Journal of Experimental Child Psychology 158 (2017) 19–31

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Group presence, category labels, and generic statements influence children to treat descriptive group regularities as prescriptive

Steven O. Roberts *, Arnold K. Ho, Susan A. Gelman

Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA

ARTICLE INFO

Article history:
Received 19 July 2016
Revised 23 November 2016
Available online 4 February 2017

Keywords:
Social cognitive development
Normative reasoning
Category labels
Generics
Novel groups
Conformity

ABSTRACT

Children use descriptive regularities of social groups (what *is*) to generate prescriptive judgments (what *should be*). We examined whether this tendency held when the regularities were introduced through group presence, category labels, or generic statements. Children (ages 4–9 years, $N = 203$) were randomly assigned to one of four conditions that manipulated how descriptive group regularities were presented: group presence (e.g., "These ones [a group of three individuals] eat this kind of berry"), category labels (e.g., "This [individual] Hibble eats this kind of berry"), generic statements (e.g., [showing an individual] "Hibbles eat this kind of berry"), or control (e.g., "This one [individual] eats this kind of berry"). Then, children saw conforming and non-conforming individuals and were asked to evaluate their behavior. As predicted, children evaluated non-conformity negatively in all conditions except the control condition. Together, these results suggest that minimal perceptual and linguistic cues provoke children to treat social groups as having normative force.

Introduction

Children readily adopt a normative stance (Rakoczy & Schmidt, 2013). They interpret another person's behavior as normative (e.g., "That is how one *should* do that") even in non-normative contexts

* Corresponding author.
E-mail address: sothello@umich.edu (S.O. Roberts).

http://dx.doi.org/10.1016/j.jecp.2016.11.013
0022-0965/

(Schmidt, Rakoczy, & Tomasello, 2011) and even after seeing the behavior only once (Schmidt, Butler, Heinz, & Tomasello, 2016). When children observe a tool being used in a certain way, they rigidly imitate the behavior, expect others to do so as well, and protest when the tool is used differently (e.g., "You *must* do this"; Kenward, 2012). Similarly, when children learn the rules of a game and subsequently observe someone violate those rules, they respond with protest and critique (e.g., "You *can't* do that"; Rakoczy, Hamann, Warneken, & Tomasello, 2010; Rakoczy & Schmidt, 2013). Indeed, children's sensitivity to norms is central to their social cognition; it tells them what kind of behaviors to expect from others (e.g., children expect group members to share properties; Kalish, 2012), guides their own behavior (e.g., children spontaneously create their own norms and teach them to others; Göckeritz, Schmidt, & Tomasello, 2014), shapes how they evaluate and respond to norm violators (e.g., they critique norm violators and are annoyed by them; Cooley & Killen, 2015; Hardecker, Schmidt, Roden, & Tomasello, 2016), and in some cases can even influence them to act antisocially (e.g., if their group members do not share, they might not share either; Engelmann, Herrmann, Rapp, & Tomasello, 2016).

Recently, Roberts, Gelman, and Ho (2016) found that children use descriptive group regularities (what *is*) to make prescriptive judgments (what *should be*). They introduced children (ages 4–13 years) to two novel groups, Hibbles and Glerks, that engaged in morally neutral behaviors (e.g., spoke a certain language, ate a certain food) and then to a series of conforming or non-conforming individuals. Children, especially the youngest (4- to 6-year-olds), evaluated non-conforming individuals negatively (e.g., if Hibbles listen to a certain kind of music or speak a certain language, then it was bad for an individual Hibble to listen to a different kind of music or speak a different language). They also found that children's responses were robust across intergroup contexts (i.e., children disapproved of non-conformity regardless of whether the novel groups were portrayed as cooperating with or competing against each other. This research provided a strong test of the power of norms; the groups were unfamiliar, the behaviors were morally neutral, and children did not belong to the groups—all factors that have been previously demonstrated to license prescriptive judgments (e.g., Abrams, Rutland, & Cameron, 2003; Blakemore, 2003; Liben, Bigler, & Krogh, 2001; Mulvey, 2016; Smetana et al., 2012)—yet children nonetheless made prescriptive judgments when faced with non-conforming individuals.

An important open question is what information signals to children that a feature is group relevant and thus normative. Roberts and colleagues (2016) provided children with several converging cues; the groups consisted of individuals in sets of three (thereby highlighting group presence), they were labeled with a common noun (i.e., Hibble or Glerk), and they were described with generic statements in which properties were attributed to categories (e.g., "Hibbles eat these kinds of berries"). When these cues were present, children made prescriptive judgments; when they were absent, children did not. Thus, although children used group regularities to generate prescriptive judgments, it remains unclear which cues or combination of cues shift a behavior from being a descriptive "is" to being a prescriptive "should." That is, are prescriptive judgments elicited by group presence, category labels, generic statements, or some combination of these three factors?

Seeing a group of individuals engage in a common behavior has profound effects on our social cognition. Take as an example Asch's (1955) classic research on social pressure (i.e., normative social influence); individuals confronted with three or more people who share a belief feel the pressure to also hold that belief even if they suspect the belief to be false. Thus, because the presence of a group is a strong predictor of conformity, one possibility is that simply seeing a group of individuals who share a common behavior will generate prescriptiveness (i.e., all individual group members *should* share that behavior). Indeed, recent research suggests that being exposed to a group of individuals who share common properties increases the likelihood of that group being perceived as coalitional as well as the likelihood of stereotyping individuals within that group (e.g., Bigler, Jones, & Lobliner, 1997; Cosmides, Tooby, & Kurzban, 2003).

Group labels (i.e., count nouns; Macnamara, 1982), even when only an individual group member is present, play a critical role in shaping young children's categorization. Labels make categories more salient, encourage children to form new categories and treat them as stable, and promote category-based inferences (Baron, Dunham, Banaji, & Carey, 2014; Gelman & Markman, 1986; Graham, Keates, Vukatana, & Khu, 2013; Waxman & Markow, 1995). For example, Waxman (2010)

introduced children to individuals (e.g., Black woman, White man), provided novel properties about those individuals (e.g., "is good at a game called zaggit"), and assessed whether participants expected newly encountered individuals to share the novel property. When individuals received category labels (e.g., "This one is a *Wayshan*"), children expected newly encountered same-race and same-gender individuals to share the same property, whereas when individuals were not labeled (e.g., "This one *eats big lunches*"), they did not expect newly encountered same-race or same-gender individuals to share the same property. Thus, labels facilitated group-based inferences. However, this finding does not speak to whether labels also generate children's prescriptive judgments toward novel groups.

