

## Jimmy's Baby Doll and Jenny's Truck: Young Children's Reasoning About Gender Norms

Clare Conry-Murray  
*Pennsylvania State University*

Elliot Turiel  
*University of California, Berkeley*

To assess the flexibility of reasoning about gender, children ages 4, 6, and 8 years ( $N = 72$ ) were interviewed about gender norms when different domains were highlighted. The majority of participants at all ages judged a reversal of gender norms in a different cultural context to be acceptable. They also judged gender norms as a matter of personal choice and they negatively evaluated a rule enforcing gender norms in schools. Older children were more likely to show flexibility than younger children. Justifications obtained from 6- and 8-year-olds showed that they considered adherence to gender norms a matter of personal choice and they viewed the rule enforcing gender norms as unfair.

The study of sex and gender has focused on differences and similarities in the psychological characteristics of males and females, as well as on how children at different ages think about gender differences and similarities. The research reported here addressed the latter question. For many years, central questions regarding concepts of gender have revolved around whether children conceptualize gender and associated activities or characteristics as fixed or changeable (see Ruble, Martin, & Berenbaum, 2006, for a review). Are perceived gender roles judged to be flexible or inflexible? The answer to this question is not straightforward. Research does indicate that children's conceptions of activities associated with gender become increasingly flexible in later childhood (ages 8 or 9 years) and that in early childhood (ages 4 or 5 years) they appear to conceptualize activities associated with males and females as relatively fixed (Levy, Taylor, & Gelman, 1995; Ruble et al., 2006; Taylor, Rhodes, & Gelman, 2009; Trautner et al., 2005). However, research also shows inconsistencies in these trends in that young children accept deviations from gender norms on some dimensions (Ruble et al., 2006) and adolescents believe that activities associated with gender should be maintained in inflexible ways (Stoddart & Turiel, 1985).

A major reason for the seeming inconsistencies in findings may be that different aspects or domains of the issue of gender were assessed in

different studies. Gelman (2010, p. 717) provides a useful illustration of how an issue can be approached from the perspective of different domains. Her example pertains to a child's desire to eat all her Halloween candy in one night. A parent might respond in a biological domain (too much candy will make you sick), in a personal domain (you are going to want to eat more candy tomorrow), and in a moral domain (you need to share your candy with your sister). Gender norms can also be approached from different domains. In some research, children's understandings of the biological and psychological sources of gender differences were assessed by posing questions about hypothetical situations in which a child is raised only by members of the opposite sex (Levy et al., 1995; Taylor, 1996; Taylor et al., 2009). It was found that young children (4–5 years) believe that the hypothetical child will nevertheless display stereotypical characteristics (e.g., in personality traits, toy preferences, physical appearance) of the biological sex, whereas older children attribute the characteristics acquired to environmental influences (i.e., the opposite sex characteristics to which the child was exposed). On the basis of these findings, it was proposed that young children have an "essentialist" conception of gender and thus are inflexible about the attributes of males and females. However, other studies suggest that young children are not always inflexible about gender-related norms and that it is

---

Correspondence concerning this article should be addressed to Clare Conry-Murray, Department of Psychology, Penn State University, Beaver, 100 University Drive, 3K RAB, Monaca, PA 15061. Electronic mail may be sent to cec23@psu.edu.

important to examine the particular type of judgments being made about gender norms. Research indicates that children's flexibility and inflexibility about gender norms varies in accord with personal choices (Smetana, 1986), societal conventions (Carter & Patterson, 1982; Stoddart & Turiel, 1985), and their moral evaluations of exclusion from groups on the basis of gender-associated activities (Horn, 2008; Killen & Stangor, 2001; Theimer, Killen, & Stangor, 2001).

Although there is a greater tendency for young children to apply gender norms in less flexible ways than older children (Ruble et al., 2006; Trautner et al., 2005), the research does indicate that conceptions of gender norms are multidimensional and associated with different domains (moral, conventional, personal) of social judgment that children begin to develop at young ages (Nucci, 2001; Smetana, 2006; Turiel, 1983, 2002). The purpose of the present study was to examine reasoning about the development of gender norms when different domains are highlighted. To this end, we examined children's reasoning about domains as related to gender through assertions of personal preferences (the *personal* domain), variations in cultural norms or *conventions*, and the *moral* implications of imposing regulations. We expected that children would not uniformly judge gender as fixed, but that they would consider the context of each domain.

The present study was guided by theory and methods stemming from a body of work that has been referred to as "social domain theory" (see Killen & Smetana, 2006; Smetana, 2006; Turiel, 2010). The research on social domains shows that from a young age children form judgments about moral issues pertaining to welfare, justice, and rights (Nucci, 2001; Smetana, 2006; Turiel, 1983, 2002). The research also shows that even very young children distinguish morality from social conventions, which are shared uniformities that coordinate interactions (such as a rule that children address teachers by their titles). Children distinguish moral and conventional rules by judging conventions as alterable and relative to social systems, while issues in the moral domain are not alterable or relative (Davidson, Turiel, & Black, 1983; Smetana, 2006; Smetana & Braeges, 1990; Turiel, 1983, 2002, 2006). For example, hitting without provocation is viewed as moral and thus usually judged to be wrong even if no rule exists, and is also judged to be wrong across cultural contexts. By contrast, social conventions, such as the rule about how to address teachers, are judged as alterable, under the

jurisdiction of authority, and can apply to particular settings (such as school) but not others (such as home).

Studies that applied the domain distinctions to judgments about gender showed that violations of gender norms are often judged to be conventional and changeable, even at young ages. For example, Smetana (1986) found that preschoolers judged sex role violations related to activities and appearance to be more permissible and alterable than moral violations. Smetana also found that gender norms were judged to be more alterable than other conventions, suggesting that some sex role issues may be seen as a matter of personal choice. A study by Carter and Patterson (1982) found age differences in judgments about changes in gender-related activities regarding toys and occupations. Children older than 5 years judged that rules pertaining to those activities could be altered and that the association of activities with gender could legitimately differ in societies other than their own. The Carter and Patterson's study also indicated that these patterns of rule flexibility and relativity were similar to children's judgments about social conventions, such as rules of etiquette (see also Damon, 1977; Turiel, 1978, 2002, 2006). This suggests that if children view gender norms as flexible conventions instead of inflexible obligations (Taylor et al., 2009), they may be more likely to see gender norms as guidelines that do not always need to be followed.

