Cognition 162 (2017) 133–142

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Mutual intentions as a causal framework for social groups

Alexander Noyes *, Yarrow Dunham

Yale University, Kirtland Hall, 2 Hillhouse Ave, New Haven, CT 06511, USA

ARTICLE INFO

Article history:
Received 12 October 2016
Revised 10 February 2017
Accepted 16 February 2017

Keywords:
Social categorization
Cognitive development
Intuitive theories
Social cognition

ABSTRACT

Children's early emerging intuitive theories are specialized for different conceptual domains. Recently attention has turned to children's concepts of social groups, finding that children believe that many social groups mark uniquely social information such as allegiances and obligations. But another critical component of intuitive theories, the causal beliefs that underlie category membership, has received less attention. We propose that children believe membership in these groups is constituted by mutual intentions: i.e., all group members (including the individual) intend for an individual to be a member and all group members (including the individual) have common knowledge of these intentions. Children in a broad age range (4–9) applied a mutual-intentional framework to newly encountered social groups early in development (Experiment 1, 2, 4). Further, they deploy this mutual-intentional framework selectively, withholding it from essentialized social categories such as gender (Experiment 3). Mutual intentionality appears to be a vital aspect of children's naïve sociology.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

From an early age, children are sensitive to the categories to which their fellow humans belong. An important question is whether and how these concepts are supported by domain-specific *intuitive theories* (Wellman & Gelman, 1992) – i.e., unique expectations and causal beliefs that structure children's understanding of social groups. Here we focus on the possibility that children's causal beliefs about social groups are distinct from other early-emerging category domains in that they are constituted by the mutual intentions of the individuals that compose them.

Past work in this area has focused on the types of properties children expect category members to share, i.e. the inductive potential of social categories. This work suggests that children approach social groups from two distinct perspectives (Rhodes, 2012). First, in some cases children have an *essentialist* perspective, believing that social categories – much like biological species – share deep similarities (Diesendruck & HaLevi, 2006; Hirschfeld, 1998). This perspective is most strongly applied to gender in the United States (Gelman, Collman, & Maccoby, 1986; Taylor, 1996) and strongly (though not mostly strongly) elsewhere (Diesendruck, Goldfein-Elbaz, Rhodes, Gelman, & Neumark, 2013). For instance, children expect children of the same gender to share diverse biological and behavioral properties and expect gender categories to be stable across time; these expectations emerge earlier for gender than for most other social categories, but eventually emerge for some other categories such as race (Haslam, Rothschild, & Ernst, 2000). This perspective, however, is not universally applied to all social categories (notably, novel social categories), and may require additional input to be activated (for example, generic language; Rhodes, Leslie, & Tworek, 2012). In other words, essentialist beliefs are not applied to all newly encountered social groups and so reflect just one way in which children construe social categories (Rhodes & Brickman, 2011).

For most other social groups, including most newly encountered groups, children apply a second perspective, wherein social categories entail uniquely social-relational or coalitional information. For example, children predict that members of a social category (such as race) are friends earlier than they predict that members of that same category share deep similarities (Shutts, Roben, & Spelke, 2013). Furthermore, children reliably expect that novel social categories mark deontic relationships, such as an intrinsic obligation not to harm (Kalish & Lawson, 2008; Rhodes & Chalik, 2013). Taken together, these findings suggest that children approach social groups via a "naïve sociology" (Hirschfeld, 1998) revealed through the domain-specific expectations they bring to bear on social groups.

However, domain-specific inferences are only one component of intuitive theories. Indeed, domain-specific inferences often stem from (or partner with) specialized causal beliefs. For example, children not only expect animals to share many of their properties, they suspect that these similarities are *caused* by a singular underlying essence (Gelman, 2003), roughly, an internal cause transmit-

* Corresponding author.
E-mail address: alexander.noyes@yale.edu (A. Noyes).

http://dx.doi.org/10.1016/j.cognition.2017.02.007
0010-0277/© 2017 Elsevier B.V. All rights reserved.

ted from parent to offspring (Keil, 1992). Furthermore, the functional properties children expect artifacts to share are related to their belief that artifact identity is constituted by the creator's intentions to bring an object with particular properties into being (Diesendruck et al., 2003). These causal beliefs are critical to understanding why children believe categories license certain types of predictions, for example the prediction that a fork shaped to resemble a spoon has *become* a spoon while a tiger groomed to resemble a lion has remained a tiger (Bloom, 1996; Gelman, 1988, 2003; Keil, 1992). Therefore, a complete understanding of children's intuitive theories requires identifying their underlying causal beliefs.

For those social categories that children believe mark deep similarities, notably gender, children's causal beliefs are similar to their causal beliefs for animal categories. That is, children believe gender categories are based in inherent and natural properties (Taylor, 1996; Taylor, Rhodes, & Gelman, 2009). As we noted, however, children are not readily essentialist about most newly encountered social groups (Rhodes & Brickman, 2011; Rhodes et al., 2012), and there is little research outlining the underlying causal beliefs children bring to bear in these non-essentialist cases. That is, we know little about the underlying causal beliefs governing groups that children believe mark obligations and allegiances but not deep similarities. We propose that in these cases *mutual intentions* serve as the causal framework that supports judgments of category membership. By "mutual intentions" we refer to the general agreement of individuals that they belong to a group. Mutual intentions involve two key components. Any specific individual belongs to the group by virtue of two facts that must be known to all parties: that she intends to belong to the group and that other group members intend for her to belong to the group.

To illustrate the causal role of mutual intentions it is helpful to draw an analogy to other socially constituted entities such as money. A dollar bill is money not because of the inherent value of its raw materials but because of the general agreement of Americans that dollar bills count as money (and thus have value) (Searle, 1995, 2006). Mutual intentions possess two key components. First, the intentions need to be shared across a group or community. Often these intentions have the structure of "we intend that X counts as Y"; for example, Americans intend that dollar bills (X) counts as money (Y). The objects or individuals (X) do not inherently possess the properties of Y or belong to the kind Y. Rather, Y is a social role that is assigned to the object or individual. Second, the intentions cannot simply be shared by coincidence. They must be shared in the sense of being mutually understood. Thus, there needs to be common knowledge that others intend X to count as Y.

Shared intentions of this sort play a fundamental role in the creation of many social institutions from group norms to linguistic symbols (Searle, 1995, 2006; Tomasello, Carpenter, Call, Behne, & Moll, 2005). We argue that an analogous causal pathway underlies social groups. Children readily treat groups marked by otherwise trivial features such as t-shirts as socially important and informative (Dunham, Baron, & Carey, 2011; Rhodes, 2012). Like money, these groups are not meaningful because of the properties of their shared (trivial) markers, nor, we argue, does it stem from shared inherent properties or individual intentions. Rather, we propose that children assume membership is constituted by the mutual intentions of the group, i.e. the mutual agreement of individual group members that the group has social meaning and that any and all individual members belong to it. The goal of this paper is to provide evidence that children are sensitive to these abstract causal relationships in the context of social groups.

Past work has demonstrated that children have an early emerging and rich understanding of intentions. Not only do they understand others' intentions and mental states from infancy (Baillargeon, Scott, & He, 2010; Buttelmann, Carpenter, & Tomasello, 2009; Onishi & Baillargeon, 2005), but they understand how intentions influence social interactions and the behavior of others (Choi & Luo, 2015). Finally, they have an ability to participate in joint attention and inter-subjectivity (Tomasello, 1992; Tomasello & Rakoczy, 2003), opening up the possibility for them to engage in collaboration and coordinate their intentions with others. In our work we ask whether children understand that intentions are not only socially relevant, but indeed structure social reality by serving as the foundation of some social groups. More specifically, we hypothesize that children recognize the causal role of mutual intentions in creating groups and instantiating group membership. Indeed, we believe that mutual intentions may motivate their other domain-specific expectations concerning group allegiance and obligations to in-group members – mutual intentions may help causally explain these other social properties.

