.05$, $SD = 0.10$, Neutral: $M = 0.25$, $SD = 0.25$, $p < 0.001$), when justifying evaluations with references to personal choice. Eleven year olds (gendered: $M = 0.14$, $SD = 0.22$, Neutral: $M = 0.15$, SD

TABLE 4 Mean justifications (and *SD*s) by age, equivalent status, and gendered/neutral

	Age				Distribution		Items		
Justification	**7**	**9**	**11**	**18**	**Equivalent**	**Unequal**	**Gendered**	**Neutral**	**Total**
Gender	0.49_{a} (0.28)	$0.47_{a,b}$ (0.20)	$0.32_{b,c}$ (0.22)	0.26_{c} (0.17)	0.40* (0.27)	0.36* (0.27)	0.58+ (0.32)	0.18+ (0.25)	0.38 (0.24)
Fair	0.13_{a} (0.18)	$0.12_{a,b}$ (0.11)	$0.22_{a,b}$ (0.17)	0.33_{b} (0.25)	0.23* (0.24)	0.16* (0.22)	0.10+ (0.17)	0.30+ (0.30)	0.20 (0.20)
Unfair	0.09_{a} (0.16)	0.15_{ab} (0.14)	0.19_{b} (0.18)	0.11_{ab} (0.13)	0.10* (0.15)	0.22* (0.27)	0.10+ (0.16)	0.17+ (0.20)	0.13 (0.16)
Choice	0.12_{a} (0.18)	$0.15_{a,b}$ (0.15)	$0.15_{a,b}$ (0.17)	0.19_{b} (0.20)	0.17* (0.22)	0.14* (0.19)	0.11+ (0.19)	0.19+ (0.26)	0.15 (0.18)

Note. Means are proportions. Within Age, means with subscripts that differ, differ at $p < 0.05$. Within Distribution types, * indicates that means differ at $p < 0.05$. Within Item Type, + indicates that means differ at $p < 0.05$.

= 0.23),) and adults (gendered: $M = 0.18$, $SD = 0.23$, Neutral: $M = 0.20$, $SD = 0.29$) referred to personal choice in both conditions.

4.3 | Perceived likability of distributions

Like the analysis of judgments, initial analyses of judgments of how many of the children would like the rewards in the distributions found no significant ethnicity, sex or authority condition effects, so these variables were dropped. A 2 (equivalent status: Equivalent, Unequal) × 2 (gendered vs. neutral) × 2 (Value difference type: qualitative vs. quantitative) × 4 age group (7, 9, 11 and adults) repeated measures ANOVA with Equivalent status, Gendered versus Neutral, and Value difference type as repeated measures was conducted.

A main effect for gendered versus neutral, $F(1,209) = 93.51$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.31$, indicated that participants judged that distributions of gendered items would be liked by more than distributions of neutral items. A main effect for Equivalency status $F(1, 209) = 13.29$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.06$, indicated that distributions of equivalent items were judged to be liked by more children than distributions of unequal items. A main effect for age, $F(3, 209) = 12.71$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.15$, indicated that that younger children judged that the distributions would be more likable, while older children and adults expected the distributions to be less likable. Means for these three main effects are in Table 5. A value difference type effect, $F(1, 209) = 11.06$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.05$, indicated that items in qualitatively different distributions (i.e., toys, $M = 2.93$, $SD = 0.57$) were seen as more likable than items in quantitatively different distributions (i.e., books, $M = 2.71$, $SD = 0.63$).

These main effects were qualified by an interaction between gender versus neutral and age, $F(3,209) = 3.19$, $p = 0.025$, $\eta^2_{\text{partial}} = 0.04$. Follow-up tests found that children at age 7 ($p < 0.001$), 9 ($p < 0.001$), and 11 ($p < 0.001$), rated distributions of gendered items as likeable to more, but adults made no distinction. See Table 5 for means.

Finally, a gendered/neutral × equivalency status interaction, $F(3, 209) = 37.31$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.15$, indicated that when items were gendered, distributions of equivalent items were judged to be liked by more children than distributions of unequal items (Equivalent: $M = 3.29$, $SD = 0.74$; Unequal: $M = 2.85$, $SD = 0.74$; $p < 0.001$). However, when items were neutral, there were no significant differences depending on equivalency status (Equivalent: $M = 2.58$, $SD = 0.88$; Unequal: $M = 2.66$, $SD = 0.71$).