Recent research has examined whether children take a moral perspective on gender by assessing their judgments about exclusion based on gender stereotypic expectations. That research shows that in straightforward situations children judge social exclusion based on gender as wrong and give moral reasons of unfairness (Killen & Stangor, 2001; Theimer et al., 2001). However, in complex situations, conventional concerns about group cohesion and the attainment of goals for in-group activities are judged to justify exclusion based on gender norms. These distinctions are made at some ages and not others. Theimer et al. (2001) found that preschool age children judge exclusion from group activities based solely on gender as unfair (e.g., a group of girls playing with dolls refuse to let a boy join them). By contrast, exclusion based on characteristics that would interfere with the goals of the group is evaluated positively (e.g., to exclude certain girls from a baseball team because they are not good players), but only for older children (Killen & Stangor, 2001). As judgments about gender can involve both moral and conventional considerations, judgments in favor of gender norms may

not be an indication that the gender norm is endorsed unequivocally; the reasons for the endorsement must also be considered.

Therefore, previous research indicates that understandings of gender-related characteristics and activities are multifaceted and draw from different domains of judgment. The research also suggests that children draw on these domains differently as they develop. The present study was designed to examine children's reasoning regarding gender norms and their variations by systematically highlighting features from the personal, conventional, and moral domains. Whereas previous research on gender norms has assessed judgments from one domain or another, those studies do not provide a set of assessments of judgments in different domains.

As previous research has shown that children as young as 2½ years are aware of gender norms regarding toys, activities and clothing (Ruble et al., 2006), participants (ages 4–8 years) in this study were presented with hypothetical situations pertaining to masculine and feminine norms on these topics: attending classes (computer or babysitting classes), receiving toys (trucks or dolls), and wearing costumes for a party (soldier or ballet). They were asked questions aimed at assessing (a) their knowledge of traditional gender norms, (b) whether they judge gender norms as conventions and variable by cultural contexts, (c) whether they take into account the personal preferences of individual children, and (d) whether they judge imposed regulations on gender-related activities as morally fair or unfair.

To assess the children's knowledge of traditional or stereotypical gender norms, they were presented with a choice as to whether a parent should choose a gender-norm-consistent or -inconsistent child for each of the activities (e.g., Should a parent choose a son or a daughter to attend a babysitting class?). Framing the question in terms of the parent not only provided a forced choice but also encouraged the children to make the choice that they perceived that an authority should choose, and it provided a decision maker whose gender was not identified. We expected that children would use gender norms in the initial question about parents' choice because the information about the children in the story was limited to their gender. This expectation is consistent with previous research using other methods. For instance, Schuette and Killen (2010) found that most children use stereotypes to decide whether a boy or a girl should participate in a gendered chore.

To assess conventional aspects of gender norms, we examined whether children judge gender norms to be variable (thus assessing flexibility) by cultural contexts. They were asked to judge whether a reversal of traditional gender norms in a different culture was acceptable. In previous studies assessing generalizability or relativity of judgments about gender norms children have been asked questions that amounts to assessment of both informational knowledge and by implication, an evaluation: Could there be another country in which the gender norms are different (Smetana, 1986)? In this study, participants were asked directly to evaluate the practices of another country in which the gender norms are the reverse of this country. We expected that children would not judge the gender norms as universally imperative and, thus, would positively evaluate the reversal of norms in another country.

To ascertain whether children give validity to personal preferences in this realm, children were then presented with a situation where one child has expressed a preference for an activity that is inconsistent with the gender norm (e.g., If the son really wants to go to a babysitting classes, even more than the daughter, who should the parents choose?). Whereas previous research asked whether it is possible for a norm violation to occur with questions like "Can girls also play football?" (e.g., Levy et al., 1995), in this study children were presented with an individual preference that violates a norm and asked who should be chosen to engage in the activity. We expected that in those circumstances most children would judge that gender norms entail a choice, but given that past research has shown age-related changes in the awareness of the possibility for deviation (Levy et al., 1995), we also expected increases with age in choices that endorse norm deviations.

To assess children's possible moral perspectives on gender norms, we posed questions regarding the regulation of gender-related activities in a school or in another country. Past research (Helwig & Jasiobedzka, 2001) has found that in judging the legitimacy of legal regulation children consider multiple features of the situation, including whether the law has a socially beneficial purpose, and individual rights and autonomy. In the present study, we expected that children would negatively evaluate regulations that enforce gender norms on the basis of their concerns with the fairness of the regulation. We expected that these effects would be more pronounced for the older than the younger children.

Our general expectation regarding young children (4- to 5-year-olds) was that, despite the "essentialist" concepts of the sources of gender-related activities as biologically determined, which was found in some studies (Taylor et al., 2009), young children would show some flexibility when accounting for cultural contexts, personal preferences, and the moral implications of imposing rules on gender-related activities. We expected that older children would display knowledge about traditional gender roles and judge that parents should choose activities for children based on those norms. We also expected that older children would show flexibility (and more consistently than the younger children) with regard to cultural contexts, personal preferences, and moral implications.

The study also allows us to consider an issue that has emerged from findings of other studies, namely, that gender-role transgressions by males are judged more negatively than transgressions by females (Levy et al., 1995). Research by Blakemore (2003) indicates that this is especially the case for deviations from male clothing norms. Although this question was not a systematic part of our research design and we had no expectations about it, we were able to assess if male norms would be judged as less flexible than female norms.