We test these ideas across three experiments. Overall, the logic of our design is based on the classic transformation paradigm (Keil, 1992), which is used to explore children's causal beliefs. In the transformation paradigm, an individual category members undergo changes to their properties. Children are asked to categorize the individual after transformation (i.e., a lion that changes its outer appearance to look like a tiger). If category membership changes, then it suggests children believe the changed property is constitutive of (i.e., causally central to) category membership. If category membership does not change, the property must not be causally central. In the present study, we ask whether changes in mutual intentions (the mutual understanding of an individual's group membership) causes an individual's group status to change. We contrast mutual intentions with other possible transformations.

In Experiment 1, we ask whether children believe mutual intentions underlie becoming a group member. Here, we compare mutual intentions to non-mutual intentions, such as individual intentions that are un-reciprocated by the members of the group. In Experiment 2, we ask two follow-up questions: First, whether children believe that changes in mutual intentions can change group membership (changing from one group to another), and second, whether they privilege the causal relevance of mutual intentions over changes in other salient properties like visual group markers (i.e., t-shirt color) that others have proposed are important (e.g., Aboud, 1988). Finally, in Experiment 3, we explore the role of mutual intentions for the special case of gender, a highly essentialized social category. Given that past literature has demonstrated that some select social categories, and gender in particular, operate via a different system of causal beliefs – namely essentialism (Gelman, 2003; Rhodes, 2012) – we predict that children will not view mutual intentions as causally central to gender categories. Thus, we expect that children believe mutual intentions are causally relevant to newly encountered social groups (such as the groups they encounter in Experiment 1 and 2), but withdraw this expectation from essentialized social categories.

In conclusion, our hypotheses fit with the broader notion that children have two intuitive theories for social groups (Rhodes, 2012), one essentialist and one more relational and coalitional; past work has characterized the causal beliefs underlying essentialist groups but has not characterized coalitional/relational groups. The goal of our paper is to characterize the causal beliefs applied to coalitional groups and to demonstrate that they are distinct from the causal beliefs applied to essentialized groups. Thus, we expect a mutual intentional framework will be selectively applied to coalitional groups (exemplified here by newly encountered groups; Experiment 1 and 2) and that an essentialist causal framework will be selectively applied to essentialized social categories (exemplified here by gender; Experiment 3). Finally, we go on to show that children require that the mutual intentions be jointly known to all parties, demonstrating that the mere coinci-

dental alignment of intentions is not sufficient to underwrite a claim of group membership (Experiment 4).

2. Experiment one

2.1. Methods

2.1.1. Participants

We follow previous research by focusing on two age groups, 4–5 year-olds and 7–9 year-olds (Rhodes & Chalik, 2013). These two age groups straddle a turning point in children's group-based reasoning over which group-based cognition appear to strengthen (e.g., Baron, Dunham, Banaji, & Carey, 2014). Sixteen children participated in each of the younger (M = 4.59, SD = 0.61) and older (M = 8.48, SD = 0.77) age groups; we based our sample sizes on previous work (e.g. Rhodes & Chalik, 2013) and on the assumption of a large effect, as revealed in past work with transformation paradigms, which generally produce consistent patterns of expectations across children (Keil, 1992). Sixteen males and sixteen females participated.

2.1.2. Stimuli

Building on past work (e.g. Rhodes & Brickman, 2011; Rhodes & Chalik, 2013), children were introduced to two novel social groups, the Flurps and Zazzes. This work has established that children treat groups of this sort as marking obligation to the in-group, but do not treat them as essentialized kinds. Unlike this past work, in Experiment 1 we removed salient visual group markers (e.g., distinct t-shirts) to highlight the contrast between mutual and non-mutual intentions. The two groups could be distinguished by verbal labels (provided by the experimenter) and their proximity to each other (all of the Flurps were playing together and all of the Zazzes were playing together). The two groups could not be distinguished by their appearance. Group members had *no* consistent appearance; each Flurp, for example, was wearing a different t-shirt color. We explored the impact of salient visual markers in Experiment 2.

2.1.3. Design and procedure

Children were introduced to 4 members of the Flurps and the 4 members of the Zazzes (presented on screen, verbally labelled, and visually segregated). In each trial children were introduced to new individuals one at a time; these individuals were labelled with an individuating label (a unique first name) and explicitly described as belonging to neither group. Children were asked which group the individual belonged to and were required to verbally confirm that the individual belonged to neither group in order to proceed. Flurps and Zazzes were both depicted by cartoon humans; pictures varied in gender composition (male and female), hair color, and clothing (as stated before, group members wore different t-shirt colors).

To test whether mutual intentions (rather than non-mutual intentions) constituted group membership, we presented children with three trial types: First, in *Individual-only* trials, an individual non-member wanted to (and declared themselves to be) the member of one of the groups, but the group members believed the individual could not join and viewed the individual as a non-member. For example, "*Amy says she wants to be a Flurp. Amy tells the Flurps she wants to be a Flurp. The Flurps tell Amy she cannot be a Flurp. The Flurps tell Amy she is not a Flurp.*" In *Group-only* trials, all members of the group viewed the individual as a member, but the individual did not want to join and viewed herself as a non-member. For example, "*The Zazzes say John can be a Zazz. The Zazzes tell John he is a Zazz. John tells the Zazzes he does not want to be a Zazz.*" In *Mutual* trials, the individual and group members mutually viewed the individual as a member of the group. For example, "*Mark says he wants to be a Flurp. Mark tells the Flurps he wants to be a Flurp. The Flurps tell Mark he is a Flurp.*"

Children received 9 total trials, 3 of each of type. In each case children had to decide whether the individual had become a member of the relevant group or not. All children received a predetermined pseudo-random order based in three 'blocks' containing one trial of each type. Based on our analysis of the role of mutual intentions, we predicted that children would only endorse membership at high rates in the *mutual* trials.

2.2. Results

Primary results are depicted in Fig. 1. We analyzed the number of trials children reported the character had joined the group (0–3) with a 2 (Age Group: 4–5; 7–9) by 3 (Condition, within-subjects: individual, group, mutual) mixed-design ANOVA. There was a main effect of condition, $F(2,87) = 22.34$, $p < 0.001$, $\eta^2 = 0.32$ but no main effect of age group, $F(1,87) = 0.18$, $p = 0.669$, $\eta^2 = 0.0013$. However, there was a significant age group by condition interaction, $F(2,87) = 3.68$, $p = 0.029$, $\eta^2 = 0.053$. Visual inspection of Fig. 1 suggests that this interaction was driven by a qualitatively similar but quantitatively starker pattern of results in older children than younger children.