4.4 | Predicting judgments of distributions with perceived likability

Participants' judgments that the distributions would be likable to children were expected to predict positive evaluations of the distributions. To test this hypothesis, the judgments of toys and books were averaged to create four means: unequal gendered, unequal neutral, equivalent gendered, and equivalent neutral distributions. Mean

TABLE 5 Ratings of how many of the children would like the rewards by age, gendered/neutral status, and equivalency status

	Item		Distribution		
Age group	Gendered	Neutral	Equivalent	Unequal	Total
7	3.40_a (0.60)	2.91_b (0.66)	3.18_a (0.60)	3.13_a (0.59)	3.14 (0.54)
9	3.18_a (0.55)	2.58_b (0.53)	2.88_a (0.67)	2.88_a (0.56)	2.88 (0.46)
11	2.98_a (0.55)	2.57_b (0.56)	2.94_a (0.61)	2.61_b (0.57)	2.77 (0.47)
Adults	2.70_a (0.43)	2.50_a (0.67)	2.74_a (0.52)	2.46_b (0.52)	2.60 (0.43)
Total	3.00_a (0.62)	2.66_b (0.61)	2.94_a (0.62)	2.78_b (0.61)	2.86 (0.52)

Note. Judgments were scored from 0 (no one would like the rewards) to 4 (everyone would like the rewards). Subscripts that differ within Item Type or within Distributions, differ at $p < 0.05$.

judgments of how many of the children in the class would like a distribution in these four conditions were entered as predictors of evaluations of the distributions. To see whether these effects existed above and beyond the effect of age, age was also included as a control in the regressions. Distribution likability judgments for each condition were entered for all four regressions. Only the distribution likability judgments of the DV condition significantly predicted judgments in each of the four regressions in the predicted direction (βs between 0.39 and 0.58, all at $p < 0.001$). Thus, judgments that more children liked the distributions predicted judgments that the distribution was acceptable over and above the effect of age, but only within the same condition. See Table 6 for standardized coefficients.

5 | DISCUSSION

Overall, differential treatment of boys and girls was seen as more acceptable when the items being distributed were consistent with gender norms, and this effect was strongest for the youngest participants at ages 7 and 9. Participants' justifications confirmed that gender norms were important to their judgments, especially for the two youngest age groups. Further, as expected, gendered distributions were judged to be liked by more people, and judgments that the distributions were liked by many predicted more positive evaluations of the distributions, over and above the effect of age. These findings are consistent with Conry-Murray (2015), despite some differences in methodology including a scale rather that dichotomous judgments and different items, indicating that these findings are relatively robust.

In a novel finding, responses were not affected by the authority of the person distributing the items. In a between-subjects design, distributions were judged similarly, whether it was a teacher or a child who distributed the items. Judgments of distributions also did not differ when the items differed qualitatively compared to quantitatively, despite that past research has shown that children under age 10 do not notice qualitative differences (Sheskin et al., 2016). While these findings should continue to be investigated, these initial results indicate neither a non-authority figure distributing the items nor numerical inequality of the items helped participants judge the distributions as unfair.

Participants may have judged that gendered items are more appealing to the recipients if they believed that boys and girls have different preferences. This explanation extends previous research on children's biased standards (McGillicuddy-De Lisi, Daly, & Neal, 2006; Olson et al., 2011) to identify one possible source of bias. Participants' assumptions about preferences can differ for boys and girls, and therefore providing them with different items that match their different preferences may have seemed reasonable to participants, (though there is reason to question the accuracy of participants' assumptions about children's preferences, Conry-Murray, 2013). The suggestion that assumptions about preferences are a mechanism that links gender norms to evaluations is further supported by the fact that no in-group bias or other sex differences were found. In other words, the bias was not general. Instead, it was found in specific cases that were related to assumptions about preferences, as the regressions showed. Future research should investigate how relevant beliefs influence other types of bias, such as that demonstrated in distributive justice based on race or status, as in past research (McGillicuddy-De Lisi et al., 2006; Olson et al., 2011).

Although references to gender norms were the most common justifications, the current study also found that participants mentioned concerns with fairness or unfairness, indicating that justice was an important factor for them, as in past research (Conry-Murray, 2015; Geraci & Surian, 2011; Turiel, 2015). Participants were most sensitive to unfairness when items were unequal and gender-neutral. It seems that larger differences in the value of the items highlighted the issue of justice. Distributions of gender neutral items may have been judged as unfair more without the distraction of gender norms. As discussed above, some may have believed that the gendered distributions were consistent with preferences and therefore fair.