## Method

### Participants

The participants were 72 children, 24 four-year-olds (13 girls, 11 boys; range = 3.6–5.6,  $M = 4.61$ ,  $SD = 0.64$ ), 24 six-year-olds (12 boys, 12 girls; range = 5.8–7.3;  $M = 6.43$ ,  $SD = 0.53$ ), and 24 eight-year-olds (12 boys, 12 girls; range = 7.11–10.0;  $M = 9.02$ ,  $SD = 0.52$ ). Participants were recruited through their preschool or elementary school in the San Francisco Bay area. Parents were told that the research was on how children think about gender norms. Most participants were middle to upper middle income, and they were a diverse sample (67% European American, 19% Asian American, 9% Latino, and 5% African American).

### Design and Procedures

Judgments about gender norms and fairness were assessed in an audiotaped interview conducted by a female interviewer that consisted of a series of hypothetical stories. Each story first described parents who had to choose either their son or daughter for an activity. "Parents" were

depicted as making the choice so that the decision maker would not be one sex or the other. There were three sets of stories involving (a) attending a class, (b) getting a toy, and (c) getting a costume for a Halloween party. Within each of the three sets there were two situations reflecting male or female norms (for a total of six stories). The first set of stories involved sending a child to a computer class (male norm) and a babysitting class (female norm), the second set involved giving a child a toy truck (male norm) and a doll (female norm), and the third set involved giving a child a soldier costume (male norm) and a ballet costume (female norm). These stories were selected because they are often associated with one sex or the other. Each participant heard all six stories, and boys and girls all heard about both male and female norms. The stories were presented in a random order.

*Assessments.* For each story six assessments were made.

1. To measure *judgments of the parents' choice* the children were asked was if they thought the parents should choose the son or daughter (e.g., "Who should the parents give the truck to, the daughter Elizabeth or the son Noah?"). This was followed by a question asking for their justifications (e.g., "Why should the parent choose the daughter [or the son]?")
2. *Knowledge of gender norms* was assessed with a question about which sex engages in the relevant activity more (e.g., "Do boys or girls usually babysit more?").
3. To assess *generalizability*, participants were asked whether or not it is acceptable that the gender norms be reversed in another country (e.g., "In another country, boys are usually the babysitters more than girls. Is that OK or not OK?").
4. *Judgments of nonnormative preferences* assessed whether participants saw the norm as alterable when a child had a nonnormative preference (e.g., "If the boy Joe loves babysitting, even more than the girl Sara, then who should go to the babysitting class?"). This question was followed by a justification question (e.g., "Why should they choose Joe?" or "Why should they choose Sara?").
5. *Rule legitimacy* assessed judgments about a rule enforcing gender norms at school (e.g., "The parents decide they want to send the boy Joe to the babysitting class. But they find out there is a rule at the school that only girls can be in the babysitters program. Is that rule

OK or not OK?"). This question was followed by a justification question ("Why is it OK?" or "Why isn't it OK?").

6. *Rule legitimacy in another country* assessed judgments of a rule enforcing gender norms in another country (e.g., "In another country they have the same rule that only girls can learn to be babysitters. Is that rule OK or not OK in the other country?"). Younger participants sometimes needed the phrase, "another country" to be explained, and "a place far, far away" was used.

*Coding and reliability.* Participants' responses included three possibilities (the boy, the girl or both, or OK, not OK or both). These were recoded to make them dichotomous for ease of analysis and because very few participants used the "both" response. Note that "both" was not a response that was offered, but children sometimes used it. The recoding was done conservatively, so that responses indicating "both" were collapsed into the response that was contrary to our hypotheses. More detail on how this was done for each assessment is provided in the following paragraphs.

Coding for the assessment called *judgments of parents' choice* reflected whether the participant's choice was based on gender norms (coded as 1), or a counter-gender norm choice or a choice of both the son and the daughter (coded as 0). *Knowledge of gender norms* was coded as 1, correct or consistent with the gender norm or 0, not consistent with the gender norm. Responses to the *generalizability* assessment were coded as 1, OK to reverse norms in another country or 0, which included responses of not OK and both OK and not OK. For the *judgments of nonnormative preferences* assessment, a response indicating that the one who likes it more should be chosen was coded as 1, and a response that the gender norm-consistent child or either child should be chosen was coded as 0. For both *rule legitimacy* and *rule legitimacy in a another country*, responses that the rule was not OK were coded as 1, and responses that the rule was OK or that the rule was both OK and not OK were coded as 0.

Justifications for *judgments of the parents' choice*, *judgments of nonnormative preferences*, and *rule legitimacy* were coded using categories derived from previous research (Davidson et al., 1983; Turiel, 1983) and adapted to responses from this study. The justification categories and brief examples of responses are listed in Table 1. The categories related to gender were grouped into a new category

called gender (including gender norms, gender-related preferences, and gender-related practices). Other categories were grouped into moral justifications (fairness/equity, and rights). Categories were also grouped into conventional justifications (culture and authority). The other justifications fell into personal choice and personal statement. An additional category for responses that were uncodable or for responses with no answer to the question was labeled unelaborated. Conventional and personal statement justifications were used in less than 10% of responses and thus were not included in analyses. Coding for justifications was 0, did not use the justification or 1, did use the justification. Up to two justifications were coded for each response, and when there were multiple justifications proportional coding was used.

The youngest children often did not provide any justifications. Specifically, the 4-year-olds gave justifications for an average of 46% of requests for justifications. Because of the low number of justification responses among the 4-year-olds, they were dropped from the analyses of justifications. Six- and 8-year-olds gave justifications more frequently, 74% and 86% of the time, respectively.

Three trained research assistants conducted the coding. Reliability was calculated based on coding of 15% of the interviews by all three coders. For evaluations, the Cohen's kappa was between .92 and .94. For justifications, the kappa was between .71 and .74.