2.2.1. Younger children - 4–5 year-olds

For younger children, a one-way ANOVA on children's category judgments (0–3) revealed an overall condition difference, $F(2,46) = 24.23$, $p < 0.001$, $\eta^2 = 52$. Younger children reported far greater group change in the mutual condition (73% of trials) than the non-mutual conditions (7%), $F(1,46) = 47.66$, $p < 0.001$, $\eta^2 = 0.51$. There was no difference, however, between the individual-only (13% of trials) and group-only (2% of trials) condition, $t(15) = -1.16$, $p = 0.264$, $d = 0.42$. One-sample t-tests comparing category judgments (0–3) to chance (1.5), revealed that younger children were above chance in the mutual condition, $t(15) = 2.25$, $p = 0.040$, $d = 0.56$, but below chance in the individual-only, $t(15) = 4.39$, $p < 0.001$, $d = 1.10$, and group-only, $t(15) = -23$, $p < 0.001$, $d = 5.75$, conditions.

2.2.2. Older children - 7–9 year-olds

Older children expressed the same overall pattern as younger children, but with even more clarity. A one-way ANOVA again revealed an overall condition difference, $F(2,45) = 225$, $p < 0.001$, $\eta^2 = 0.91$. Older children reported far greater group change in the

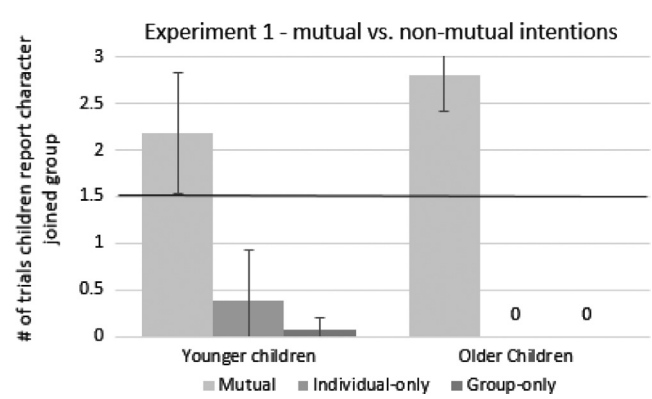

Fig. 1. Average number of trials (out of 3) that children reported that the unaffiliated character successfully joined the relevant group, as a function of age group (4–5 vs 7–9) and experimental condition (mutual intentions, individual-only intention, and group-only intention). Both age groups were above chance in the mutual condition and below chance in the non-mutual condition. Error bars are 95% confidence intervals.

mutual condition (93% of trials) than the non-mutual conditions (0%), $F(1,46) = 460$, $p < 0.001$, $d = 0.91$. There was no difference, however, between the individual-only and group-only condition; in fact, older children never reported that the non-member character joined in these cases. In the mutual condition, older children were also above chance, $t(15) = 7.00$, $p < 0.001$, $d = 1.75$, and they were below chance for both other conditions (indeed, older children never reported change in these conditions).

2.3. Discussion

The results of this study support the causal role of mutual intentions. Children across all age ranges believed that a non-member character could become a group member only through the mutual agreement of all group members *and* the individual him or herself. Children did not believe that *either* the group's intentions *or* the individual's intentions were sufficient to create group membership. We take these findings to provide initial evidence of a causal role for mutual intentions in determining group membership, but despite the clear pattern, several additional questions deserved further investigation. First, Experiment 1 involved *joining* a group rather than *changing* from one group to another; the latter transformation case might be a stronger test of the role of mutual intentions because licensing a transformation in the latter case also requires revoking a past group membership, something children might be reluctant to do (Solomon, 2002). A change in membership also provides a cleaner analog to past transformation studies, which generally probed intuitions about category change rather than initial category assignment. Second, many real-world groups involve salient group markers that visually signal category membership, but these cues were not present in Experiment 1. This leaves open the question of whether children will think mutual intentions trump other cues to group membership.

Thus, in Experiment 2 all characters started out as a member of one group and all transformations involved a possible shift from one group to the other, bringing us closer in line with past transformation studies and allowing us to examine whether children are willing to override a prior group membership when presented with information about mutual intentions (e.g., Keil, 1992). In addition, we added salient visual markers (red and blue t-shirts) as contrasting properties. This means that at the start of each trial the target child shares a salient visual property with their original group – a property originally introduced as a group marker – such that children, if they are to follow mutual intentions, have to actively override the visual similarity in making that judgment. Indeed, some influential past accounts (Aboud, 1988; Katz, 1983) have claimed that salient perceptual cues are among the most powerful early determinants of children's social category judgments. If children override such salient cues based on the much more abstract mutual intention-based cues, it would further bolster our claims that such intentions are central to the causal theory underlying social groups. Therefore, Experiment 2 provides a considerably stronger test of the role of mutual intentions in determining category membership.

3. Experiment two

3.1. Methods

3.1.1. Participants

We focus on the same two age groups as Experiment 1. Post-hoc analyses revealed that the sample size was appropriate for the older children, who were at ceiling, but may have been too small for the younger children on certain measures; the smallest effect in younger children in Experiment 1 was medium in size ($d = 0.56$) for the critical test comparing their categorization of individuals in the mutual condition to chance. Thus, to ensure sufficient power (>0.80) on all critical measures in experiment 2, we increased the sample size of younger children to 30, giving us 0.84 power to detect an effect of that size. Therefore, 16 children participated in the older age range (M = 8.73, SD = 0.85) and 30 children participated in the younger age range (M = 5.14, SD = 0.59. Thirty females and 16 males participated. All children in this study and in Experiment 3 were recruited from two Connecticut science museums with diverse attendance.

3.1.2. Stimuli

Stimuli were identical to Experiment 1 except for the addition of blue and red t-shirts to mark the groups.

3.1.3. Design and procedure

The procedure is similar to Experiment 1 but with several important changes. First, children were introduced to two novel social groups that were visually marked by shirt color and were explicitly called the "red team" and "blue team" in addition to the same noun labels used previously, i.e. the Zazzes and the Flurps. Children were introduced to characters who began already on one of the two group and thus were initially wearing the shirt color of that group.

Children heard stories where one of the characters' properties change. In an *appearance-only trial*, the character loses their group shirt and puts on the opposite group shirt, but they intend to remain in the same group. For example, "*Amy says she wants to be a Flurp but one day she lost her blue shirt and now she wears a red shirt. The Zazzes say Amy can be a Zazz. Amy says she does not want to be a Zazz. Amy says 'I'm a Flurp.'*" In an *intention-only trial*, the character intends to join the opposite group, but does not change their group t-shirt (and thus continues to look like the original group members). For example, "*Mark says he does not want to be a Zazz. . .Mark doesn't have a blue shirt so he still wears his red shirt. Mark says he wants to be a Flurp. Mark says 'I'm a Flurp.' The Flurps say Mark can be a Flurp.*" In an *appearance + intention trial*, the character intends to join the opposite group and changes their t-shirt to match their intention. For example, "*Sarah says she does not want to be a Zazz. . .so she decides to wear a blue shirt. Sarah says she wants to be a Flurp. Sarah says 'I am a Flurp'. The Flurps say Sarah can be a Flurp.*"

To reduce the number of conditions in order to focus on how intentions compete with other group markers, we fixed the intentions of the potential new group across all trials by indicating, across all trials, that the members of the potential new group always intended for the character to join their group. Thus, by comparing the *appearance-only* condition (the only condition without mutual intentionality) and the other two conditions we gain a sense of the casual power of mutual intentionality. By comparing the *intention-only* and *appearance + intention* condition we gain a sense of the additional influence of a visual cue.

Children received 9 total trials, 3 of each of these types. In each case, children were asked to report which group the character belonged to after the transformation. All children received a predetermined pseudo-random order based in three 'blocks' containing one trial of each type. We predicted that mutual intentions would play the dominant role in children's judgments; thus, we expected no difference between both conditions with mutual intentionality (*intention-only* and *appearance + intention*), and expected both of those conditions to be endorsed more frequently than the *appearance-only* condition.