Labels may license stronger inferences when they are coupled with generic statements (e.g., expressions that refer to categories). For example, the generic statement "Hibbles eat this kind of berry" may license stronger inferences than the specific statement "These Hibbles eat this kind of berry" because the generic implies that the category is closely linked to the property, general across time and contexts, and indicative of commonalities shared among individual group members (Cimpian & Erickson, 2012; Gelman, Ware, & Kleinberg, 2010; Graham, Nayer, & Gelman, 2011; Prasada, 2000). Indeed, when 4-year-olds were presented with a completely novel category (i.e., Zarpies), they were more likely to make category-based inferences when Zarpies were introduced via generics (e.g., "Zarpies like to sing") than when they were introduced via specific labels (e.g., "This Zarpie likes to sing") or no labels (e.g., "This likes to sing"; Gelman et al., 2010; Rhodes, Leslie, & Tworek, 2012), demonstrating that generic statements, compared with labels alone, generate more category-based inferences and suggesting that they may likewise generate more prescriptive judgments.

The current research tested the extent to which group presence, category labels, and generic statements bolstered children's prescriptive judgments of group regularities. To do this, we presented children with two kinds of groups and randomly assigned children to one of four conditions in which we manipulated how the group regularities were conveyed: (a) group presence (each group included three individuals who were described without labels or generic statements), (b) labels (each group included one individual who was described with a category label but not a generic statement), (c) generic (each group included one individual who was described with a label and a generic statement), or (d) control (each group included one individual who was described without a label or generic statement). Because we wanted group membership to be apparent in all conditions, members of different groups were spatially segregated and wore distinct clothing patterns. Prior research indicates that such cues by themselves do not foster prescriptive judgments (Roberts et al., 2016, Study 2). We predicted that (a) children would use group presence, category labels, and generic statements to make prescriptive judgments; (b) children would make fewer prescriptive judgments when not provided with these cues (thereby replicating Roberts et al., 2016); and (c) children would make more prescriptive judgments when hearing generic statements than when hearing only labels. No additional a priori predictions were made regarding the relative effects of group presence, labels, and generic statements. We focused on the age groups of 4 to 6 years and 7 to 9 years given that Roberts and colleagues (2016) found that this was when important developmental changes occurred on this task and that other research shows captures significant changes in children's social category concepts (Quintana, 1998; Rhodes & Gelman, 2009; Roberts & Gelman, 2015, 2016, 2017).

Method

Participants

Two age groups of children were included ($N = 203$): 4- to 6-year-olds ($n = 106$, 51% female, $M_{age} = 5$ years 6 months, range = 4 years 1 month to 6 years 11 months) and 7- to 9-year-olds ($n = 97$, 61% female, $M_{age} = 8$ years 4 months, range = 7 years 0 months to 9 years 11 months). All children were recruited in the U.S. Midwest at two university-affiliated museums. The sample was mostly White/European American (60% White/European American, 12% Asian American/Asian/Pacific Islander, 10% multiracial, 5% Black/African American, 3% Hispanic/Latino, and 12% other or not reported).

Materials and procedure

The task was adapted from Roberts et al. (2016) and presented via Qualtrics. Children were randomly assigned to one of four conditions: *group presence* (i.e., sets of three individuals per group, spatially segregated, distinguished by clothing patterns, and presented without category labels or generic statements), *label* (one individual per group, spatially segregated, distinguished by clothing patterns, and presented with category labels), *generic* (one individual per group, spatially segregated, distinguished by clothing patterns, and presented with category labels and generic statements), and *control* (one individual per group, spatially segregated, distinguished by clothing patterns, and presented without labels or generic statements). For sample trials, see Figs. S1 and S2 in the online supplementary material.

Group presence

Children (4–6 years: $n = 25$; 7–9 years: $n = 24$) were introduced to two novel groups that were presented without category labels or generic statements: "I'm going to tell you about these ones [*pointing*] and these ones [*pointing*]." Each group consisted of three individuals (one set of three located on the left side of the screen and another set located on the right side of the screen), with group membership portrayed by clothing pattern (i.e., orange rectangles vs. green stripes). Next, children received eight test trials. Across all conditions, there were four behavioral domains—Food, Games, Language, and Music—and each behavior matched its corresponding group in color (e.g., green clothing pattern corresponded with a green musical note). On each trial, children were given regularities for both groups and then shown a conforming or non-conforming individual. For example, "These ones [*pointing to the group with orange rectangles*] eat this kind of berry, and these ones [*pointing to the group with green stripes*] eat this kind of berry. Look [*revealing and pointing to an individual who either conformed or did not conform to the group*], this one is eating this kind of berry."

Label

Children (4–6 years: $n = 31$; 7–9 years: $n = 24$) were introduced to two individuals who were spatially segregated, distinguished by clothing pattern (identical clothing to those in the group presence condition), and presented with contrasting category labels: "I'm going to tell you about these two. This one is a Hibble [*pointing*], and this one is a Glerk [*pointing*]." Next, children were told a fact about each individual and then introduced to other individuals who either conformed or did not conform to the property exhibited by the individual with the same name (eight trials). For example, "This Hibble eats this kind of berry [*pointing*], and this Glerk eats this kind of berry [*pointing*]. Look [*revealing and pointing to an individual who either conformed or did not conform to the labeled group*], this Hibble is eating this kind of berry [*pointing*]."

