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
Available online 24 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.

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/

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-