American Psychological Association

Journal of Experimental Psychology: General

ISSN: 0096-3445

2020, Vol. 149, No. 7, 1344–1359
http://dx.doi.org/10.1037/xge0000711

Collective Recognition and Function in Concepts of Institutional Social Groups

Alexander Noyes and Frank C. Keil
Yale University

We investigate ordinary concepts of institutional groups: stable, cooperative, and socially constructed entities like clubs, companies, and academic departments. We use a transformation paradigm to examine participants' causal beliefs about how groups exist and persist over time. We consider whether participants believe groups are grounded in collective recognition or function. Participants' default views about groups see them as persisting because the members or a relevant third-party collectively recognize the members as belonging to a group (Studies 1–4). Social groups are dual-character though (Studies 5–8). There is a second sense: the *true* group. This *true* judgment is grounded in whether the group realizes its basic function. This sense is more influenced by participants' own ideological commitments. Thus, participants can disagree about whether a group *truly* exists even if they agree about the bare facts. We discuss implications for theories of conceptual representation.

Keywords: concepts, identity, social groups, causal reasoning

Social groups are fundamental to the lives of humans. In many modern societies, we are exposed to countless institutional groups. A graduate student in psychology will contend with multiple layers of institutional groups during her academic work (*Class of 1989, Cognitive Area, Department of Psychology, Yale University*). She will vote in, pay taxes to, or follow the rules of her town, state, and nation of residence (*Atkinson*, *NH*, *U.S.A.*). She will interact with various companies (*Google, Amazon, and Facebook*), and academic organizations (*NSF*, *APA*, *SPSP*). Here, we investigate how participants mentally represent these entities, and in particular, how they conceptualize their causal structure—how do groups exist, what is their nature, and what underlies their persistence over time?

We define institutional groups as entities that are relatively stable, cooperative, and socially constructed. This contrasts with temporary collectives like mobs and audiences. It also contrasts with groups grounded in shared similarities, like "blue-eyed people" (see Rabbie & Horwitz, 1988 for a similar distinction). Teams, companies, schools, and nation-states are all stable, cooperative, socially constructed entities. We exclude entities like *White people* that although partially fit these criteria involve notably different processes (see Mallon, 2016)—namely, that these entities are constructed because people mistakenly believe they are natural kinds.

Theoretical Framework and Rationale

Institutional social groups are widespread and consequential. Even though our lives are dominated by interactions with institutional groups, this pervasive facet of ordinary cognition has been barely studied from the perspective of conceptual representation and causal reasoning. Institutional groups, along with other entities such as money and social roles, constitute a conceptual domain with unique causal architecture. Theories of concepts have been centrally concerned with which types of concepts are represented using nonobvious causal processes, and which concepts are represented using mere similarity (Bloom, 1998; Keil, 1989; Gelman, 2003, 2013; Malt & Sloman, 2007). Though institutional entities are often assumed to lack interesting causal structure (see Keil, 1989), philosophical theories explain institutions via rich, causal structure (see Guala, 2016 and Millikan, 1998) and thereby suggest a more elaborated set of causal beliefs underlying institutional concepts. This possibility raises the related question of the extent to which people are systematically biased toward certain causal interpretations of the world, such as those that favor internal or inherent features (Cimpian & Salomon, 2014; Gelman, 2003). Institutional kinds are ordinary examples of social construction; their properties depend on external social arrangements rather than their internal or inherent properties (Guala, 2016; Searle, 2010). Therefore, institutional kinds are an important test case for theories of conceptual representation; they can show us how participants reason about social construction, and to what extent social construction is entertained in lay theories. Consequently, studying

This article was published Online First November 7, 2019.

Alexander Noyes and Frank C. Keil, Department of Psychology, Yale University.

Alexander Noyes is funded by a fellowship from the Graduate Research Fellowship Program of the National Science Foundation. The funding source had no involvement in the design, nor in the collection, analysis, and interpretation of data, nor the writing and submission of the article. The data and ideas appearing in the article have not been published previously, and have not been disseminated in any other form previously. The authors have no conflicts of interest to disclose. All studies were conducted under Protocol 1311013027 approved by Yale's IRB Board.