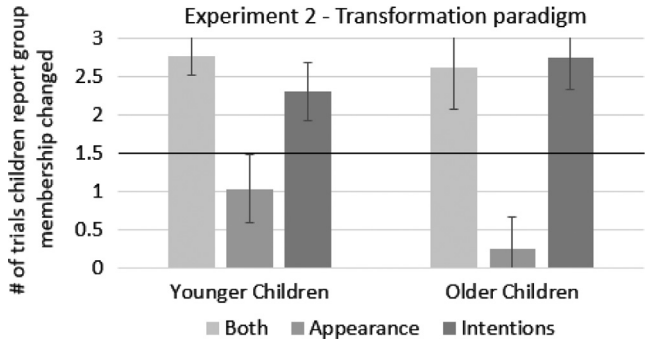

Fig. 2. Average number of trials (out of 3) that children reported a character changed group status, as a function of age (4–5 vs 7–9) and experimental condition (appearance-change, intention-change, and both). Error bars are 95% confidence intervals.

3.2. Results

An overview of results is presented in Fig. 2. We analyzed results with a 2 (Age group, between-subject: 4–5 vs. 7–9) by 3 (Condition, within-subject: intentions + appearance, intentions-only, appearance-only) mixed-design ANOVA. There was a main effect of condition, $F(2,129) = 19.69$, $p < 0.001$, $\eta^2 = 0.23$, no main effect of age group, $F(1,129) = 0.05$, $p = 0.819$, $\eta^2 = <0.001$, and no significant interaction between age group and condition, $F(1,129) = 1.09$, $p = 0.339$, $\eta^2 = 0.025$. Thus, both older and younger children appeared to respond similarly and, as predicted, were generally most swayed by mutual intentions. Despite the non-significance of the interaction between condition and age, to facilitate comparison with Experiment 1 we present results for each age group separately.

3.2.1. Younger children - 4–5 year-olds

A one-way ANOVA revealed an overall effect of condition, $F(2,84) = 3.70$, $p = 0.029$, $\eta^2 = 0.081$. Younger children reported far greater group change when intentions and appearance both changed than when only appearance changed, $t(29) = 7.24$, $p < 0.001$, $d = 1.81$. They also reported more change when intentions and appearance both changed than when only intentions changed, $t(29) = 2.31$, $p = 0.028$, $d = 0.54$. This difference, while smaller, was contrary to our hypothesis. Critically, however, children reported far more change when intentions alone changed than when appearance alone changed, $t(29) = 4.03$, $p < 0.001$, $d = 1.14$, suggesting intentions are a far stronger factor in children's membership judgments. Critically, and supporting our hypothesis, younger children reported more change than expected by chance when both intentions and appearance changed, $t(29) = 10.22$, $p < 0.001$, $d = 1.87$, and when only intentions changed, $t(29) = 4.29$, $p < 0.001$, $d = 0.78$, but reported less change than expected by chance when only appearance changed, $t(29) = 2.15$, $p = 0.040$, $d = 0.39$.

To gain a finer-grained understanding of broader patterns in children's responding we also examined the response profiles of individual children. We classified children as *mutual intention responders* when they met three conditions: they reported change in a majority of intention-only trials, reported change in a majority of intention + appearance trials, and reported no change in a majority of appearance only trials. On the assumption of chance responding this would be one of eight possible answer patterns and so quite infrequent (12.5%), but for our purposes it reflects viewing changes in mutual intentions alone as constituting group membership, and so is on our hypothesis the pattern we would expect to appear most frequently. Other possible response patterns that we considered were *appearance-based responders* (reporting change whenever appearances changes and not reporting change when it doesn't) and *no-change* or *always-change responders* (generally reporting no change or change irrespective of trial type). Strongly suggesting the presence of principled strategies, 57% of younger children fell into the mutual intention pattern, far more than expected by chance, $\chi(1, N = 30) = 11.06$, $p < 0.001$, OR = 9.15. No other answering pattern appeared more often than chance: Only 10% of children conformed to an appearance-based strategy, 3.33% conformed to a no-change strategy, 20% to an all-change strategy, and the remaining 6.66% reported change only when appearance and intentions both changed.

3.2.2. Older children - 7–9 year-olds

A one-way ANOVA also revealed an effect of condition, $F(2,42) = 5.67$, $p = 0.007$, $\eta^2 = 0.27$. Children reported that the character changed groups far more when intentions and appearance changed than when only appearance changed, $t(15) = 7.89$, $p < 0.001$, $d = 2.61$, and far more when intentions changed than when only appearance changed, $t(15) = 9.68$, $p < 0.001$, $d = 3.23$. As predicted, older children reported that the character changed equally often when intentions and appearance changed as when only intentions changed, $t(15) = 1$, $p = 0.333$., $d = 0.11$, suggesting that for older children visual cues do not increase their confidence in a change mediated by mutual intentions.

As for older children's individual strategies, 81% of children were mutual intention responders. This is far greater than expected by chance, χ (1, N = 16) = 12.55, $p < 0.001$, OR = 30.33. No other strategy appeared more often than chance. No children conformed to an appearance-based strategy, 6.25% conformed to a no-change strategy, 6.25% to an all-change strategy, and 6.25% reported change only when intention *alone* changed.

3.3. Discussion

These results strongly support the causal role of mutual intentions. Across both age groups, children readily believed that individuals could change their group membership. Children were selective, however, about the conditions under which they believed group membership could change. Children only believed group membership changed when there were mutual changes in intentions (both the individual and the opposite group viewed the individual as a member of the opposite group). When the characters put on the opposite group t-shirt in the absence of mutual intentions, not even younger children believed that group change was plausible. Therefore, children believe that mutual intentions constitute group membership.

Experiment 2 also hints at some modest shifts in behavior as a function of age, constituting a strengthening of this general pattern. While both age groups viewed mutual intentions as more important than appearance, and as both necessary and sufficient for category change, younger children were more likely to factor appearance into their decisions. That is, younger children reported more appearance-based change than older children (though they remained below chance), and younger children were especially persuaded by the possibility of group change when mutual intentional changes were paired with appearance change. By contrast, for older children, appearance cues did not add anything once information concerning mutual intentions had been established. Overall, however, even the youngest children were reliably able to look past salient visual markers and focus on the role of intentions, and the interaction between age and condition did not reach significance. Therefore, our results suggest continuity rather than dramatic change.

Critically, while the social categories we have focused on so far are ones that children interpret as socially rich (for example marking obligations and allegiances; Rhodes & Chalik, 2013), absent

additional input, they do not essentialize them (Rhodes et al., 2012). Still, a central features of social categories is that children sometimes do apply an alternative causal framework of essentialism. Recent work has suggested that social categories are not essentialized by default. That is, essentialism is selectively applied and appears to require additional input such as generic language (Rhodes et al., 2012). Supporting this contention, children interpret racial categories as socially relevant long before they essentialize them, though they generally do come to essentialize them later in life, as do adults (Haslam et al., 2000; Rhodes, 2012; Shutts et al., 2013). The one social category for which essentialism appears early and powerfully, however, is gender (Taylor, 1996; Taylor et al., 2009). Gender thus provides an ideal comparison case: a social category to which children apply an alternative causal framework, and to which they might therefore no longer rely on mutual intentions in determining category membership.