## Results

Responses to the six questions were analyzed using a 3 (story type: classes, toys, costumes)  $\times$  2 (normative gender of story: female or male)  $\times$  3 (age: 4, 6, 8)  $\times$  2 (sex of participant) repeated measures analysis of variance (ANOVA), with story type and gender of story as repeated measures. Justifications with 10% or more of responses were also analyzed. Justifications were analyzed with repeated measures ANOVAs by justification category, story type, normative gender of story child, sex and child age, with justification category, story type and story gender as repeated measures. (ANOVA-based procedures were adopted because they have been shown to be appropriate for analyzing this type of data; Wainryb, Shaw, Laupa, & Smith, 2001). Significant main effects were analyzed using Bonferroni matched-pair *t* tests. Interactions were followed up with ANOVAs using restricted samples, and Bonferroni-corrected significance levels. One-sample

Table 1  
Justification Categories

Category	Justifications	Definitions and examples
Gender	Gender norms	References to traditional gender norms. References to the preferences of others for actions that are in accord with traditional roles. Preferences may also be based on observations of the frequency of actions by males and females <i>Examples:</i> Everyone wants girls to be babysitters. Boys do that more.
	Gender-related preferences	References to preferences that are inferred from traditional gender norms. The idea that preferences are in line with traditional norms <i>Examples:</i> She’s a girl so she’ll like the doll more. Boys don’t want that anyway.
	Gender-related capabilities	References to abilities that are inferred from traditional gender roles. The idea that someone would be more capable because of their gender <i>Example:</i> Girls are better with kids, more gentle.
Moral	Fairness/equity	A comparison implying equality or very similar treatment/opportunities/abilities. References to being the same, sharing, turn taking. Also includes assertions that the other sex should have a chance. Concerns that one sex will be excluded <i>Examples:</i> If they have that rule, then girls won’t fix computers. Girls and boys should both be able to go to a computer class. Boys should have a chance to try out babysitting.
	Rights	References to a right that applies to the whole gender or human category, use of the word “right(s)” <i>Example:</i> Girls have the right to go to a computer class.
Personal choice		Personal preferences. Individuals should be able to choose about this issue. The issue is not legitimately regulated <i>Examples:</i> It’s up the child to choose what they want to do. Everyone (even boys) likes dolls.
Conventions	Culture	Culture-specific norms that could change. Traditions that are tied to a specific location or context <i>Examples:</i> Boys probably do some things there and girls do different things there. Because they’re in a different culture.
	Authority	Authorities or consensus can make rules and they should be followed <i>Example:</i> It’s OK because it’s the rule.
Personal statement		A reference to the participant’s preferences or personal experiences or self interest <i>Examples:</i> I don’t like boy babysitters. My dad fixes computers, not my mom.
Unelaborated		Missing and uncodable responses

*t* tests were used to determine if responses differed from chance (.50).

*Knowledge of Gender Norms*

We begin with an analysis of the question assessing participants’ knowledge about the gender norms (e.g., Do boys or girls usually babysit more?). This assessment was administered after the first assessment but we describe the findings first to provide a sense of children’s knowledge about gender before moving to their judgments. We expected most participants to be aware of the gender norms and the majority of participants (90% overall) did give gender norm consistent responses. However, a

Story × Sex interaction,  $F(2, 126) = 3.94, p = .022, \eta = .06$ , indicated that while most participants of both sexes identified the traditional gender norms for toys (males:  $M = 0.94, SD = 0.05$ ; females:  $M = 0.96, SD = 0.04$ ) and costumes (males:  $M = 0.98, SD = 0.03$ , and females:  $M = 0.91, SD = 0.03$ ), females were more likely to say that the counter-gender norm children usually do the classes more (males:  $M = 0.97, SD = 0.05$ ; females:  $M = 0.79, SD = 0.04$ ). There were no age effects for this question, and *t* tests indicated that participants at all ages responded above chance for an average score over all stories on this question (4-year-olds:  $M = 0.81, SD = 0.24$ ; 6-year-olds:  $M = 0.95, SD = 0.09$ ; 8-year-olds:  $M = 0.97, SD = 0.06$ , all  $ps < .001$ ).

*Judgments of the Parents' Choice and Justifications*

Table 2 presents the proportions of participants' responses in each age group for each story type to the question of who the parents should choose. As can be seen in Table 2, the majorities in each age group made the gender norm-consistent response, as expected given that there was little other information to base a decision on. However, a story main effect,  $F(2, 120) = 9.54, p < .001, \eta = .14$ , indicated that participants were less likely to state that the parents should choose the norm-consistent choice for the stories about classes ( $M = 0.69, SD = 0.04$ ) than in either of the other types of stories (toys:  $M = 0.83, SD = 0.04$ ; costumes:  $M = 0.88, SD = 0.03$ ). The  $t$  tests for all ages combined revealed that responses to the computer class question were only marginally above chance,  $t(69) = 1.95, p = .055$ , although all other stories were significantly above chance ( $ps < .001$ ).

Justifications for these choices were primarily focused on gender norms. A main effect for

justification,  $F(3, 132) = 87.50, p < .001, \eta = .67$ , indicated that gender-related justifications were used significantly more than the other justifications ( $p < .001$ ), as Table 3 shows. A Justification  $\times$  Story Type interaction,  $F(6, 264) = 5.05, p = .001, \eta = .10$ , indicated that the gender norms justification,  $F(2, 88) = 7.50, p = .002, \eta = .15$ , was used less for the stories related to classes ( $M = 0.49, SD = 0.06$ ) than for costumes ( $M = 0.76, SD = 0.05, p = .002$ ).

*Generalizability*

To assess whether or not participants saw the gender norm as generalizable to another country, participants were asked to evaluate a practice in another country that is counter to traditional gender norms. Table 2 presents the proportion of responses for each age group by story type. Overall, most (79%) said it is all right to reverse the gender norm practice in another country, above chance for each story over all ages ( $ps < .001$ ). We