Correspondence concerning this article should be addressed to Alexander Noyes, Department of Psychology, Yale University, Box 208205 [Courier: 2 Hillhouse Avenue], New Haven, CT 06520-8205. E-mail: Alexander.Noyes@yale.edu

institutional kinds may prompt critical revisions of prior assumptions about human concepts. Thus, institutional social groups are an ecologically significant set of entities that can provide greater insight into the nature of ordinary conceptual representation.

How do intuitional groups exist? Philosophers have taken two broad perspectives that may correspond to participants' intuitive theories of institutional groups. The first is collective recognition (e.g., Gilbert, 2013; Searle, 1995, 2010). A group exists because its members or others (some relevant third-party) recognize the individuals to be a group. Though this proposition might seem trivial in one sense, it suggests groups differ fundamentally from other types of entities in the world that exist independently of human being's beliefs (i.e., natural kinds like tigers and gold; and to an extent standard artifact: see Noyes, Keil, & Dunham, 2018). This approach also helps to explain the normative dimensions of groups. Why do members of a reading group feel obligated to show up to meetings and to read the assigned novels? Perhaps members recognize themselves as a cooperative entity with joint commitments to one another (Gilbert, 2013). Thus, a group continues to exist so long as individuals (members or a relevant third-party) collectively recognize its existence.

An alternative approach argues that institutional groups are defined by the functions they realize (Guala, 2016). Institutions develop in response to coordination problems. One coordination problem is that individuals want to read with other people, but to accomplish this goal they need to read the same books and find time to meet. The reading group solves this problem. Therefore, the function of a reading group is causally antecedent to the members collectively recognizing themselves as a group. Functional accounts capture why groups exist in the first place and explain why institutional groups are often characterized by their functions (Rosenberg, 2016). According to this approach, a reading group exists so long as people continue to read together and discuss literature. This perspective of social institutions also assumes unique causal architecture, with institutions referring to stable patterns of human behavior, rather than to fixed properties like physical structure (Guala, 2016).

People might also believe that both collective recognition and function underlie groups—just in different ways. Certain concepts are "dual character," meaning that there are actually two distinct senses of the concept present simultaneously (Knobe, Prasada, & Newman, 2013). *Scientist* is one example of a dual-character concept: Participants judge that a career scientist that cares nothing about science is "technically" a scientist, and a hobbyist with a deep passion for science is a "true" scientist. These are not mere hedges (Malt, 1990). People can refer to either sense without a qualifier; the qualifiers rather help to make explicit which sense is being referred to (Leslie, 2015). Thus, the concept of scientist has two different senses (the "technical" and the "true"), and someone can be one or the other or both. A "technical" scientist appears to be someone who is collectively recognized as a scientist (they occupy the official institutional position). A "true" scientist is someone that realizes the basic function of a scientist (Leslie, 2015), or at least is committed to realizing it (Del Pinal & Reuter, 2017). By extension, institutional groups might be dual-character. Institutional groups might have two senses of continuity, the sense in which they continue to be collectively recognized, and the sense in which they continue to realize the same basic function.

Notably, all three of these accounts differ from intuitive theories of natural kinds and artifacts (Guala, 2016). Institutional groups clearly differ from natural kind concepts: People believe that natural kinds are grounded in microstructural properties like genes and chemical composition (Gelman, 2003; Keil, 1989). But institutional entities also differ from standard artifacts like chairs and hammers. People believe standard artifacts are grounded in their physical properties and the intention the creator had when they built it (Bloom, 1996, 1998; Keil, 1989). Therefore, artifacts depend on people in a historical sense (they had to be made), but then they persist independently of how people currently use them or regard them (Bloom, 1996; Noyes, Keil, & Dunham, 2018). Institutional groups depend on ongoing properties: Either in the continual recognition of the group (i.e., collective recognition account), or in the ongoing enactment of the group's function (i.e., function account). Institutional groups are continuously made to exist. They are socially constituted in a way that natural kinds and artifacts are not. Institutional groups cannot be subsumed into prior theories of conceptual representation. Therefore, insofar as participants concepts conform to one of these three patterns, institutional social groups may reveal additional flexibility and diversity in people's intuitive theories.