Previous literature, pioneered by Kohlberg (1966), explores whether and when children believe that gender is stable across the lifespan and cannot change. One method used in past research is transformation-like paradigms where boys put on girl clothes or girls put on boy clothes (e.g., Bem, 1989). Generally, this research finds that children in our younger age range (4–5 years) sometimes report that a person's gender can change. By our older age range (7–9 years), children robustly report that gender cannot change. Previous research has never explored the role of intentions. Therefore, it is not clear whether younger children entertain the possibility of gender change because they care about mutual intentions or because of other beliefs (Ruble et al., 2007) or cognitive deficits in inferring kind membership from kind-typical properties (Gelman et al., 1986). Therefore, we sought to test whether children withhold a mutual intentional framework to gender, a category that they already richly essentialize in this age range.

4. Experiment 3

4.1. Method

4.1.1. Participants

We focus on the same two age groups and sample sizes as Experiment 2 in order to promote cross-study comparisons.16 children participated in the older age range (M = 8.67, SD = 0.81) and 30 younger children participated in the younger age range (M = 4.96, SD = 0.66). 25 males and 21 females participated.

4.1.2. Design and procedure

The design and procedure were largely identical to experiment two except children were presented to a group of boys and a group of girls instead of Flurps and Zazzes. To ensure that the appearance change accorded with children's intuitions about gender, individuals changed their gendered appearance (i.e., hair, clothing style, and color of clothes) in the trials involving appearance change rather than merely changing t-shirts.

4.2. Results

The primary results are presented in Fig. 3. We analyzed results with a 2 (Age group: 4–5 vs. 7–9) by 3 (Condition, appearance + intentions, intentions-only, appearance-only) mixed-design ANOVA. There was a main effect of age group, $F(1, 129) = 18.17$, $p < 0.001$, $\eta^2 = 0.12$, such that older children reported less gender change than younger children, $t(42.09) = 4.048$, $p < 0.001$, $d = 1.17$. There was no significant effect of condition, $F(2,129) = 1.09$, $p = 0.34$, $\eta^2 = 0.01$. There was no significant two-way interaction, F (2, 129) = 0.38, $p = 0.688$, $\eta^2 = 0.0050$. To facilitate comparison with prior studies we again present analyses broken down by age group; sample sizes are equal for all age-specific analyses.

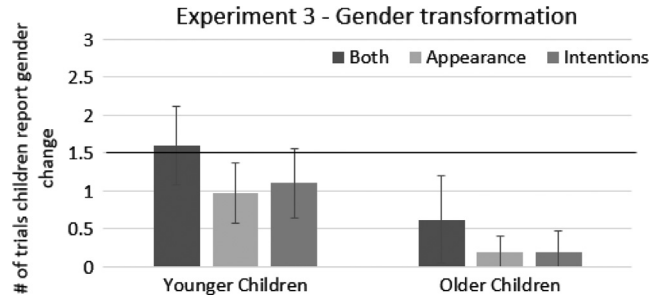

Fig. 3. Average number of trials (out of 3) that children reported a character changed gender by age group and condition (both appearance and intention change, appearance-only change, and intention-only change). Error bars are 95% confidence intervals.

4.2.1. Younger children - 4–5 year-olds

There was no effect of condition for younger children, $F(2,84) = 0.668$, $p = 0.516$, $\eta^2 = 0.02$. Overall, younger children reported slightly less gender change than expected by chance (40.74%), $t(89) = 2.12$, $p = 0.037$, $d = 0.22$.

Next we examined individual-level data, focusing again on the same three strategies as we did in experiment two (mutual intentional, appearance-based, and a no change/essentialist strategy). This most common strategy was reporting no change across all conditions (37%). We interpret this pattern of responding as indicative of an essentialist causal framework because essentialist categories are generally immutable (as essences are not easily changed); therefore, essentialist categories are not changed by changes in intentions or appearances. The frequency of essentialist responses was marginally more common than expected by chance, χ (1, N = 30) = 3.51, $p = 0.061$, OR = 4.05; this response profile was also considerably more common than what occurred in Experiment 2, where they appeared only 3.33% of the time, χ (1, N = 60) = 8.44, $p = 0.003$, OR = 16.79. The number of mutual intentional responses (26.7%) was no different than chance, χ (1, N = 30) = 0.57, $p = 0.449$, OR = 1.14. Furthermore, it was significantly less common than it was in Experiment 2 (60%), χ (1, N = 60) = 5.50, $p = 0.019$, OR = 4.13. Finally, 16.67% of children conformed to an appearance-only strategy, which did not differ from chance, χ (1, N = 30) = 0.0084, $p = 0.927$, OR = 1.40.

4.2.2. Older children - 7–9 year-olds

Again, there was no significant effect of condition, $F(2,42) = 1.42$, $p = 0.254$, $\eta^2 = 0.067$. Older children reported far less gender change than expected by chance (11.11%), $t(47) = 10.73$, $p < 0.001$, $d = 1.55$.

Turning towards individual-level data, 75% of older children reported no change cross all three conditions (i.e., an essentialist strategy), which is significantly more often than expected by chance, χ (1, N = 16) = 10.29, $p = 0.001$, OR = 21.00. No other strategy appeared more often than expected by chance. Furthermore, this was significantly more essentialist responses than appeared in Experiment 2, where only 6.25% of children conformed to that pattern, χ (1, N = 32) = 12.96, $p < 0.001$, OR = 21.00.

4.3. Discussion

Despite heavy reliance on mutual intentions for social groups in Experiments 1 and 2, younger and older children generally rejected mutual intentions as causally relevant to gender, though there was some evidence that younger children were occasionally moved by

them. However, given the lack of differences between conditions, it is also possible that children were simply more likely to endorse the possibility of group change irrespective of condition. More broadly, both age groups appeared to rely on an essentialist strategy, generally rejecting the possibility of gender change. Indeed, children treated changes in mutual intentions as similarly irrelevant to gender status as changes to superficial appearance. This pattern of results is similar to how children treat animals in the transformation paradigm, rejecting change to category status even after dramatic external transformations (Keil, 1992).

The cross-study comparison further emphasizes children's different approaches to different social categories. Children were much less likely to apply a mutual intentional strategy and much more likely to apply an essentialist strategy to gender than they were to a novel social group. This is consistent with past claims that gender essentialism is early emerging and relatively robust (Taylor, 1996; Taylor et al., 2009) and with findings concerning children's tendency not to essentialize simple novel groups such as those used here (Rhodes et al., 2012). Critically, then, the difference across studies demonstrates theoretical specificity: mutual intentions are critical to novel social groups (which children tend to treat as social coalitions; Rhodes, 2012) but not to essentialized social kinds.

Interestingly, we find that younger children's responses are considerably less essentialist using the transformation paradigm than are older children's responses. This is consistent with previous research investigating beliefs about gender constancy. In tasks that involve changes in appearance, very young children tend to report more gender change than older children (Bem, 1989; Ruble et al., 2007). This contrasts with other measures of gender essentialism where gender essentialism decreases with age. Specifically, tasks that examine beliefs about whether gendered properties (e.g., a boy's preference for trucks versus dolls) are flexible and sensitive to environmental input find that older children accept more flexibility in preferences than younger children (e.g., Taylor, 1996). This dissociation may stem from children having difficulty inferring category membership from properties (like gendered appearance) compared to inferring properties from category membership (Gelman et al., 1986). Furthermore, children's essentialist beliefs may actually make understanding gender variance difficult. Because younger children are inflexible about gender stereotypes and infer that gendered properties are inherent (e.g., caused by "girl-ness") rather than caused by the environment, they may find gender inconsistency more confusing that older children (Ruble et al., 2007).