Table 2  
Mean Judgments (%) and Standard Deviations by Age

Story	4		6		8	
	M	SD	M	SD	M	SD
Parental choice should be norm consistent						
Class	0.69 (70)	0.06	0.74 (72)	0.06	0.65 (63)	0.06
Toy	0.87 (87)	0.06	0.87 (85)	0.06	0.76 (73)	0.07
Costume	0.89 (88)	0.06	0.93 (92)	0.06	0.81 (80)	0.06
Total	0.81 (82)	0.04	0.85 (83)	0.04	0.74 (72)	0.04
Norm can be reversed in another country						
Class	0.60 (61)	0.07	0.86 (85)	0.07	0.90 (92)	0.07
Toy	0.58 (60)	0.06	0.86 (83)	0.06	0.91 (92)	0.06
Costume	0.65 (65)	0.07	0.76 (77)	0.07	0.98 (98)	0.07
Total	0.61 <sub>a</sub> (62)	0.05	0.82 <sub>b</sub> (82)	0.05	0.93 <sub>b</sub> (94)	0.05
Choose counter stereotypical child who wants it more						
Class	0.78 (75)	0.06	0.92 (92)	0.05	0.98 (98)	0.05
Toy	0.73 (71)	0.06	0.93 (94)	0.05	0.95 (96)	0.05
Costume	0.68 (63)	0.06	0.92 (92)	0.06	0.92 (91)	0.06
Total	0.73 <sub>a</sub> (70)	0.05	0.92 <sub>b</sub> (92)	0.05	0.95 <sub>b</sub> (95)	0.05
Rule enforcing norm is not OK						
Class	0.44 (44)	0.07	0.75 (75)	0.07	0.98 (98)	0.07
Toy	0.63 (60)	0.07	0.83 (83)	0.06	1.00 (100)	0.07
Costume	0.60 (55)	0.07	0.75 (75)	0.07	0.93 (94)	0.07
Total	0.55 <sub>a</sub> (53)	0.06	0.78 <sub>b</sub> (78)	0.06	0.97 <sub>b</sub> (97)	0.06
Rule enforcing norm is not OK in another country						
Class	0.53 (44)	0.08	0.77 (75)	0.07	0.95 (96)	0.08
Toy	0.53 (42)	0.07	0.74 (72)	0.07	0.97 (96)	0.07
Costume	0.52 (51)	0.08	0.73 (74)	0.07	0.94 (96)	0.08
Total	0.53 <sub>a</sub> (43)	0.07	0.74 <sub>b</sub> (74)	0.06	0.96 <sub>c</sub> (96)	0.06

Notes. All means are proportions. Means in the same row with different subscript letters differ at  $p < .05$ .

expected children of all ages to judge gender norms as conventions that are alterable in another country. However, a main effect for age,  $F(2, 59) = 9.42, p < .001, \eta = .24$ , indicates that 4-year-olds were less likely than 6-year-olds ( $p < .01$ ) and 8-year-olds ( $p < .001$ ) to say it is all right to reverse the gender norm in another country. Whereas the majority (60%–65%) of the 4-year-olds stated it was all right to reverse the norm, these results were not above chance according to a  $t$  test ( $p = .126$ ). In an unexpected finding, an interaction between sex and the normative gender in the stories  $F(1, 118) = 5.26, p = .025, \eta = .08$ , indicated that females,  $F(1, 30) = 6.14, p = .019, \eta = .17$ , were more likely to judge as OK the reversal of a male norm in another country ( $M = 0.90, SD = 0.04$ ) than the reversal of a female norm in another country ( $M = 0.76, SD = 0.06$ ).

*Judgments of Nonnormative Preferences and Justifications*

The data in Table 2 display the findings for responses to the question of who the parents should choose when one child preferred the counternormative activity or toy or costume. Most participants (86%) judged that the child with the preference should be chosen. A main effect for age,  $F(2, 58) = 6.07, p = .004, \eta = .17$ , indicated that 4-year-olds were less likely than 6-year-olds ( $p < .05$ ) and 8-year-olds ( $p < .01$ ) to judge that parents should choose the countergender norm child if that child shows a preference. Still, as can be seen in Table 2, the majority of 4-year-olds did choose the countergender norm child, but to a lesser extent than the older children. The  $t$  tests showed that all ages responded above chance ( $p < .01$  for each age group).

A main effect for justification,  $F(3, 129) = 343.40, p < .001, \eta = .89$ , indicates that the 6- and

8-year-olds used the personal choice justification more frequently than the other justifications (see Table 3).

*Rule Legitimacy and Justifications*

In response to the question of whether a school can make a rule enforcing gender norms related to the activities, most (76%) said it is not all right to make that rule. We expected that most children would judge this type of rule to be unfair and that this tendency would increase with age. A main effect for age,  $F(2, 62) = 11.78, p < .001, \eta = .28$ , confirmed that 4-year-olds were less likely than 6-year-olds ( $p = .029$ ) and 8-year-olds ( $p < .001$ ) to negatively judge the school rule enforcing gender norms (see Table 2). The 4-year-olds did not perform above chance over all the stories, but the 6- and 8-year-olds did perform above chance ( $ps < .001$ ). A main effect for story type,  $F(2, 124) = 4.47, p = .013, \eta = .07$ , indicated that a school rule enforcing gender norms about bringing a toy to school ( $M = 0.82, SD = 0.04$ ) was judged not OK more than rules about which classes girls and boys can take ( $M = 0.72, SD = 0.04, p < .01$ ).

A main effect for justification,  $F(3, 129) = 50.51, p < .001, \eta = .54$  indicated that the 6- and 8-year-olds used moral justifications more than the other justifications ( $p < .001$ ). Table 3 shows the percentages of justifications used by these two age groups.

*Rule Legitimacy in Another Country*

Participants were also posed with the question of whether a school can make a rule enforcing gender norms related to the activities in the context of another country. A majority (76%) said the rule enforcing gender norms was not OK, even in another country, above chance for each story ( $ps < .01$ ). As expected, there was again a main

Table 3  
Percentages of Justifications Used by Question for 6- and 8-Year-Olds

Question	Gender-related norms		Personal choice		Moral		Unelaborated	
	6	8	6	8	6	8	6	8
Gender normative parental choice	64	62	5	9	4	6	20	19
Countergender norm child desiring act more	4	3	81	86	1	3	10	10
Rule enforcing gender norm	9	1	13	23	60	67	14	9

Note. Percentages may not add to 100 because coding categories were not included when they have less than 10% of responses.