In one prior study on causal reasoning about social groups, participants seemed to construe groups as grounded in collective recognition (Noyes & Dunham, 2017). Participants (children, ages 4–9) believed an individual was a member of a group if the individual thought she was, the group members thought she was, and this was common knowledge among them. This study could not dissociate between the three different accounts outlined above though (collective recognition, function, and dual-character). The study investigated how individuals join groups, not how groups persist over time. Moreover, it did not test the importance of function, and groups were not described in functional terms. Thus, we do not yet know whether adults believe groups are grounded in collective recognition, function, or some combination of the two.

Hypotheses

We hypothesized that participants' intuitive theories of institutional social groups are dual-character: We predicted that collective recognition is the privileged (and *technical*) sense of groups and that function is a less privileged *true* sense of groups (this terminology is taken from the dual character literature; Knobe et al., 2013; Leslie, 2015). We suspected groups would be dual-character because collective recognition and function both capture important aspects of groups and their causal structures, but they can come apart in practice. We expected that collective recognition would underlie the default sense—the sense people rely on in most contexts to decide if a group exists—because collective recognition is a more proximate explanation of group existence. For example, when individuals recognize themselves as a group (or become recognized), normative properties like obligations are created (Gilbert, 2013; Searle, 1995, 2010). Function is a deeper explanation for why groups exist, occurring earlier in a group's etiology. But because function is more distal, it is too imprecise to individuate groups. A basic function can be realized by multiple entities: For example, Google and Bing are both search providers. And once a group is formed, it can take on a life of its own: For example, Google's function expanded into hardware and robotics.

Nevertheless, because function offers a deeper explanation of why groups exist, substantial deviations in function (e.g., if Google no longer hosted a search engine) undermines something important about a group's nature. Therefore, we predicted that concepts of institutional groups integrate two senses, one grounded in collective recognition and one grounded in function, and that they would be integrated such that collective recognition underlies default *technical* judgments and function underlies additional *true* judgments. This basic framework aligns well with the dual character account, where default *technical* judgments rely on criteria like "is hired as a scientist," whereas the additional *true* judgments rely on the realization of function (Knobe et al., 2013; Leslie, 2015).

Empirical Approach

Our guiding method was the transformation paradigm (e.g., Keil, 1989; Rips, 1989). The transformation paradigm is an established method for uncovering causal beliefs. A classic example shows a raccoon being costumed or surgically transformed as a skunk. Participants are asked whether the animal is a skunk now. Affirmative answers indicate the changed property must be central to category membership because it made a difference, negative answers indicate the changed property must be peripheral because it did not. This method has been used extensively within cognitive psychology for both generic and singular concepts across diverse domains. It is suited well for social groups. Social groups seemingly persist across drastic changes: The Lakers moved from Minneapolis to Los Angeles, and its entire roster and management have changed many times. And yet, groups can seemingly dissolve almost overnight: like the famous breakup of the Bell System in 1982. Transformations are commonplace for groups and provide an ecologically valid tool for assessing ordinary concepts.

A distinction can be made between singular and general concepts (following Rips, Blok, & Newman, 2006). General concepts like *Catholics* and *Italians* represent categories with many members, and singular concepts like *the Catholic Church* and *Italy* represent groups as individuals. We take the second perspective here: We are not interested in (e.g.) what makes Donatella Versace an Italian, but what makes Italy the same nation across changes in its properties. This contrasts with Noyes and Dunham (2017) which focused on groups as general concepts and probed how participants determine what makes an individual a member of the group. Researchers have explored how people track the identity of animals, artifacts and persons (Gutheil, Bloom, Valderrama, & Freedman, 2004; Gutheil, Gelman, Klein, Michos, & Kelaita, 2008; Gutheil & Rosengren, 1996; Molouki & Bartels, 2017; Rips et al., 2006; Strohminger & Nichols, 2014); but the identity literature has not investigated institutional groups.