Generic

As in the label condition, children (4–6 years: $n = 26$; 7–9 years: $n = 25$) were introduced to two individuals who were spatially segregated and distinguished by clothing pattern: "I'm going to tell you about these two. This one is a Hibble [*pointing*], and this one is a Glerk [*pointing*]." Next, children were told a fact about each individual and then introduced to conforming or non-conforming individuals (eight trials). Unlike the label condition, however, the initial facts were provided in the form of generic statements. For example, "Hibbles eat this kind of berry [*pointing*], and Glerks eat this kind of berry [*pointing*]. Look [*revealing and pointing to an individual who either conformed or did not conform to the generically referenced group*], this Hibble is eating this kind of berry [*pointing*]."

Control

Children (4–6 years: $n = 24$; 7–9 years: $n = 24$) were introduced to two individuals who were spatially segregated, distinguished by clothing pattern, but presented without category labels or generic statements (see also Roberts et al., 2016, Study 2): "I'm going to tell you about these two—this one [*pointing*] and this one [*pointing*]." Children were then told a fact about each individual and introduced to conforming and non-conforming individuals. For example, "This one eats this kind of berry [*pointing*], and this one eats this kind of berry [*pointing*]. Look [*revealing and pointing to an individual who*

either conformed or did not conform to the other individual with the same appearance], this one is eating this kind of berry [*pointing*]."

Across all conditions, each participant saw eight trials. Four trials depicted a conforming individual, and four trials depicted a non-conforming individual. Trials were presented in random order, and across participants the left–right position of the groups/individuals was counterbalanced. In the label and generic conditions, we also counterbalanced which label was associated with which clothing pattern.

Measures and coding

First, children were asked whether or not the behaviors of the conforming/non-conforming individuals were "okay" or "not okay" (evaluation: e.g., "Is it okay or not okay for this Hibble to eat this kind of berry?"). We calculated the frequency with which these evaluations occurred for both conformity and non-conformity trials (scores for each could range from 0 to 4), focusing on the frequency of "not okay" evaluations, which reflected disapproval, as the dependent variable. (Note that the frequencies of "okay" and "not okay" responses were precise inverses.)

Second, children were asked how bad or how good a specific behavior was. Children who evaluated behaviors as "not okay" were presented with a scale of three increasingly unhappy faces and were asked, "Is it a little bad, pretty bad, or very, very bad?" (1 = *a little bad*, 2 = *pretty bad*, 3 = *very*, *very bad*), and children who evaluated behaviors as "okay" were presented with a scale of three increasingly happy faces and were asked, "Is it a little good, pretty good, or very, very good?" (1 = *a little good*, 2 = *pretty good*, 3 = *very*, *very good*).

Third, children were asked to explain their evaluations (e.g., "Why is it not okay for this one to eat this kind of berry?"). Responses were coded into five types based on previous research (Rhodes, 2014; Roberts et al., 2016): (a) norm-based, (b) group-based, (c) individual-based, (d) similarity-based, and (e) other (see Table 1 for a description of the coding scheme). Codes were not mutually exclusive, so children could appeal to multiple explanation types within a single response. These responses were coded by two research assistants who were blind to hypotheses of the study (Cohen's kappa = .77), and disagreements were resolved by discussion. We calculated the percentage of times that each type of explanation was provided, out of the total number of trials, for each response type (e.g., a child who disapproved of non-conformity on four trials and gave a norm-based explanation on one of those trials was coded as having given norm-based explanations for disapproved non-conformity 25% of the time).

As a comprehension check, at the end of the task, children were asked, "What does it mean for something to be not okay?" (open-ended) and "Does 'not okay' mean that someone should or should

Table 1
Description of the coding scheme.

Code	Description	Examples
Norm	Explicitly mentions that there is a rule or an obligation to which the individual must adhere	"They are supposed to play with that kind of toy." "They have to listen to that kind of music." "They aren't allowed to do that."
Group	References the groups or the category labels	"Because that's what Glerks do." "Because it's a Hibble." "Because that's what the rest of the group is doing."
Individual	References mental states, including thoughts, emotions, motivations, and traits. States that the behavior is about personal choice rather than group membership	"They can do whatever they want." "They can do that if they like it or want to." "Different people like different things."
Similarity	Mentions similar or dissimilar physical appearances	"Because they are orange." "Because it looks like that one." "Green goes with green."
Other	Gives an explanation that does not fit the others' codes	"I don't like them." "It is weird."

not do something?" (0 = *should*, 1 = *should not*). Only children who indicated norm-based reasoning in their open-ended question (e.g., "You aren't supposed/allowed to do it," "It means something bad will happen") and said that "not okay" means that someone "should not" do something were included in the final sample. Seven children did not meet these criteria and therefore were excluded from the final sample.

For exploratory purposes, parents were also given the option of completing a survey (81% response rate) adapted from Feldman (2003) that assessed their views on authoritarian parenting ($\alpha = .51$), conformity ($\alpha = .41$), and respect for common norms ($\alpha = .61$). However, these measures yielded low reliabilities and were not related with children's responses and therefore are not reported further.

Results

There were no significant effects of any of the counterbalancing factors or of the behavioral domains, so the data were collapsed across these variables. All significant effects at the $p < .05$ level were followed by Bonferroni-corrected pairwise comparisons.

Disapproval toward non-conformity and conformity

We first tested whether children were more disapproving of non-conformity than of conformity and whether their rates of disapproval varied across conditions. To test this, we conducted a 2 (Age Group: 4–6 or 7–9 years) × 2 (Behavior: conformity or non-conformity) × 4 (Condition: group presence, label, generic, or control) repeated measures analysis of variance (ANOVA) with behavior as a within-participants variable, age group and condition as between-participants variables, and the frequency of "not okay" evaluations as the dependent variable. A main effect of behavior, $F(1,195) = 139.83$, $p < .001$, $\eta_p^2 = .42$, showed that non-conformity was indeed disapproved of more frequently than conformity, and a main effect of age group, $F(1,195) = 10.23$, $p = .002$, $\eta_p^2 = .05$, showed that 4- to 6-year-olds were more disapproving than 7- to 9-year-olds. There was also a main effect of condition, $F(3,195) = 4.93$, $p = .003$, $\eta_p^2 = .07$. Planned comparisons showed that, compared with the control condition, disapproval was higher in the generic condition ($p = .001$). There was also a significant interaction of behavior and condition, $F(3,195) = 8.34$, $p < .001$, $\eta_p^2 = .11$. Planned comparisons revealed that non-conformity was disapproved of more frequently than conformity in the group presence, label, and generic conditions ($ps < .001$) but not in the control condition. For conformity, disapproval did not differ significantly between any of the conditions ($ps = 1.00$). For non-conformity, disapproval was

Table 2
Means, standard errors, and one-sample *t*-test statistics comparing average disapproval frequencies against chance (i.e., 2) across age groups, conditions, and behavior type.