TABLE 6 Regression results with perceived likability of distributions predicting evaluations of distributions

Criterion variable	Predictor variable	Standardized coefficient	*p*	R^2
Equivalent neutral	Likability of equivalent neutral distributions	0.39	<0.001	0.16
Equivalent gendered	Likability of equivalent neutral distributions	0.49	<0.001	0.20
Unequal neutral	Likability of equivalent neutral distributions	0.47	<0.001	0.24
	Likability of equal gendered distributions	−0.14	0.030	
Unequal gendered	Likability of equivalent neutral distributions	0.58	<0.001	0.30

Despite their apparent concern with justice in their justifications, participants' evaluations also indicated that they often saw it as acceptable to provide different items to boys and girls. This was most common when the items were equivalent, perhaps because each child received an item. However, in each case, the items the boys received differed from the items the girls received. Judging distributions that provided different but gendered items required participants to balance concerns with fairness and concerns with adherence to gender norms. For many participants, the consistency between the gendered items and the sex of the recipient may have overshadowed information about the inequality of providing different items based on sex.

While concerns with both gender norms and fairness were evident in the justifications provided by every age group, there were important differences in the degree to which each group prioritized gender or fairness. Age differences indicated that younger children were more strongly influenced by the gendered nature of the items distributed. Children ages 7 and 9 approved of gendered distributions more than neutral distributions, while 11 year olds similarly distinguished the two but with a smaller effect size. Justifications also indicated that children at ages 7 and 9 were especially focused on gender. At these ages, participants were more likely than emerging adults to use gender-related justifications. Although some research (Conry-Murray & Turiel, 2012) has found increasing flexibility about gender after age 7 or 8, the current study found few differences between 7 and 9 year olds. However, differences between the 7 and 9 year olds, and older children and emerging adults did show an overall increase in flexibility, consistent with past literature (Taylor et al., 2009). Research on gender essentialism (Levy et al., 1995; Taylor, 1996; Taylor et al., 2009) shows that children exaggerate the importance of the category of gender more than emerging adults. In the current study, emerging adults did not seem to consider the gender norms of the items as a justification for providing different resources to boys and girls. This is evidence that in middle childhood, children are more likely than emerging adults to consider gender as relevant to judgments of distributions.

Personal choice is important to children, but the younger children saw it as more relevant when the items were gender-neutral. Children at ages 7 and 9 used the personal choice justification more in gender neutral distributions, while they used the gender norms justification more when items were gendered. If young children assume that gender can be used as proxy for a stated personal preference, it makes sense that they would not need to refer to personal choice since they only needed to mention gender. In fact, children at all ages rated gendered items as more likable. Only emerging adults rated gendered and neutral items as equally likable. Still, similar to Conry-Murray (2015), the current study found that the older children, like the adults, were sensitive to the need to refer to personal preferences. Emerging adults and 11 year olds both mentioned personal choice at the same rates across both gendered and gender-neutral conditions. For emerging adults and older children, the diversity of possible personal preferences was relevant, even when gendered items were provided.

In middle childhood, children may be less critical of information about norms, and gender is a particularly strong norm. Research on distributive justice in the context of other social norms (Schmidt et al., 2016) has shown that children consider the rules of a game that have been agreed upon by the group, when deciding on rewards. However, it was not until age 8 that children evaluated the additional information and rejected idiosyncratic reasons for inequality. Conventional concerns with well-established conventional norms, including gender norms may be more influential than the rules of a temporary game, but both appear to be evaluated more carefully after age 8.

When preferences are made explicit, even children as young as 4 use them to assign rewards, even when the preferences are counter to gender norms (Conry-Murray, 2013, 2017; Conry-Murray & Turiel, 2012). However, when information is not made explicit, younger children may not evaluate gender norms if they are unaware that preferences can vary for different individuals. Children's understanding of norms is less flexible than emerging adults, perhaps because children construct their judgments based on different assumptions related to whether gender reflects individual preferences.

Highlighting the quantitative differences in the rewards did not affect judgments for any age group, not even the 7 year olds, who might be expected to benefit from the explicit numerical inequality in the distributions, since numerical equality has been found to be understood earlier than quality-based equality in past research (Sheskin et al., 2016). Judgments of qualitative versus quantitative item-differences did not differ within any age group, even where they would most likely to be found, in the unequal distributions, despite that toys were seen as more likable than books.