5. Experiment 4

Study 1 and Study 2 demonstrate that children believe that group membership has a strongly intentional basis. Children believe that individuals join groups only when they and other group members consensually intend for them to join (Study 1). Further, children believe that individuals' group statuses change when there are changes in intentions even when other salient group membership cues remain unchanged (Study 2).

Recall, however, that mutual intentions require not *only* that people privately (or coincidentally) share their intentions about an object or individual. *Mutual* intentions in the sense used here also requires common knowledge of shared intentions. Only under conditions of common knowledge does the social role (of group membership, for example) become assigned to an individual. By analogy, consider the case of being engaged. If two individuals who have been dating for a long time both want to become engaged, the mere coincidence of their intentions to get engages does not make them engaged. The individuals only acquire the social status of "fiancés" when they make their intentions known to each other (e.g., "Will you marry me?" "Yes!").

In Study 1 and 2 all of the intentions *are* publicly expressed, and thus these results are consistent with the importance of common knowledge. However, these results are also consistent with a slightly different and more general causal belief. That is, perhaps when assigning group statuses children only care about the alignment of intentions, whether that alignment is publicly known or not. Or even more generally, perhaps children believe the mere overlap of *any* psychological property is sufficient to confer group status. To examine whether children's causal theory is more specialized, in the sense of also requiring shared knowledge of the intentions, in Experiment 4 we directly investigate the role of common knowledge in children's causal theories of social groups. Thus, we compare contexts identical to the *Mutual* condition of Study 1, where intentions are shared and publicly stated (*Shared Intentions + Common Knowledge*), to contexts where intentions are privately and coincidentally shared but not publicly stated (*Shared Intentions - Common Knowledge*).

5.1. Method

5.1.1. Participants

We focus on the same two age groups as Experiment 2. Piloting with adults suggested the role of common knowledge may be a subtler distinction, so we opted for the larger sample size for both age groups (30 children).

5.1.2. Design and procedure

The basic procedure of the experiment was largely similar to the *Mutual* condition of Study 1. Children were introduced to a novel group of Flurps called the "Flurp group." Children then encountered 6 characters who wanted to join the group. The characters *always* expressed their intent to join and the group members *always* expressed their intent for the character to join. Indeed, the characters always declared that the character was a member of the group. In each context children were asked whether the character was a Flurp or not a Flurp. What varied in this experiment was whether these verbal declarations of intentions and group status occurred in the same room (*Public condition*) or in separate rooms (*Private condition*), as two within-subject conditions (Fig. 4). This allowed the stimuli that we presented to children to be remarkably similar across conditions. In both conditions the character stands a similar distance away from the group and is socially unengaged. The character and group members make the same verbal statements about their desires. The only thing that differs is that in the *public* condition we highlight the presence of mutual knowledge and in the *private* condition we highlight the lack of that mutual knowledge. That is, we either say "*But look – s/he and these Vawns are in different rooms so they can't hear each other*" or "*And look – they are in the same room. So, these Vawns know s/he wants to be (in the group).*"

5.2. Results

The primary results are presented in Fig. 5. We analyzed results with a 2 (Age group: 4–5 vs. 7–9) by 2 (Condition, *public* vs. *private*) mixed-design ANOVA. There was a main effect of condition, $F(1,59) = 101.76$, $p < 0.001$, $\eta^2 = 0.38$, such that children were more likely to categorize a character as a Flurp in the *public* condition ($M = 2.70$, $SD = 0.80$) than in the *private* condition ($M = 1.03$, $SD = 1.32$). There was no main effect of age group, $F(1,59) = 0.115$, $p = 0.78$, $\eta^2 = <0.001$; however, there was a significant interaction between age group and condition, $F(2,59) = 7.49$, $p = 8.18$, $\eta^2 = 0.04$. Visually (by inspecting Fig. 5), the interaction seems to result from a strengthening of pattern across the age

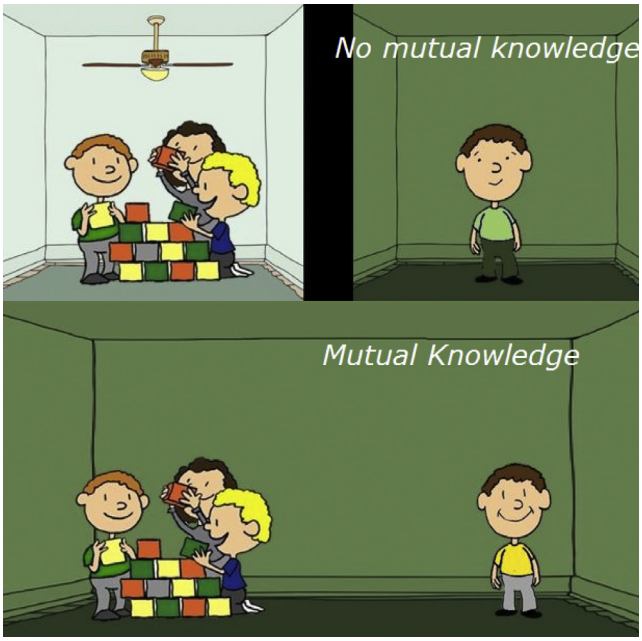

Fig. 4. Comparison of stimuli across public and private conditions of Experiment 4. In the *private* condition (top), the group and individuals have no knowledge of each other's intentions. This is because the wall prevents them from hearing each other. In the *public* condition (bottom) the characters can hear each other and thus have knowledge of each other's intentions.

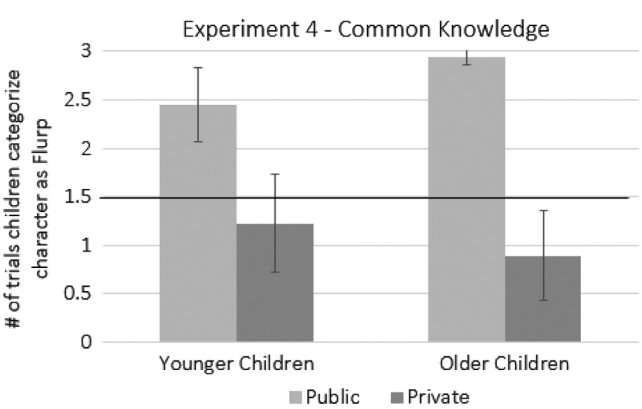

Fig. 5. Children's categorization of a character as a member of the group across two conditions. In the *public* condition there was no common knowledge, it in *private* condition there was. Error bars are 95% confidence intervals.

groups, such that older children differentiated the conditions more strongly than younger children. Indeed, the participant-level difference score between the two conditions is higher for older children than younger children, $t(58.7) = 2.74$, $p = 0.008$, $d = 0.70$.

5.2.1. Younger children - 4–5 year-olds

We analyzed younger children's categorization using a one-way ANOVA on condition (*public* vs. *private*). There was the expected main effect of condition, $F(1,30) = 25.22$, $p < 0.001$, $\eta^2 = 0.20$, such that children were more likely to categorize a character as a Flurp in the *public* condition ($M = 2.45$, $SD = 1.06$) than in the *private* condition ($M = 1.23$, $SD = 1.38$). In the public condition, children categorized a character as a Flurp significantly more often than expected by chance, $t(30) = 5.00$, $p < 0.001$, $d = 0.90$. In the private condition, children were at chance, $t(30) = 1.10$, $p = 0.279$, $d = 0.199$.