Age (years)	Condition	Behavior	*M* (*SE*)	*t*	*p*	*d*
4–6	Group presence	Non-conformity	2.44 (0.34)	1.29	.21	0.26
		Conformity	0.64 (0.21)	−6.56	<.001	1.31
	Label	Non-conformity	2.06 (0.29)	0.23	.823	0.04
		Conformity	0.81 (0.21)	−5.56	<.001	0.99
	Generic	Non-conformity	2.85 (0.31)	2.71	.012	0.53
		Conformity	0.92 (0.28)	−3.81	.001	0.75
	Control	Non-conformity	1.38 (0.32)	−1.97	.061	0.40
		Conformity	1.04 (0.24)	−3.92	.001	0.80
7–9	Group presence	Non-conformity	2.25 (0.37)	0.67	.509	0.13
		Conformity	0.17 (0.10)	−18.65	<.001	3.81
	Label	Non-conformity	2.13 (0.39)	0.32	.755	0.06
		Conformity	0.25 (0.11)	−16.13	<.001	3.29
	Generic	Non-conformity	2.44 (0.34)	1.27	.217	0.25
		Conformity	0.16 (0.07)	−24.59	<.001	4.92
	Control	Non-conformity	0.75 (0.27)	−4.62	<.001	0.94
		Conformity	0.25 (0.15)	−11.63	<.001	2.37

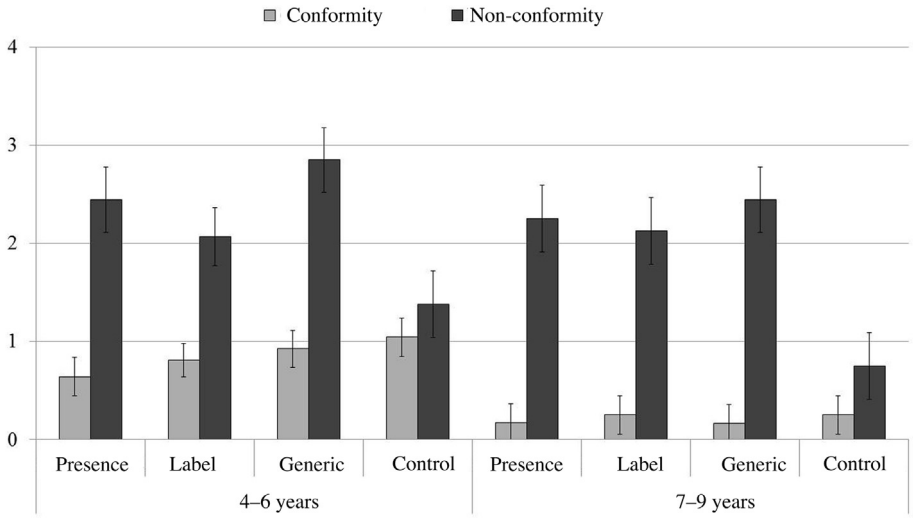

Fig. 1. Mean frequency of disapproval ("not okay") of conformity and non-conformity across age groups and conditions. Scores could range from 0 to 4. Error bars display standard errors.

lowest in the control condition compared with the other three conditions ($ps \leqslant .012$) and disapproval did not differ across the three experimental conditions ($ps > .55$). There were no interactions involving age group.

We next conducted one-sample *t* tests to test responses against chance (i.e., 2). When looking at disapproval for conformity, both age groups across all four conditions were below chance (thereby indicating approval). When looking at disapproval for non-conformity, 4- to 6-year-olds in the generic condition were above chance, 4- to 6-year-olds in the control condition were marginally below chance, and 7- to 9-year-olds in the control condition were below chance. All other groups were at chance. These statistics are presented in Table 2, and the data are presented graphically in Fig. 1. Non-parametric Wilcoxon signed-ranks tests provided further insight into the chance-level responses (see Table 3 for all data and statistics). Overall, these analyses revealed that in the experimental conditions, regarding non-conformity, half of the children most often approved and half of the children most often disapproved (although this was not the case for 4- to 6-year-olds in the generic statement condition). See the Discussion for insight into these response patterns.

Negativity toward non-conformity

We next focused only on the children who disapproved of non-conformity on at least one trial and were asked how bad the behavior was (1 = *kind of bad*, 2 = *pretty bad*, 3 = *very*, *very bad; n* = 138). A univariate ANOVA with age group (2: 4–6 or 7–9 years) as a between-participants variable and negativity as the dependent variable (i.e., average rating score across non-conformity trials on which children indicated "not okay"; scores could range from 1 to 3) yielded a main effect of age group, $F(1,130) = 21.46$, $p < .001$, $\eta_p^2 = .14$, showing that 4- to 6-year-olds ($M = 2.47$, $SE = 0.08$) were more negative than 7- to 9-year-olds ($M = 1.85$, $SE = 0.09$). This effect was qualified by a significant interaction of age group and condition, $F(3,138) = 3.76$, $p = .02$, $\eta_p^2 = .07$. The 4- to 6-year-olds were more negative than the 7- to 9-year-olds in the control condition (4–6 years: $M = 2.70$, $SE = 0.18$; 7–9 years: $M = 1.34$, $SE = 0.22$, $p < .001$) and the group presence condition (4–6 years: $M = 2.66$, $SE = 0.15$; 7–9 years: $M = 2.09$, $SE = 0.17$, $p = .012$). The two age groups did not differ in their negativity in either the label condition (4–6 years: $M = 2.19$, $SE = 0.14$; 7–9 years: $M = 2.00$, $SE = 0.17$, $p = .41$) or the generic condition (4–6 years: $M = 2.33$, $SE = 0.14$; 7–9 years: $M = 1.93$, $SE = 0.15$, $p = .057$), with ratings in both age groups corresponding roughly to the evaluation that the non-conformity was "pretty bad."