To summarize, we used vignettes centered on various transformations to understand causal beliefs about individual social groups. In Study 1 we examined whether institutional groups are grounded in collective recognition in a way that natural kinds and artifacts are not. In Study 2 we compared the importance of collective recognition and function for institutional groups. In Study 3 we ruled out alternative explanations for Study 2's results. In Study 4 we examined the separate role that members and relevant third-parties play in collective recognition. Studies 1–4 only examined participants' default judgments about social groups and did not examine whether their concepts are dual-character. In Study 5 we tested whether they are. In Studies 6–7 we examined the discriminant validity of Study 5 by examining whether its results extend to natural kinds and artifacts. In Study 8 we considered the possibility that participants' own values might influence their judgments about social groups, which further informs whether social groups are dual-character as well as revealing some real-world social policy implications.

Study 1

In Study 1 we began with a preliminary investigation into whether people believe social groups are grounded in something current and ongoing about human beings and their social arrangements (i.e., collective recognition) and believe that entities from other domains, like artifacts and natural kinds, are not. We hypothesized that participants would indicate that a social group had terminated once it divided into two subgroups no longer recognized as belonging together by the community; in contrast, we hypothesized that participants would indicate that an animal population or artifact set would continue to exist. This study was not designed to distinguish between collective recognition and function. It was only meant as a coarse validation that social groups depend on something ongoing and conventional in ways that other entities do not. We test between collective recognition and function in the subsequent experiments (Studies 2–4). We examine the role of function for objects and animals in Studies 6–7. We also add more precision to the nature of collective recognition in Studies 2–4.

Method

Participants. We recruited participants from Amazon Mechanical Turk. We recruited 200 "hits." One-hundred and 83 participants actually took the survey. We selected our sample size of 200 "hits" to be able to reliably detect (power = .95) within-subject comparisons that were moderately weak (d = .30) after accounting for substantial participant loss due to exclusions or failure to complete the survey. That is, we would achieve 95% power with at least 75% of the recruited sample, and would achieve 90% power if we had at least 60% of the recruited sample. We suspected effect sizes would be small because we were using a noisy Internet sample with relatively complex stimuli and judgments. This is the power analysis we use to guide the remainder of our studies. We did not collect demographic data for this study. See later Participant sections for more demographic information. MTurk samples are close to representative for race and gender, but skew younger, liberal, and nonreligious (Berinsky, Huber, & Lenz, 2012; Burnham, Le, & Piedmont, 2018; Huff & Tingley, 2015). Research that compares the demographics of MTurk workers to undergrad convenience samples finds that they are more representative than undergrad convenience samples (e.g., Berinsky et al., 2012).

Design and procedure. We compared social groups with animal populations and object sets. Participants evaluated whether the animal population, object set, or social group continued to exist after it was split into two different locations and was no longer recognized as being a single entity by the community. Thus, we could look at whether changes in ongoing social properties influence the existence of animals, artifacts, and social groups. We

randomized domain within-subject to increase power and reduce recruitment costs.

Participants either responded to a reading club, chair set, and squirrel population, or to an anthropology department, pencil set, and sparrow population. In no cases did results differ between items ($ps > .50$), so we ignore item differences from here onward. For each vignette, participants then responded to a single judgment about whether the entity still existed. The test question was a 6-point Likert scale: 1= *totally disagree* to 6 = *totally agree*. Participants rated the statement: "The original X no longer exists." Therefore, higher ratings indicate a higher belief that the group has terminated (and thus a higher belief that the group's existence depends on ongoing collective recognition).

Here and elsewhere we selected social groups grounded in a few criteria. We sought groups that were in the intermediate range of institutional complexity. They were less complex than a nation-state but more complex than an ad hoc group. We selected groups that would feel coherent and realistic, and thus involve familiar elements, but not actually be a group people know from the real world.