There were no differences in judgments based on whether the person distributing the items was an authority figure or not. The lack of difference between the child-distributor and teacher-distributor conditions indicates that teachers were not seen as having special knowledge that would explain the reason for their distribution of different items to boys and girls, as in some past research (Laupa, 1991). The current study included different contexts, and party gifts may be seen as being more discretionary than a teacher giving gifts to a class. However, this would seem to enhance differences rather than diminish them and no differences were found.

There were also no sex differences, either based on participant sex or based on the sex of the children receiving rewards, indicating that in-group bias or differences in the degree to which norms apply to boys and girls were not evident in this data.

The current findings suggest that children may not challenge differential treatment when they believe gender norms justify it, even when one group is provided with more valuable items. Parents and teachers should be especially sensitive to differential treatment that may affect both boys' and girls' opportunities, and they may want to intercede to ensure that all children are treated fairly and provided with equal opportunities.

Although the current study found evidence of children's support for treating boys and girls differently, other research shows that they can also advocate for fair treatment when given the opportunity. Grocke, Rossano, and Tomasello (2015) found that children who had the opportunity to change a game to make it more fair, took advantage of this and adapted the game to make it more fair. Research assessing children's own distributions of gendered stickers (Conry-Murray, 2017) also shows a concern for fairness, even when it is in conflict with gender norms. In that study, participants had to choose between an equal distribution and a gender-norms consistent distribution (e.g., they divided up four masculine and two feminine stickers between a boy and a girl). Children as young as 7 provided a masculine sticker to a girl or a feminine sticker to a boy, rather than giving one sex more stickers than the other, indicating that their concern with fairness was dominant. The assessment in the current study did not allow for children to propose different distributions.

A limitation of the current study is the focus on children at the oldest ages of gender essentialism. Future research with more simplified methods should examine younger children's reasoning to determine whether children at the height of gender essentialism accept gender-norm consistent inequalities at even greater levels, especially since there is evidence that children who have less advanced in theory of mind have difficulty distributing rewards based on effort when the effort is in a counterstereotypical task (Rizzo & Killen, 2018).

In conclusion, the current study shows that children judge distributions of different items to girls and boys more positively when those items are related to gender, and children at age 7 and 9 are especially prone to consider gender norms as a justification for differential treatment of boys and girls. Acceptance of unequal treatment appears to be related to perceptions of personal preferences. Over and above the effect of age, judgments that more children would like the rewards significantly predicted positive evaluations of the distributions. Assumptions that gender norms accurately predict preferences appear to lead to an acceptance of unequal treatment. This is a concern because it could mean that children will not challenge unequal treatment of boys and girls. Future research should examine whether sensitivity to the variety of preferences predicts more critical judgments of unequal treatment of boys and girls.

ORCID

Clare Conry-Murray http://orcid.org/0000-0003-0245-1289

REFERENCES

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology*, *48*, 492–498. https://doi.org/10.1037/a0026598

Charafeddine, R., Mercier, H., Clément, F., Kaufmann, L., Reboul, A., & Van der Henst, J. (2016). Children's allocation of resources in social dominance situations. *Developmental Psychology*, *52*, 1843–1857. https://doi.org/10.1037/dev0000164

Conry-Murray, C. (2013). Children's reasoning about gender-atypical preferences in different settings. *Journal of Experimental Child Psychology*, *115*, 210–217. https://doi.org/10.1016/j.jecp.2012.09.007

Conry-Murray, C. (2015). Boys get the math games and girls get the reading games: Children's reasoning about distributive justice and gender. *Merrill-Palmer Quarterly*, *61*, 319–344. https://doi.org/10.13110/merrpalmquar1982.61.3.0319

Conry-Murray, C. (2017). Children's distributive justice: The role of gender norms in different settings. *European Journal of Developmental Psychology*, *14*, 62–74. https://doi.org/10.1080/17405629.2016.1152176

Conry-Murray, C., & Turiel, E. (2012). Jimmy's baby doll and Jenny's truck: Young children's reasoning about gender norms. *Child Development*, *83*, 146–158. https://doi.org/10.1111/j.1467-8624.2011.01696.x

Damon, W. (1977). *The social world of the child*. San Francisco, CA: Jossey-Bass.