Table 3
Number of participants who more often approved ("okay"), more often disapproved ("not okay"), or approved and disapproved equally (tie) as a function of age group, condition, and behavior.

Age (years)	Condition	Behavior	Okay	Not okay	Tie	*Z*	*p*	*r*
4–6	Group presence	Non-conformity	10	14	1	1.50	.13	.30
		Conformity	20	1	4	−3.81	<.001	.76
	Label	Non-conformity	12	13	6	0.239	.81	.04
		Conformity	24	3	4	−3.79	<.001	.68
	Generic	Non-conformity	7	18	1	2.51	.012	.49
		Conformity	20	5	1	−3.06	.002	.60
	Control	Non-conformity	16	7	1	−1.79	.07	.37
		Conformity	16	3	5	−3.07	.002	.63
7–9	Group presence	Non-conformity	9	12	3	0.688	.49	.14
		Conformity	23	0	1	−4.63	<.001	.95
	Label	Non-conformity	12	12	0	−0.56	.57	.11
		Conformity	23	0	1	−4.51	<.001	.92
	Generic	Non-conformity	9	14	2	1.36	.17	.28
		Conformity	25	0	0	−4.72	<.001	.94
	Control	Non-conformity	21	3	0	−3.12	.002	.64
		Conformity	22	1	1	−4.58	<.001	.93

Positivity toward non-conformity

We next focused only on the children who approved of non-conformity on at least one trial and were asked how good the behavior was (1 = *kind of good*, 2 = *pretty good*, 3 = *very, very good*; $n = 125$). A univariate ANOVA with age group (2: 4–6 or 7–9 years) as a between-participants variable and positivity as the dependent variable (i.e., average rating score across non-conformity trials on which children indicated "okay"; scores could range from 1 to 3) yielded only a main effect of age group, $F(1,117) = 4.37$, $p = .039$, $\eta_p^2 = .04$, showing that 4- to 6-year-olds were more positive than 7- to 9-year-olds (4–6 years: $M = 2.33$, $SE = 0.07$; 7–9 years: $M = 2.09$, $SE = 0.08$) when approving of non-conformity.

Positivity toward conformity

Lastly, we focused on the children who approved of conformity on at least one trial and were asked how good the behavior was (1 = *kind of good*, 2 = *pretty good*, 3 = *very, very good*; $n = 196$). A univariate ANOVA with age group (2: 4–6 or 7–9 years) as a between-participants variable and positivity as the dependent variable (i.e., average rating score across conformity trials on which children indicated "okay"; scores could range from 1 to 3) yielded a main effect of age group, $F(1,188) = 8.71$, $p = .004$, $\eta_p^2 = .04$, showing that 4- to 6-year-olds were more positive than 7- to 9-year-olds (4–6 years: $M = 2.59$, $SE = 0.05$; 7–9 years: $M = 2.38$, $SE = 0.05$) when approving of conformity.

Explanations

We next analyzed the explanations that children provided after they were asked why they approved or disapproved of a given behavior (e.g., "Why is it okay for this one to eat this kind of berry?"). Explanations were given for three response types: disapproved non-conformity, approved non-conformity, and approved conformity. Responses for disapproved conformity were not analyzed because this response was rarely given (see Table 3). We focused on the frequency of the four primary explanation types (norm-based, group-based, individual-based, and similarity-based) within each response type (e.g., approved conformity), not including "other" miscellaneous explanations, and did not statistically compare across response types because not all participants provided each type of response. The percentage of times each explanation type was provided, out of the total number of trials, was analyzed via repeated measures ANOVAs in which age group and condition were between-participants variables, explanation type was a within-participants variable, and the percent-

age of given explanations was the dependent variable. The Huynh–Feldt correction, which adjusts the degrees of freedom for the calculated *F* values (Field, 2011; Huynh & Feldt, 1976), was used because the explanation data violated the repeated measures assumption of sphericity. See Tables 4 and 5 for all explanation data.

Explanations about disapproved non-conformity

Focusing on children who evaluated non-conformity as not okay ($n = 130$), we found a significant main effect of age group, $F(1, 122) = 6.70$, $p = .011$, $\eta_p^2 = .05$, a significant difference in explanation type, $F(2.77, 337.62) = 17.05$, $p < .001$, $\eta_p^2 = .12$, an interaction of age group and explanation type, $F(2.67, 337.62) = 4.20$, $p = .006$, $\eta_p^2 = .03$, an interaction of explanation type and condition, F

Table 4
Percentage of 4- to 6-year-olds' explanation types for each behavior, across conditions and evaluations types, separately for "not okay" and "okay" responses.

Condition	Behavior	Evaluation	Percentage of explanation types [*M* (*SE*)]				
			n	Norm	Group	Individual	Similarity
Group presence	Non-conformity	Not okay	20	6 (7)	25 (8)	10 (4)	35 (9)
	Non-conformity	Okay	11	0	5 (5)	19 (10)	14 (9)
	Conformity	Okay	24	10 (5)	17 (7)	15 (6)	29 (7)
Label	Non-conformity	Not okay	23	18 (7)	22 (7)	7 (3)	31 (8)
	Non-conformity	Okay	21	1 (3)	4 (4)	33 (7)	18 (7)
	Conformity	Okay	29	7 (4)	17 (6)	25 (6)	31 (7)
Generic	Non-conformity	Not okay	22	18 (7)	11 (7)	2 (3)	33 (8)
	Non-conformity	Okay	10	15 (5)	9 (6)	15 (11)	21 (10)
	Conformity	Okay	25	8 (5)	17 (7)	7 (6)	25 (7)
Control	Non-conformity	Not okay	11	15 (9)	0	2 (5)	41 (12)
	Non-conformity	Okay	19	1 (3)	0	12 (8)	20 (7)
	Conformity	Okay	23	15 (5)	23 (7)	8 (6)	19 (7)

Note. Scores represent percentage of each explanation type out of the total number of trials. Individual explanations could have been coded as of more than one type, and explanations that did not fit any of the coded types are not reported (which is why the percentages can add to more or less than 100). Data for disapproved ("not okay") conformity are not presented because this response was rarely given.