Results and Discussion

Participants said that social groups terminated when they split and were no longer recognized, $M = 3.82$, 95% CI [3.59, 4.06], this was significantly above the midpoint: $t(182) = 2.77$, $p = .006$, $d = .20$. In contrast, participants said animal populations, $M = 3.23$, 95% CI [2.99, 3.47], and object sets, $M = 2.91$, 95% CI [2.67, 3.14], persisted. Both of these were below chance, (animals) $t(182) = -2.26$, $p = .025$, $d = .17$, (objects) $t(182) = -4.97$, $p < .001$, $d = .37$. Thus, social groups were rated as less continuous across the presented transformations than both animal populations, $t(363.87) = 3.56$, $p < .001$, $d = .37$, and object sets, $t(363.91) = 5.49$, $p < .001$, $d = .57$. Animal populations were nonsignificantly less continuous than object sets, $t(364) = 1.91$, $p = .057$, $d = .20$. Overall then, we find support for the conclusion that social groups depend on ongoing conventions in a way that animal populations and object sets do not. In the subsequent studies we seek more detail into the nature of these conventions. Particularly, whether collective recognition, function, or some combination of these underlie the continuity of social groups.

Study 2

Study 2 had two primary goals. First, we wanted to contrast the importance of collective recognition and function. Second, we wanted to compare these in real-world scenarios, using plausible vignettes that did not explicitly mention collective recognition or function. To accomplish this goal, we presented a group of participants with a long list of possible issues a group might disagree about. We then had participants rate the issues on whether they would disrupt the group's ability to maintain a common identity, and whether the issues were relevant to the core purpose of the group. We then presented another group of participants with vignettes derived from these ratings in which the group split into two subgroups because the members disagreed about the issue. Participants judged whether the group no longer existed after the split. If a group splits on an issue that disrupts their collective identity, then participants can infer that group members no longer collectively recognize themselves as members of one group. If a group splits on an issue relevant to the core purpose of the group, then participants can infer that the members no longer realize the same basic function. Thus, we were able to measure the contribution of collective recognition and function using realistic vignettes that never mentioned these variables.

Method

Participants. We recruited participants from MTurk. We recruited 50 participants for issue ratings and 200 for the experimental conditions. One-hundred and 86 completed the study. We did not collect demographic information for the first sample. The second sample was M age = 34.51 years old; 85 women and 101 men. One-hundred and 40 participants were White (137 non-Hispanic, three Hispanic), 16 were Black (14 non-Hispanic, two Hispanic), 16 were Asian (15 non-Hispanic, one Hispanic), eight were multiracial (seven non-Hispanic, one Hispanic), one was Native American, and five preferred not to indicate their race (two non-Hispanic, three Hispanic).

Stimuli. We described a hypothetical reading group. Participants in the first sample rated 15 hypothetical issues the reading group could disagree about. They judged whether the issue would make it difficult for the group to maintain a common identity, which we used as a proxy for collective recognition, and whether the issues were relevant to the core purpose of the group, which we used as a proxy for realizing the same basic function. The collective recognition question read:

> *Imagine the members disagreed about the following issues. Rate how difficult it would be for the members to maintain a common identity if they disagreed about this* (very easy to very hard, 6 points).

The function question read:

> *Imagine the members disagreed about the following issues. Rate how important this issue is to the core purpose of the organization* (very not central to very central, 6 points).

We selected four issues. One issue disrupted both collective recognition and function, one issue disrupted only collective recognition, one issue disrupted only function, and one issue disrupted neither. We had a set of mutually constraining criteria. We tried to maximize the magnitudes: Any condition that was supposed to disrupt collective recognition should be rated among the most disruptive. Any condition that was supposed to be undisruptive should be rated among the least disruptive. But we also tried to maximize similarity across stimuli. For example, we wanted the "both" condition to be as equally disruptive of both collective recognition and function as the single conditions were of collective recognition and function, respectively. This similarity was important for interpreting the pattern of results. The resulting issues were whether abortion is morally justifiable or not, whether the group should read fiction or nonfiction, whether the group should read LGBT books or not, and whether the members should raise their hands before speaking or not. Abortion: collective recognition disruption = 4.40, 95% CI [3.96, 4.83], function disruption = 2.90, 95% CI [2.42, 3.38]. Fiction/nonfiction: collective recognition disruption = 2.64, 95% CI [2.32, 2.96], function disruption = 4.08, 95% CI [3.64, 4.52]. LGBT books: collective recognition disruption = 3.80, 95% CI [3.42, 4.18], function disruption =