effect for age,  $F(2, 53) = 11.63$ ,  $p < .001$ ,  $\eta = .31$ , which indicated that 4-year-olds were less likely than 6-year-olds ( $p = .044$ ) and 8-year-olds ( $p < .001$ ) to negatively evaluate the rule enforcing gender norms (both feminine and masculine) in another country. Six-year-olds were also less likely to negatively evaluate the enforcement of gender norms than were 8-year-olds ( $p = .044$ ). (See Table 2 for the findings presented by story type and age.) Like the previous question, these results were above chance for the 6- and 8-year-olds ( $ps < .001$ ), but the 4-year-olds did not perform above chance. There was also an Age  $\times$  Gender interaction  $F(2, 53) = 4.37$ ,  $p = .017$ ,  $\eta = .14$ , which indicated that 4-year-old girls were less likely to negatively evaluate the rule in another country than boys of the same age  $F(1, 16) = 12.06$ ,  $p = .003$ ,  $\eta = .43$  (significant after Bonferroni corrections). Among 4-year-olds, the mean proportions of judgments that the rule enforcing both feminine and masculine gender norms was not OK were .29 for girls ( $SD = 0.10$ ) and .76 for boys ( $SD = 0.08$ ).

### Discussion

It has been proposed that children's conceptions about gender norms shift with age from judgments that they are fixed or invariant to judgments that they are flexible and variable. For example, Ruble et al. (2006, p. 865) stated, "Until approximately 7–8 years of age . . . stereotypes are held rigidly, perhaps younger children do not seem to recognize that there can be individual variations in masculinity and femininity within the male and female categories." Some of the strongest assertions of this type have come from researchers who found that young children conceive of gender activities as fixed and uninfluenced by the environment, and that older children develop conceptions about variable environmental influences on such activities. However, judgments about the biological or environmental sources of behavioral characteristics tap into only one feature of thinking about gender. In this study, which was designed to further examine conceptions of gender norm flexibility when different domains are involved, we obtained findings showing that there are variations in how younger and older children think about gender. We found that younger children do accept that gender norms exist and should be followed. We also found this to be the case for older children, as well. When asked whether boys or girls usually engage in the activities, most at all ages gave responses consistent with

gender norms. The results show that gender norms have normative and binding characteristics in children's thinking since the majority also responded that parents should choose for their children the activities consistent with gender norms.

However, we also found that children were, for the most part, flexible about the application of gender norms under certain circumstances. Like the parents in Gelman's (2010) illustrative example (presented earlier) of how an issue (a child's desire to eat all her Halloween candy in one night) can be approached from the perspective of different domains, the children in this study applied gender norms differently for the different domains. We found that children judged gender norms to be alterable conventions in other cultures. We also found that personal preferences take precedence over the application of gender norms. Most children thought that parents should base their choices of a boy or girl for the activities on personal preferences. It was thought that a child expressing a preference for the nongender norm activity should be allowed to act on that preference. Note that the term *should* can have different meanings—moral and nonmoral. Although children responded that, for example, parents should choose children by gender norms (when preferences are not stated), they did not necessarily treat it as a moral should. This is shown by the findings that the children rejected the imposition of rules on gender-related activities in a school, and gave moral reasons for doing so. The most direct evidence of the concern with fairness comes from the older children, from whom we were able to elicit reasons for their rejection of the imposition of rules. The majority of the children in the two older groups judged enforcing gender norms with a school rule as unfair for each type of activity and gave justifications were based mainly on moral considerations. The most frequently stated reasons for the negative evaluations pertained to equal treatment and fairness. These findings are consistent with findings in research showing that children judge exclusion based on gender (or race) to be unfair (Killen & Stangor, 2001; Theimer et al., 2001). Unlike previous research, these findings indicate that even young children are concerned with fairness in gender activities in some situations.

The findings of this study also indicate that gender norms are not viewed as fixed at young ages, as some have suggested (Levy et al., 1995). In this study, children were asked about the acceptability of the violation in a different context, as opposed to studies asking about the possibility of violations

(e.g., Levy et al., 1995). A majority of the children, including the 4-year-olds (though fewer than the older children), accepted violations when they were presented to them. A separate question that was not addressed here is whether children are aware that others have preferences that are inconsistent with norms. Still, there were age differences that are consistent with an increase in flexibility with age. More of the 6- and 8-year-olds judged a reversal of the gender norms in another country to be acceptable than did the 4-year-olds. It is important to note that while flexibility did increase with age, the majority of the 4-year-olds (60%–65%) still judged that a reversal of the gender norms in another country was acceptable. Although this finding was not above chance (50%) for the 4-year-olds, gender norms are not uniformly judged as inflexible at this age.

A seeming anomaly in the findings regarding judgments about alternative norms in another country is that female participants (90% of them) were more likely to positively judge a reversal of male norms in another country (e.g., stating that it is OK if boys play with dolls more than girls) than female norms (e.g., stating that it is OK if girls play with trucks more than boys; 76%). The reason for this finding is unclear. However, it is possible that females focus on the norms that affect them most directly more than male norms.

Similar results of flexibility were obtained in judgments about the personal preferences of individuals. Participants were posed with the question of whom the parents should choose if one child had a greater preference for activities that went against the gender norms. Again, children in the two older groups were more likely to make a counter-gender norm choice under those circumstances (92%–98%) than were those in the youngest group. Nevertheless, the majority (63%–75%) of the 4-year-olds did think the parents should choose on the basis of the children's preferences, and they performed above chance on this measure. The reasons for the judgments by the two older groups were clear-cut. The prevalent justifications for these choices were based on the idea that a child should be able to choose desired activities (the personal choice justification).

It appears, therefore, that gender norms are not uniformly judged as inflexible even at the youngest ages represented in this study. Still, the findings were not as clear-cut for the 4-year-olds as for the older children. These mixed results with the youngest children might reflect differences in developmental levels among them—which may be due in

part to their understandings of minds (theory of mind). It may be that some of the 4-year-olds did not accept that a child could actually have a counterstereotypical preference. Although children younger than the participants in this study usually have an understanding of diverse preferences, some research (Terwogt & Rieffe, 2003) shows that understandings of others' preferences may be more difficult when those preferences are counter to gender norms. If some participants did not accept that children sometimes have counternorm preferences, it would impact judgments about a counterstereotypical choice.