Table 5
Percentage of 7- to 9-year-olds' explanation types for each behavior, across conditions and evaluation types, separately for "not okay" and "okay" responses.

Condition	Behavior	Evaluation	Percentage of explanation types [*M* (*SE*)]				
			n	Norm	Group	Individual	Similarity
Group presence	Non-conformity	Not okay	16	22 (8)	41 (9)	2 (4)	47 (10)
	Non-conformity	Okay	12	2 (4)	19 (5)	56 (10)	25 (10)
	Conformity	Okay	22	6 (5)	23 (7)	31 (7)	23 (8)
Label	Non-conformity	Not okay	13	23 (9)	69 (10)	0	21 (11)
	Non-conformity	Okay	13	0	10 (5)	47 (10)	8 (9)
	Conformity	Okay	24	10 (5)	38 (7)	30 (6)	17 (7)
Generic	Non-conformity	Not okay	18	21 (7)	49 (8)	2 (4)	10 (09)
	Non-conformity	Okay	11	9 (4)	14 (5)	30 (10)	7 (9)
	Conformity	Okay	25	8 (5)	21 (7)	14 (6)	22 (7)
Control	Non-conformity	Not okay	7	7 (11)	0	14 (6)	54 (15)
	Non-conformity	Okay	21	5 (3)	0 (4)	54 (7)	15 (7)
	Conformity	Okay	23	0	11 (7)	33 (6)	22 (7)

Note. Scores represent percentage of each explanation type out of the total number of trials. Individual explanations could have been coded as of more than one type, and explanations that did not fit any of the coded types are not reported (which is why the percentages can add to more or less than 100). Data for disapproved ("not okay") conformity are not presented because this response was rarely given.

(8.30,337.62) = 3.55, $p < .001$, $\eta_p^2 = .08$, and an interaction of age group, explanation type, and condition, $F(8.30,337.62) = 1.96$, $p = .048$, $\eta_p^2 = .05$. The 4- to 6-year-olds gave mostly similarity-based explanations and did so at similar rates across all conditions, whereas the 7- to 9-year-olds gave mostly group-based explanations and did so most often in the label and generic conditions (although in the group presence condition they gave mostly similarity-based explanations).

Explanations about approved non-conformity

Focusing on participants who approved of non-conformity ($n = 118$), we found a significant main effect of age group, $F(1,110) = 9.71$, $p = .002$, $\eta_p^2 = .08$, a significant difference in explanation type, $F(2.16,237.53) = 26.40$, $p < .001$, $\eta_p^2 = .19$, and a significant interaction of age group and explanation type, $F(2.16,237.53) = 2.16$, $p = .001$, $\eta_p^2 = .07$. Both age groups gave mostly individual-based explanations, although 7- to 9-year-olds did so to a greater extent than 4- to 6-year-olds.

Explanations about approved conformity

Focusing on participants who approved of conformity ($n = 195$), we found an interaction of age group and explanation type, $F(2.86,535.38) = 3.27$, $p = .021$, $\eta_p^2 = .02$. The 4- to 6-year-old group gave more similarity-based explanations than norm-based explanations, and the 7- to 9-year-old group gave more group-based, individual-based, and similarity-based explanations than norm-based explanations.

Discussion

Children take what *is* to infer what *should be* (Kenward, 2012; Rakoczy & Schmidt, 2013), even with regard to unfamiliar third-party groups engaged in morally neutral behaviors (Roberts et al., 2016). That is, if young children are shown a group that is described as sharing a common property (i.e., descriptive regularity), children infer that individual group members should share that property and that it is bad if they do not (i.e., prescriptive judgment). In the current research, we tested the extent to which this descriptive-to-prescriptive tendency was facilitated by attention to group presence, category labels, or generic statements. Indeed, when group regularities were conveyed by showing that multiple individuals engage in a common behavior (i.e., group presence), words that highlighted an individual's group membership (i.e., category labels), or statements that associated group membership with specific properties (i.e., generic statements), children disapproved of non-conformity (consistent with our first prediction). In a control condition that included none of these factors, however, children were significantly less likely to disapprove of non-conformity (consistent with our second prediction), demonstrating that when the emphasis was on individuals and not groups, children were less prescriptive (see also Study 2 in Roberts et al., 2016). The ease with which group-based prescriptive reasoning is elicited, combined with its early emergence in younger children, suggests that such reasoning may be a fundamental aspect of human social cognition.

We also note that the generic statement condition was the only condition in which 4- to 6-year-olds disapproved of non-conformity above chance levels and in which they disapproved more often than they approved. At the same time, response patterns in the generic condition did not differ significantly from those in the other experimental conditions. Future research would be needed to determine whether or not generic statements may be more powerful than group presence or category labels alone in fostering prescriptive judgments (consistent with our third prediction and with previous research; Gelman et al., 2010; Rhodes et al., 2012). For example, a more fine-grained scale to assess the degree of negativity that children feel toward non-conformity may be more sensitive to condition differences in this age group. At a minimum, these data are consistent with prior research indicating that generics help children to learn about groups, guide their expectations about individuals, and help them to generate inferences about future behaviors (e.g., Cimpian & Erickson, 2012; Gelman, 2003; Gelman et al., 2010; Graham et al., 2011; Prasada, 2000) and further suggesting that generics may preferentially license children's prescriptive judgments.

One unexpected finding was that younger children rarely used norm-based reasoning (e.g., "They shouldn't do that") to explain why they disapproved of non-conformity, in contrast to earlier work

(Roberts et al., 2016). One possibility is that although group presence, category labels, and generic statements elicit disapproval and negativity, they might only be collectively strong enough to elicit norm-based justifications; perhaps a combination of the three conveys a stronger regularity that licenses children to make strong normative justifications. Future work is needed to test this empirically. For now, however, the current data indicate that group presence, category labels, and generic statements license prescriptiveness in the form of disapproval and negativity.