Elenbaas, L., & Killen, M. (2016). Children rectify inequalities for disadvantaged groups. *Developmental Psychology*, *52*, 1318–1329. https://doi.org/10.1037/dev0000154

Gelman, S. A., Taylor, M. G., & Nguyen, S. P. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monographs of the Society for Research in Child Development*, *69*(1). https://doi.org/10.1111/j.1540-5834.2004.06901007.x

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, *14*, 1012–1020. https://doi.org/10.1111/j.1467-7687.2011.01048.x

Grocke, P., Rossano, F., & Tomasello, M. (2015). Procedural justice in children: Preschoolers accept unequal resource distributions if the procedure provides equal opportunities. *Journal of Experimental Child Psychology*, *140*, 197–210. https://doi.org/10.1016/j.jecp.2015.07.008

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*, 1787–1798. https://doi.org/10.1111/j.1467-8624.2010.01510.x

Killen, M., & Stangor, C. (2001). Children's reasoning about social inclusion and exclusion in gender and race peer group contexts. *Child Development*, *72*, 174–186. https://doi.org/10.1111/1467-8624.00272

Lagattuta, K. H., Nucci, L., & Bosacki, S. L. (2010). Bridging theory of mind and the personal domain: Children's reasoning about resistance to parental control. *Child Development*, *81*, 616–635. https://doi.org/10.1111/j.1467-8624.2009.01419.x

Laupa, M. (1991). Children's reasoning about three authority attributes: Adult status, knowledge, and social position. *Developmental Psychology*, *27*, 321–329. https://doi.org/10.1037/0012-1649.27.2.321

Levy, G. D., Taylor, M. G., & Gelman, S. A. (1995). Traditional and evaluative aspects of flexibility in gender roles, social conventions, moral rules, and physical laws. *Child Development*, *66*, 515–531. https://doi.org/10.2307/1131594

Lytton, H., & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin*, *109*, 267–296. https://doi.org/10.1037/0033-2909.109.2.267

McGillicuddy-De Lisi, A. V., Daly, M., & Neal, A. (2006). Children's distributive justice judgments: Aversive racism in Euro-American children? *Child Development*, *77*, 1063–1080. https://doi.org/10.1111/j.1467-8624.2006.00919.x

Olson, K. R., Dweck, C. S., Spelke, E. S., & Banaji, M. R. (2011). Children's responses to group-based inequalities: Perpetuation and rectification. *Social Cognition*, *29*, 270–287. https://doi.org/10.1521/soco.2011.29.3.270

Rizzo, M. T., & Killen, M. (2018). Theory of mind is related to children's resource allocations in gender stereotypic contexts. *Developmental Psychology*, *54*, 510–520. https://doi.org/10.1037/dev0000439

Schmidt, M. H., Svetlova, M., Johe, J., & Tomasello, M. (2016). Children's developing understanding of legitimate reasons for allocating resources unequally. *Cognitive Development*, *37*, 42–52. https://doi.org/10.1016/j.cogdev.2015.11.001

Sheskin, M., Nadal, A., Croom, A., Mayer, T., Nissel, J., & Bloom, P. (2016). Some equalities are more equal than others: Quality equality emerges later than numerical equality. *Child Development*, *87*, 1520–1528. https://doi.org/10.1111/cdev.12544

Smetana, J. (2013). Moral development: The social domain theory view. In P. Zelazo (Ed.), *Oxford handbook of developmental psychology* (pp. 832–863). Oxford, UK: Oxford University Press.

Smetana, J., Jambon, M., Conry-Murray, C., & Sturge-Apple, M. (2012). Reciprocal associations between young children's moral judgments and their developing theory of mind. *Developmental Psychology*, *48*, 1133–1144. https://doi.org/10.1037/a0025

Taylor, M. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development*, *67*, 1555–1571. https://doi.org/10.1111/j.1467-8624.1996.tb01814.x

Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys, cows will be cows: Children's essentialist reasoning about gender and animal development. *Child Development*, *80*, 461–481. https://doi.org/10.1111/j.1467-8624.2009.01272.x

Turiel, E. (2015). Morality and prosocial judgments and behavior. In D. A. Schroeder, W. G. Graziano, D. A. Schroeder, & W. G. Graziano (Eds.), *The Oxford handbook of prosocial behavior* (pp. 137–152). New York, NY: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399813.013.022

How to cite this article: Conry-Murray C. Children's reasoning about unequal gender-based distributions. *Social Development*. 2018;00:1–17. https://doi.org/10.1111/sode.12342