The findings that most children accept the legitimacy of personal choices in the realm of gender norms indicate that adherence to gender norms is not judged to be morally obligatory. Children's moral judgments are not based on personal choices (Turiel, 1983, 2002). However, children do judge restrictions on activities considered to be legitimately in the personal realm as sometimes unwarranted (Nucci, 2001). In the present study, restrictions of personal choice regarding gender norms were judged to be unfair in the two questions inquiring as to whether gender norms should be enforced by rules in a school and in another country. Previous research has shown that children use gender norms to make decisions when there is little other information but they are also increasingly aware of and concerned with gender equity. When Schuette and Killen (2010) probed children 5–10 years of age on stereotypic choices by suggesting that children who had not had a chance to try a stereotypical chore should be given the opportunity, the older children were more likely to show flexibility by endorsing the gender norm inconsistent child. The present study also obtained age differences. The youngest children were more accepting of rules enforcing gender norms than the older groups. However, the majority of the 4-year-olds (55%) negatively evaluated rules enforcing gender norms. While this percentage was not above chance, it indicates a trend toward rejection of the enforcement of gender norms.

These findings were not consistent across all the types of stories. Children at all ages evaluated the rule enforcing gender norms for enrollment of girls in the babysitter class and boys in the computer class less negatively than in the stories about costumes and toys, and descriptive statistics show that this was particularly true of the 4-year-olds, who often (44% of the time) positively evaluated a school rule dictating which classes boys and girls could take. One possible interpretation of these

findings is that the young children discriminated between regulations they saw as within the jurisdiction of a school (classes) and those outside its jurisdiction (Laupa & Turiel, 1993). Another possibility is that if children believe that no preferences are counternormative (Terwogt & Rieffe, 2003), rules enforcing gender norms might not seem unfair. In the absence of data on justifications from the 4-year-olds for these evaluations, we cannot be certain of this interpretation.

An additional feature of the findings of this study is that children often considered the specific context of each story, sometimes judging the different stories differently. It was found that children at all ages assumed that the activities of boys and girls differ with regard to toys and costumes. It was thought that boys would play with a fire truck and wear soldier costumes more than girls; it was thought that girls would play with a baby doll and wear a ballet costume more than boys. By contrast, the female participants did not make as clear distinctions between babysitting and computer class activities. The females assumed that girls engage in the nongender norm activity more than did the male participants. It appears that the female participants at all ages thought that girls also take computer classes. One plausible interpretation of these findings is that girls are more aware than boys of the interests that girls have in computers and, perhaps, of the opportunities girls have to use computers. If this interpretation were correct, it would mean that with regard to certain activities (but not all) boys think along gender-related norms more than girls. Another explanation for this finding is that girls may be more likely to engage in male gender norms activities than the reverse, since girls would be seen as moving up the hierarchy, but boys would prefer not to move down in a hierarchy (Horn, 2003, 2006). If this is so, computers may be an area where girls are challenging traditional norms, and perhaps girls themselves would be the first to notice it.

A difference among activities, consistent with our interpretation, was also obtained in responses to the question of which child the parent should choose. For costumes and toys, most made the gender-consistent choice. However, participants were less likely to choose the gender-consistent choice for both the babysitting and computer classes. As would be expected from the judgments of whom the parents should choose, the 6- and 8-year-olds used the gender norm justification less for the stories related to babysitting and computer classes. More research is needed to determine why the sto-

ries on classes were judged to be less influenced by gender norms, but one possible explanation is that children judge classes to have higher stakes, since they could affect future opportunities. Still, our findings are in line with research by Schuette and Killen (2010) showing that children often make judgments in line with gender norms, and that as they get older they are more likely to consider issues of fairness, especially when those considerations are made salient.

Overall, the findings of this research demonstrate that children do not judge gender norms to be consistently fixed or morally obligatory. By the ages of 6–8 years children understand the uniformity of gender-related activities—as shown by their assumptions about the likelihood of activities differently associated with males and females and their judgments that parents should choose activities for their children along those lines. However, they do not regard those activities to be obligatory for boys and girls in generalizable ways. They accept the possibility of a reversal of the association of the activities with gender in another country. They also take personal choices into account as reasons for not adhering to gender norms. Insofar as the older children invoked moral obligations it was to reject the regulation of gender-related activities—which is contrary to the view that children see gender norms as so inflexible as to be moral (Taylor et al., 2009).

As we have seen, the evidence for the older children supports these conclusions. The findings for the youngest children in this study were not as clear-cut. On the one hand, the majority of the youngest children responded like the older ones. On the other hand, a larger proportion of younger than older children treated gender norms as fixed in that they negatively evaluated the reversal of gender norms in another country, did not accept personal choices as the basis for parental choices, and positively evaluated regulations pertaining to gender norms. As discussed earlier, this could be the result of an incomplete understanding of non-normative preferences (Terwogt & Rieffe, 2003), differences in beliefs about the jurisdiction of school authorities (Laupa & Turiel, 1993), or developmental changes in the understanding of the category of gender (Levy et al., 1995). More research is necessary to investigate the individual differences found among the youngest children. If possible, it would also be useful to formulate procedures for obtaining data on young children's justifications. It may also be that children's social contexts have a bearing on how they construe norms associated with gender. As an example, Rhodes and Gelman (2009) found

differences in gender categorizations in children from politically conservative and liberal areas in the United States. It may be that people in politically conservative communities are less accepting of variations in traditional gender norms than are those from politically liberal communities (Horn, 2008). Although we did not assess variables like political orientation in this study, it could be said, in general terms, that participants were from a politically liberal part of the country. However, the results suggest that political affiliation had minimal influence on the children's judgments. This is because the results paralleled findings from a number of other studies conducted in several regions of the country (Carter & Patterson, 1982; Smetana, 1986). As in those studies, we found both that the youngest children were less accepting of variations in gender norms than the older ones and that the older children judged gender norms as flexible. Nevertheless, it would be worthwhile in future research to use our assessments in comparisons of individuals with different religious and political orientations. Such research should examine commonalities and differences in social construals. Another way that social contexts may intersect with judgments stems from labeling of groups, which appear to increase in-group bias (Patterson & Bigler, 2006). Therefore, future research should also examine how children construe social groupings and status differences in their communities.