Also of note, younger children tended to respond in a more extreme manner than older children; they were more disapproving than older children on both conformity and non-conformity trials, and they were more negative on the negativity scale and more positive on the positivity scale. One interpretation of these patterns is that younger children are more attentive to both conformity and non-conformity and, therefore, feel especially positive about conformity and especially negative about non-conformity. This interpretation aligns with recent research suggesting that 3-year-olds are significantly more troubled by norm violations than 5-year-olds (Hardecker et al., 2016). Regardless, of primary interest here was the directionality of children's responses, which varied as a function of the specific behaviors they saw (i.e., non-conformity vs. conformity). That is, younger children's more extreme responses are independent of the finding that they were systematic in their responses; in the experimental conditions, they were more disapproving of non-conformity than of conformity and were more approving of conformity than of non-conformity.

An important question for future research is why the children interpret group presence, labels, and generic statements as indicative of what is normatively correct. One possibility is that each of these features provides children with the information that a given property is systematically shared among a group of individuals rather than possessed by only one individual. Once children take note of a descriptive regularity, they may reason that adherence and conformity to group regularities is critical for individual and group functioning (Tomasello, 2016) and subsequently interpret the regularity as having prescriptive force. In other words, children may be prepared to detect group regularities that could be signaled multiply by group presence, category labels, and generic statements.

A second possibility (not mutually exclusive from the first possibility) is that such cues encourage children to essentialize the social group in question (see Bigler & Liben, 2006, for a related theoretical model). For example, when children see a group of individuals who share common features (e.g., people with dark skin live on the other side of town), they may come to conceptualize those features as essential to their identity and subsequently use group membership as a basis for prescriptiveness (e.g., people with dark skin *should* live on the other side of town). The explanation data support this possibility, showing that in the absence of labels, children justified their prescriptiveness through similarity-based explanations. Similarly, labels and generic statements could signal to children that such group membership is salient and important, subsequently influencing them to interpret group membership as essential and therefore prescriptive. Indeed, in both of these conditions, both age groups most often appealed to group-based explanations (i.e., they often appealed to the category label, which was given to them in both of these conditions). This finding is consistent with recent research suggesting that when children (ages 5–10 years) are given race-based labels, they are especially likely to essentialize race (Roberts & Gelman, in press). Future research is needed to systematically test the extent to which essentialist reasoning mediates the descriptive-to-prescriptive tendency detected here, which may also help to understand the non-random individual response patterns. Recall from the non-parametric tests that children were often split in their evaluations of non-conformity; half of the children most often disapproved, whereas the other half most often approved (excluding 4- to 6-year-olds in the generic statement condition). One possibility is that children who interpreted the properties as "essential" to the group were more likely to disapprove of non-conformity. Indeed, recent work suggests that children who attribute behaviors to groups (e.g., Hibbles eat that kind of berry because of something about *Hibbles*) rather than to properties (e.g., Hibbles eat that kind of berry because of something about that *berry*) are more likely to conceptualize non-conformity as a norm violation (see Tworek & Cimpian, 2016).

We also stress two important limitations to the current study. First, the data were derived exclusively from U.S. children (for recent research on U.S. children's normative reasoning, see also Conry-Murray & Turiel, 2012; Cooley & Killen, 2015; Josephs, Kushnir, Gräfenhain, & Rakoczy, 2016; Kalish, 2012; Roberts et al., 2016) and therefore limit our understanding of children's normative rea-

soning across human populations. Cross-cultural research with non-U.S. samples may yield more nuanced insights (e.g., children from societies in which group norms are valued more highly may be more sensitive to non-conformity; see Wang, 2016). Second, we focused only on social categories, thereby leaving open the question of whether similar patterns would be found for other sorts of categories (e.g., non-human animals, artifacts). Future research would do well to compare children's descriptive-to-prescriptive tendency across different categorical domains, which would reveal the generality or specificity of this tendency (for insights, see Rhodes & Gelman, 2009).

In conclusion, the current research contributes to a growing body of literature on children's reasoning about norms (Engelmann et al., 2016; Riggs & Young, 2016; Schmidt, Hardecker, & Tomasello, 2016), showing that such reasoning influences children's descriptive-to-prescriptive tendency even under minimal contexts with minimal cues to group regularities. Stated more plainly, when children see a group of individuals, hear a category label, or hear that label explicitly associated with behaviors, they believe that individuals within that group *should* conform and that it is bad if they do not.

Acknowledgments

This research was supported by a National Science Foundation Graduate Research Fellowship and a Ford Foundation Predoctoral Research Fellowship to Steven O. Roberts and a National Institute of Child Health and Human Development grant (HD-36043) to Susan A. Gelman. We thank the participants in the study and also thank Madeline Adelman, Justine Ezell, Yara Gayar, Melissa Granville, Antonio Malkoun, Hyeongjun "Jack" Kim, Rebecca Strauss, and Tiffany Valencia for assistance with data collection. We thank Craig Smith and the Living Laboratory at the University of Michigan, and we are also grateful for the feedback of Yarrow Dunham, E. Margaret Evans, Alexandra Freund, Stephanie Rowley, Angeline Lillard, and Henry Wellman.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jecp.2016.11.013.