Inspection of the data suggested pronounced developmental change in this age range, so we performed an unplanned exploratory analysis investigating age effects within the younger sample. A 1 (Age) by 2 (Condition, *public* vs. *private*) mixed-design ANOVA revealed a significant two-way interaction, $F(1,29) = 7.87$, $p = 0.009$, $\eta^2 = 0.06$. The median age was 5.08, so we broke-down the condition effect for children older than or equal to this age (M age = 5.49, $N = 14$) and children younger than this age (M age = 4.50, $N = 17$). The 5-year-olds ($M = 2.00$, $SD = 1.30$) were more likely to distinguish between condition than the 4-year-olds ($M = 0.588$, $SD = 1.06$), $t(25.086) = 3.26$, $p = 0.003$, $d = 1.19$. The 5-year-olds ($M = 2.64$, $SD = 0.93$), $t(13) = 4.60$, $p < 0.001$, and 4-year-olds, $t(16) = 2.82$, $p = 0.012$, were both above chance in the *public* condition. By contrast, in the *private* condition, only 5-year-olds were actually below chance ($M = 0.643$, $SD = 1.15$), $t(13) = 2.79$, $p = 0.015$, whereas 4-year-olds were (non-significantly) above chance ($M = 1.71$, $SD = 1.40$), $t(16) = 0.605$, $p = 0.554$. The strength of these results support the possibility that children do not appreciate the necessity of common knowledge until 5 years of age; the interpretation of these results should be cautious, however, because they are exploratory and low in sample size.

To examine individual-level data, we categorized participants into three groups, those who categorized a character as a Flurp more often in the public condition (expected distinguishers), those who did not distinguish between condition (non-distinguishers), and those who categorized a character as a Flurp more often in the private condition (unexpected-distinguishers). Younger children either distinguished in the expected direction (51.61%) or did not (48.39%). No younger child (0%) ever distinguished in the unexpected direction. Overall, this distribution is significantly different than chance, χ (2, N = 31) = 12.7, $p = 0.002$.

Using the same age breakdown reported above for exploratory purposes, the same developmental patterns are observed 78.57% of 5-year-olds distinguished between conditions in the expected direction, whereas only 29.41% of 4-year-olds did so. This difference was significant, χ (1, N = 31) = 5.59, $p = 0.018$, $OR = 8.80$. This supports the conclusion that 4-year-olds do not yet appreciate the importance of common knowledge, whereas 5-year-olds do.

5.2.2. Older children – 7–9 year-olds

We analyzed older children's categorization using a one-way ANOVA on condition (*public* vs. *private*). There was the expected main effect of condition, $F(1,29) = 91.09$, $p < 0.001$, $\eta^2 = 0.62$, such that children were more likely to categorize a character as a Flurp in the *public* condition ($M = 2.97$, $SD = 0.18$) than in the *private* condition ($M = 0.833$, $SD = 1.23$). In the public condition, children categorized a character as a Flurp significantly more often than expected by chance, $t(29) = 44$, $p < 0.001$, $d = 8.17$. In the private condition, children were below chance, $t(29) = 2.96$, $p = 0.006$, $d = 0.833$. (In the interests of completeness, there was no hint of further age-related change *within* this age group, again via median split, $F(1,28) < 0.001$, $p = 0.99$, $\eta^2 < 0.001$.)

We next examined individual-level data using the same three categories (expected distinguishers, non-distinguishers, unexpected distinguishers). 80% of children were expected distinguishers, categorizing a character as a Flurp more often in the *public* condition than in the *private condition. 20%* of children were non-distinguishers and 0% were unexpected distinguishers. This distribution was significantly different than chance, χ (2, N = 30) = 16.65, $p < 0.001$.

5.3. Discussion

We find support for our predictions in both age groups. Children systematically restricted their attributions of group status to conditions where there were shared intentions *and* common knowl-

edge. Children did not believe that a character had joined a group simply because the character coincidentally shared the same intentions as the members of the other groups. Rather, these intentions had to be known by each party, reflecting the truly *social* nature of this causal pathway.

These results powerfully support the mutual intentionality account proposed here. The importance of common knowledge is uniquely predicted by our account. These data are incompatible with a more simplified account based in the coincidental overlap of intentions or other shared psychological or internal states. If these alternative accounts were correct, children would have attributed group status in both condition. In each condition the character and group members all shared the same intentions, and all wanted the character to be a group member. Indeed, they all verbalized these very intentions. However, only when these intentions are said in the presence of each other, and thus are mutually known by each party, do those intentions confer group status.

Of interest, our results suggest young children (as young as 5-years-old) may understood performative speech acts (Austin, 1975). In both conditions, the group utters an assertion about the world "C (character's name) is a Flurp." However, by age 5, children recognized that this speech act only conferred group status under the right conditions (when the utterers knew the intentions of C, and C could hear the speech act). Performative speech acts play an important role in Searle's (2009) theory of social construction, and children's understanding of them would be a natural extension of the mutual intentionality framework outlined here. To fully document an understanding of performative speech acts would require more conditions; however, we believe these data provide a strong foundation for developing research to test performative speech acts in children. To our knowledge no such research exists to date.

6. General discussion

Together these results support the hypothesis that young children believe that mutual intentions are causally important for and even causally constitutive of novel social groups. Children readily applied this causal framework whether the groups were visually marked or not and whether or not visual properties were consistent with or conflicted with the new category membership. Moreover, children were sensitive to quite subtle distinctions, such as whether the intentions were known by all parties. Across all four experiments we observed developmental solidification of early-emerging tendencies rather than substantial qualitative change. This suggests that mutual intentionality plays a foundational role in children's intuitive sociology, in that it is early emerging, readily applied to newly encountered social groups, and persists throughout childhood.

Importantly, we also demonstrate that even our younger group of children did not apply a mutual intentional causal framework to a familiar and essentialized social category, gender. This is consistent with past work showing that gender essentialism is early emerging (e.g., Taylor, 1996). Our work is therefore consistent with the perspective that children have two distinct intuitive theories that they draw on when reasoning about social categories (Rhodes, 2012), one coalitional and one essentialist. We clarify this past work by demonstrating that the causal framework applied in the (non-essentialist) coalitional case is based in mutual intentions.

Thus, children believe social groups come into existence *socially*, through the mutual (and mutually understood) intentions of multiple individuals. Once such groups are socially constituted in this manner, children believe that such groups license a variety of relational and moral predictions, including beliefs about patterns of friendship/association (Shutts et al., 2013) and patterns of moral (Rhodes & Chalik, 2013) and conventional (Schmidt, Rakoczy, & Tomasello, 2012) obligations. We suggest that at least some of these predictions may follow from or be licensed by the causal role of mutual intentions. Group norms are based in mutual intentions (Rakoczy & Schmidt, 2013; Schmidt, Rakoczy, Mietzsch, & Tomasello, 2016) and thus can be constituted by the same causal pathway as group membership; research suggests that children recognize that group norms are context- and community-bound (Rakoczy & Warneken, 2008; Schmidt et al., 2012; Wyman, Rakoczy, & Tomasello, 2009) and may even believe they are socially constructed (Schmidt & Tomasello, 2012; Schmidt et al., 2016; Turiel, 1983). As for the predictions concerning social alliances, even infants understand that intentions influence social interactions and social bonds (Choi & Luo, 2015); therefore, children may readily infer that shared intentions imply friendship relationships or additional intentions to support one another. Taken together, these findings provide initial support for the claim that mutual intentions structure some of the predictions that children make about social groups. Nevertheless, more work is needed to elucidate whether and how these predictions are related to the causal role of mutual intentions. For example, it is possible that some social predictions are fundamental enough to social group reasoning to apply irrespective of the causal framework at play (i.e., to apply to both mutual intention and essentialist groups).