The research we have reported supports previous findings that young children consider multiple factors when reasoning about gender (Killen, Sinno, & Margie, 2007) and rules (Helwig & Jasiobedzka, 2001). When particular facets of gender are highlighted, including the alterability of conventions and personal preferences, children are flexible even at a young age, and indeed, they can be critical of rules that codify gender norms. Although much previous work has focused on the development of inflexibility in gender norms, future work should also examine the ways that children are flexible in terms of gender, perhaps studying social influences on flexibility about gender.

## References

- Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex Roles, 48*, 411-419. doi:0360-0025/03/0500-0411/0
- Carter, D. C., & Patterson, C. J. (1982). Sex roles as social conventions: The development of children's conceptions of sex-role stereotypes. *Developmental Psychology, 18*, 812-824. doi:10.1037/0012-1649.18.6.812
- Damon, W. (1977). *The social world of the child*. San Francisco: Jossey-Bass.
- Davidson, P., Turiel, E., & Black, A. (1983). The effect of stimulus familiarity on the use of criteria and justifications in children's social reasoning. *British Journal of Developmental Psychology, 1*, 49-65.
- Gelman, S. (2010). Modules, theories, or islands of expertise? Domain specificity in socialization. *Child Development, 81*, 715-719. doi:10.1111/j.1467-8624.2010.01428.x
- Helwig, C. C., & Jasiobedzka, U. (2001). The relation between law and morality: Children's reasoning about socially beneficial and unjust laws. *Child Development, 72*, 1382-1393. doi:10.1111/1467-8624.00354
- Horn, S. S. (2003). Adolescents' reasoning about exclusion from social groups. *Developmental Psychology, 39*, 71-84. doi:10.1037/0012-1649.39.1.71
- Horn, S. S. (2006). Group status, group bias, and adolescents' reasoning about the treatment of others in school contexts. *International Journal of Behavioral Development, 30*, 208-218. doi:10.1177/0165025406066721
- Horn, S. S. (2008). The multifaceted nature of sexual prejudice: How adolescents reason about sexual orientation and sexual prejudice. In S. R. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 66-86). New York: Oxford University Press.
- Killen, M., Sinno, S., & Margie, N. (2007). Children's experiences and judgments about group exclusion and inclusion. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 35, pp. 173-218). San Diego, CA: Elsevier.
- Killen, M., & Smetana, J. (Eds.). (2006). *Handbook of moral development*. Mahwah, NJ: Erlbaum.
- 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. doi:10.1111/1467-8624.00272
- Laupa, M., & Turiel, E. (1993). Children's concepts of authority and social contexts. *Journal of Educational Psychology, 85*, 191-197. doi:10.1037/0022-0663.85.1.191
- 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. doi:10.2307/1131594
- Nucci, L. P. (2001). *Education in the moral domain*. New York: Cambridge University Press.
- Patterson, M. M., & Bigler, R. S. (2006). Preschool children's attention to environmental messages about groups: Social categorization and the origins of intergroup bias. *Child Development, 77*, 433-444. doi:10.1111/j.1467-8624.2006.00906.x
- Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology, 59*, 244-274.
- Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In N. Eisenberg, W. Damon, &

- R. Lerner (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 858–932). Hoboken, NJ: Wiley.
- Schuette, C., & Killen, M. (2010). Children's evaluations of gender-stereotypic household activities in the family context. *Early Education and Development, 20*, 693–712. doi:10.1080/10409280802206908
- Smetana, J. (1986). Preschool children's conceptions of sex-role transgressions. *Child Development, 57*, 862–871. doi:10.2307/1130363
- Smetana, J. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 119–153). Mahwah, NJ: Erlbaum.
- Smetana, J., & Braeges, J. (1990). The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly, 36*, 329–346.
- Stoddart, T., & Turiel, E. (1985). Children's conceptions of cross-gender activities. *Child Development, 56*, 1241–1252. doi:10.2307/1130239
- Taylor, M. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development, 67*, 1555–1571. doi:0009-3920/96/6704-0017S01.00
- 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. doi:10.1111/j.1467-8624.2009.01272.x
- Terwogt, M., & Rieffe, C. (2003). Stereotyped beliefs about desirability: Implications for characterizing the child's theory of mind. *New Ideas in Psychology, 21*, 69–84. doi:10.1016/S0732-118X(02)00031-4
- Theimer, C. E., Killen, M., & Stangor, C. (2001). Young children's evaluations of exclusion in gender-stereotypic peer contexts. *Developmental Psychology, 37*, 18–27. doi:10.1037//0012-1649.37.1.18
- Trautner, H. M., Ruble, D. N., Cyphers, L., Kirsten, B., Behrendt, R., & Hartmann, P. (2005). Rigidity and flexibility of gender stereotypes in childhood: Developmental or differential? *Infant and Child Development, 14*, 365–381. doi:10.1002/icd.399
- Turiel, E. (1978). The development of concepts of social structure: Social convention. In J. Glick & A. Clarke-Stewart (Eds.), *The development of social understanding* (pp. 25–107). New York: Gardner.
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.
- Turiel, E. (2002). *The culture of morality: Social development, context, and conflict*. Cambridge, UK: Cambridge University Press.
- Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 119–153). Mahwah, NJ: Erlbaum.
- Turiel, E. (2010). The development of morality: Reasoning, emotions, and resistance. In R. M. Lerner (Series Ed.) & W. Overton (Vol. Ed.), *Handbook of life-span human development: Vol. 1. Cognition, biology, and methods* (pp. 554–583). New York: Wiley.
- Wainryb, C., Shaw, L. A., Laupa, M., & Smith, K. R. (2001). Children's, adolescents' and young adults' thinking about different types of disagreements. *Developmental Psychology, 37*, 373–386. doi:10.1037/0012-1649.37.3.373

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.