References

Abrams, D., Rutland, A., & Cameron, L. (2003). The development of subjective group dynamics: Children's judgments of normative and deviant in-group and out-group individuals. *Child Development, 74*, 1840–1856.
Asch, S. E. (1955). Opinions and social pressure. *Scientific American, 193*, 31–35.
Baron, A. S., Dunham, Y., Banaji, M., & Carey, S. (2014). Constraints on the acquisition of social category concepts. *Journal of Cognition and Development, 15*, 238–268.
Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development, 68*, 530–543.
Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. In R. V. Kail (Ed.). *Advances in child development and behavior* (Vol. 34, pp. 39–89). San Diego: Elsevier.
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.
Cimpian, A., & Erickson, L. C. (2012). Remembering kinds: New evidence that categories are privileged in children's thinking. *Cognitive Psychology, 64*, 161–185.
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.
Cooley, S., & Killen, M. (2015). Children's evaluations of resource allocation in the context of group norms. *Developmental Psychology, 51*, 554–563.
Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences, 7*, 173–179.
Engelmann, J. M., Herrmann, E., Rapp, D. J., & Tomasello, M. (2016). Young children (sometimes) do the right thing even when their peers do not. *Cognitive Development, 39*, 86–92.
Feldman, S. (2003). Enforcing social conformity: A theory of authoritarianism. *Political Psychology, 24*, 41–74.
Field, A. (2011). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.
Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday life*. New York: Oxford University Press.
Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183–209.
Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology, 61*, 273–301.
Göckeritz, S., Schmidt, M. F. H., & Tomasello, M. (2014). Young children's creation and transmission of social norms. *Cognitive Development, 30*, 81–95.

Graham, S. A., Keates, J., Vukatana, E., & Khu, M. (2013). Distinct labels attenuate 15-month-olds' attention to shape in an inductive inference task. *Frontiers in Psychology, 3*. http://dx.doi.org/10.3389/fpsyg.2012.00586.
Graham, S. A., Nayer, S. L., & Gelman, S. A. (2011). Two-year-olds use the generic/nongeneric distinction to guide their inferences about novel kinds. *Child Development, 82*, 493–507.
Hardecker, S., Schmidt, M. F. H., Roden, M., & Tomasello, M. (2016). Young children's behavioral and emotional responses to different social norm violations. *Journal of Experimental Child Psychology, 150*, 364–379.
Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics, 1*, 69–82.
Josephs, M., Kushnir, T., Gräfenhain, M., & Rakoczy, H. (2016). Children protest moral and conventional norms when they believe actions are freely chosen. *Journal of Experimental Psychology, 141*, 247–255.
Kalish, C. W. (2012). Generalizing norms and preferences within social categories and individuals. *Developmental Psychology, 48*, 1133–1143.
Kenward, B. (2012). Over-imitating preschoolers believe unnecessary actions are normative and enforce their performance by a third party. *Journal of Experimental Child Psychology, 112*, 195–207.
Liben, L. S., Bigler, R. S., & Krogh, H. R. (2001). Pink and blue collar jobs: Children's judgments of job status and job aspirations in relation to sex of worker. *Journal of Experimental Child Psychology, 79*, 346–363.
Macnamara, J. (1982). *Names for things*. Cambridge, MA: MIT Press.
Mulvey, K. L. (2016). Evaluations of moral and conventional intergroup transgressions. *British Journal of Developmental Psychology, 34*, 489–501.
Prasada, S. (2000). Acquiring generic knowledge. *Trends in Cognitive Science, 4*, 66–72.
Quintana, S. M. (1998). Children's developmental understanding of ethnicity and race. *Applied and Preventive Psychology, 7*, 27–45.
Rakoczy, H., Hamann, K., Warneken, F., & Tomasello, M. (2010). Bigger knows better: Young children selectively learn rule games from adults rather than from peers. *British Journal of Developmental Psychology, 28*, 785–798.
Rakoczy, H., & Schmidt, M. F. H. (2013). The early ontogeny of social norms. *Child Development Perspectives, 7*, 17–21.
Rhodes, M. (2014). Children's explanations as windows into their intuitive theories of the social world. *Cognitive Science, 38*, 1687–1697.
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.
Rhodes, M., Leslie, S.-J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 13526–13531.
Riggs, A. E., & Young, A. G. (2016). Developmental changes in children's normative reasoning across learning contexts and collaborative roles. *Developmental Psychology, 52*, 1236–1246.
Roberts, S. O., & Gelman, S. A. (2015). Do children see in Black and White? Children's and adults' categorizations of multiracial individuals. *Child Development, 86*, 1830–1847.
Roberts, S. O., & Gelman, S. A. (2016). Can White children grow up to be Black? Children's reasoning about the stability of emotion and race. *Developmental Psychology, 52*, 887–893.
Roberts, S. O., & Gelman, S. A. (2017). Multiracial children's and adults' categorizations of multiracial individuals. *Journal of Cognition and Development., 18*, 1–15. http://dx.doi.org/10.1080/15248372.2015.1086772.
Roberts, S. O., & Gelman, S. A. (in press). Now you see race, now you don't: Verbal cues influence children's reasoning about the stability of race, *Cognitive Development*.
Roberts, S. O., Gelman, S. A., & Ho, A. K. (2016). So it is, so it shall be: Group regularities license children's prescriptive judgments. *Cognitive Science*. http://dx.doi.org/10.1111/cogs.12443.
Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young children see a single action and infer a social norm. Promiscuous normativity in 3-year-olds. *Psychological Science, 27*, 1360–1370.
Schmidt, M. F. H., Hardecker, S., & Tomasello, M. (2016). Preschoolers understand the normativity of cooperatively structured competition. *Journal of Experimental Child Psychology, 143*, 34–47.
Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2011). Young children attribute normativity to novel actions without pedagogy or normative language. *Developmental Science, 14*, 530–539.
Smetana, J. G., Rote, W. M., Jambon, M., Tasopoulos-Chan, M., Villalobos, M., & Comer, J. (2012). Developmental changes and individual differences in young children's moral judgments. *Child Development, 83*, 683–696.
Tomasello, M. (2016). Cultural learning redux. *Child Development, 87*, 643–653.
Tworek, C. M., & Cimpian, A. (2016). Why do people tend to infer "ought" from "is"? The role of biases in explanation. *Psychological Science, 27*, 1109–1122.
Wang, Q. (2016). Why should we all be cultural psychologists? Lessons from the study of social cognition. *Perspectives on Psychological Science, 11*, 583–596.
Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-month-old infants. *Cognitive Psychology, 29*, 257–302.
Waxman, S. R. (2010). Names will never hurt me? Naming and the development of racial and gender categories in preschool-aged children. *European Journal of Social Psychology, 40*, 593–610.