While our work suggests that children apply a mutual intentional framework to newly encountered social groups and withhold it from gender, we have certainly not exhausted the myriad range of social categories that exist. Perhaps there are social categories that are neither based in mutual intentions nor essentialism (or which are not easily classifiable); for example, children treat language as both more stable (Kinzler & Dautel, 2012) and richly social (Kinzler, Shutts, DeJesus, Spelke, 2009; Rhodes & Chalik, 2013) than race. But belonging to a language group depends on the deployment of a specific set of competencies, namely language ability. It is therefore unclear whether *either* an essentialist or a mutual -intentional stance fully accounts for children's beliefs about language ability. Thus, language may present a particularly important test case for future research. At the very least, other essentialized social categories, particularly novel groups towards which essentialism has been induced (Rhodes et al., 2012), are worth testing in the paradigm employed here: as essentialism increases, belief in the causal role of mutual intentions should decrease. Finally, while we believe many social groups are based in *mutual intentions*, some social categories may be based in individual intentions; for example, categories based in individual preferences, hobbies, and beliefs may rely solely on an individual's mental states. Exploring whether children respect this distinction at different points in development would be valuable.

Nevertheless, we believe that a causal theory based in mutual intentions underlies many social groups. Further, some of the apparent diversity among social groups may stem from differences in how mutual intentions are structured rather than differences in their causal nature *per se*. For instance, one interpretation of the causal origins of occupations like doctors is mutual intentions: a person is a doctor both because of their intentions to practice medicine and their recognition by others (such as medical institutions and boards) that they are a doctor (Searle, 1995). The difference between occupations like doctors and the simpler social groups presented here may be the institutional complexity that governs their mutual intentions. For example, doctors are only recognized after receiving extensive training and can be certified only through very particular routes; the mutual intentions that grounds their status operates via the mutual sanctioning of these institutional procedures; an adult-like understanding of social groups requires

recognizing these formal structures. Future work should investigate children's developing understanding of these richer institutional relationships, perhaps by first exploring children's understanding of pledges and initiation rituals which are the simplest way of formalizing mutual intentions.

In conclusion, while much work has focused on the social categories that children view as natural (Gelman, 2003), more work is needed to understand the kinds that children believe are socially constructed (Rhodes & Gelman, 2009), especially given that these represent many or even most social groups. Here we propose and empirically support a causal pathway by which children believe socially constructed kinds can come into existence and by which category memberships can emerge and change: mutual intentions. Our findings, in combination with many others, support social groups as a distinct conceptual domain. These domain-specific expectations prepare children for attending to and navigating the unique demands of the social world, sorting through its multitudes of allegiances, obligations, and conventions. Indeed, through an understanding of mutual intentions, children can come to appreciate how this social world is shaped and created by the mental states they share with others.

Supplementary materials

Complete scripts and data files can be accessed at the Open Science Framework: https://osf.io/nzzcj/#.

Acknowledgements

We thank the members of the Social Cognitive Development Lab for their contributions to the project. We also thank Marjorie Rhodes for providing us with the stimuli set adapted here.

References

Aboud, F. E. (1988). *Children and prejudice*. New York, NY: Blackwell.
Austin, J. L. (1975). *How to do things with words*. Cambridge, MA: Harvard University Press.
Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*, 110–118.
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.
Bem, S. L. (1989). Genital knowledge and gender constancy in preschool children. *Child Development*, 649–662.
Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition, 60*, 1–29.
Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112*, 337–342.
Choi, Y. J., & Luo, Y. (2015). 13-month-olds' understanding of social interactions. *Psychological Science, 26*, 274–283.
Diesendruck, G., & HaLevi, H. (2006). The role of language, appearance, and culture in children's social category-based induction. *Child Development, 77*, 539–553.
Diesendruck, G., Goldfein-Elbaz, R., Rhodes, M., Gelman, S., & Neumark, N. (2013). Cross-cultural differences in children's beliefs about the objectivity of social categories. *Child development, 84*, 1906–1917.
Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development, 82*, 793–811.
Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*, 65–95.
Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press.
Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development, 57*, 396–404.
Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39*, 113–127.
Hirschfeld, L. A. (1998). *Race in the making: Cognition, culture, and the child's construction of human kinds*. Cambridge, MA: MIT Press.
Katz, P. A. (1983). Developmental foundation of gender and racial attitudes. In R. Leahy (Ed.), *The child's construction of social inequality* (pp. 41–78). New York: Academic Press.
Kalish, C. W., & Lawson, C. A. (2008). Development of social category representations: Early appreciation of roles and deontic relations. *Child Development, 79*, 577–593.
Keil, F. C. (1992). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
Kinzler, K. D., & Dautel, J. B. (2012). Children's essentialist reasoning about language and race. *Developmental Science, 15*, 131–138.
Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition, 27*, 623.
Kohlberg, L. A. (1966). A cognitive-developmental analysis of children's sex role concepts and attitudes. In E. E. Maccoby (Ed.), *The development of sex differences* (pp. 82–173). Stanford, CA: Stanford University Press.
Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255–258.
Rakoczy, H., & Schmidt, M. F. (2013). The early ontogeny of social norms. *Child Development Perspectives, 7*, 17–21.
Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology, 44*, 875.
Rhodes, M. (2012). Naïve theories of social groups. *Child Development, 83*, 1900–1916.
Rhodes, M., & Brickman, D. (2011). The influence of competition on children's social categories. *Journal of Cognition and Development, 12*, 194–221.
Rhodes, M., & Chalik, L. (2013). Social categories as markers of intrinsic interpersonal obligations. *Psychological Science, 24*, 999–1006.
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, 109*, 13526–13531.
Ruble, D. N., Taylor, L. J., Cyphers, L., Greulich, F. K., Lurye, L. E., & Shrout, P. E. (2007). The role of gender constancy in early gender development. *Child development, 78*, 1121–1136.
Schmidt, M. F., & Tomasello, M. (2012). Young children enforce social norms. *Current Directions in Psychological Science, 21*, 232–236.
Searle, J. R. (1995). *The construction of social reality*. New York, NY: The Free Press.
Searle, J. R. (2006). Social ontology: Some basic principles. *Anthropological Theory, 6*, 12–29.
Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition, 124*, 325–333.
Schmidt, M. F., Rakoczy, H., Mietzsch, T., & Tomasello, M. (2016). Young children understand the role of agreement in establishing arbitrary norms—but unanimity is key. *Child Development, 87*, 612–626.
Shutts, K., Roben, C. K. P., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development, 14*, 35–62.
Solomon, G. E. A. (2002). Birth, kind and naïve biology. *Developmental Science, 5*, 213–218.
Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development, 67*, 1555–1571.
Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development, 80*, 461–481.
Tomasello, M. (1992). The social bases of language acquisition. *Social Development, 1*, 67–87.
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675–691.
Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to mutual intentionality. *Mind & Language, 18*, 121–147.
Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, MA: University Press.
Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology, 43*, 337–375.
Wyman, E., Rakoczy, H., & Tomasello, M. (2009). Normativity and context in young children's pretend play. *Cognitive Development, 24*, 146–155.