3.74, 95% CI [3.36, 4.12]. Raising hands: collective recognition disruption = 2.32, 95% CI [1.96, 2.68], function disruption = 2.56, 95% CI [2.19, 2.93]. Note, that we pretested other groups. We selected the reading group because it happened to produce a desirable spread of item ratings so we could pick stimuli that conformed well to the four conditions. Each of the four vignettes had the same basic format. Here is the one for hand raising:

> *Imagine that one year the members of Cedar Women's Literary Club discussed raising their hands before speaking or not. The members disagree about raising their hands or not. Some members want to raise their hand first and others do not. They decide to split, so that hand-raising members meet Monday and nonraising members meet Friday.*

Design and procedure. We compared the contributions of collective recognition and function with the continuity of the social group. Participants responded to the four vignettes within-subject. Participants judged whether the original group no longer existed after the split. A within-subject design has a few benefits. It increased statistical power and reduced recruitment costs. It also ensured that participants interpreted the test question the same across trials. We introduced participants to the purpose of the study and the reading group (see above). Participants then read through each of the four vignettes. Participants judged the persistence of the original group following the change: *Please rate the truth of the following statement:* "*The original Cedar Women's Literary Club no longer exists.*" The test question was a 6-point Likert scale: 1 = *totally disagree* to 6 = *totally agree*. Therefore, higher ratings indicate a higher belief that the group has terminated.

Results and Discussion

We compared the importance of collective recognition with that of function. We used a multilevel mode with participant as a random effect: R package "lmer" and "lmerTest." The outcome variable was participants' judgments about whether the group stopped existing Likert Scale: 1–6. The predictors were collective recognition disruption (high vs. low), function disruption (high vs. low), and their interaction. There was a significant main effect of collective recognition disruption, $b = .51$, $SE = .09$, $p < .001$. There was no main effect of function disruption, $b = .09$, $SE = .086$, $p = .293$, nor significant interaction, $b = -.04$, $SE = .12$, $p = .76$. Participants judged that high disruption to collective recognition ($M = 3.84$, $SD = 1.56$) terminated the original group more than low disruption ($M = 3.38$, $SD = 1.52$), $t(369.74) = 3.04$, $p = .002$, $d = .51$ (see Figure 1).

These findings suggest that collective recognition is more important to the continuity of groups than function. Recall that a sample of participants rated some issues as making it very difficult for the group to maintain its common identity. Thus, a split based on one of those issues would undermine the members' collective recognition of themselves as a group. Other issues were rated as very relevant to the group's core purpose, and thus splits based on those issues should disrupt the group's function. When given descriptions of splits based on these various issues, participants rated the group as terminating when collective recognition was disrupted, not when function was disrupted. These results support the collective recognition account.

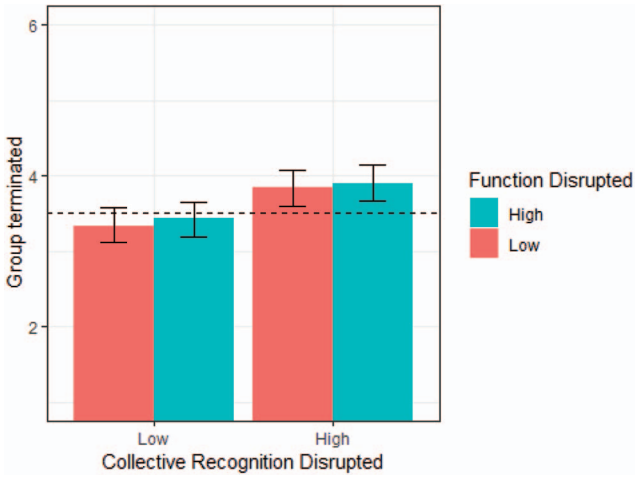

Figure 1. Ratings of whether the group terminated after a split; 1 = *totally disagree* to 6 = *totally agree* that the original group no longer exists. Error bars are 95% bootstrapped confidence intervals. See the online article for the color version of this figure.

In a supplemental study, we replicate Study 2 with an anthropology department. An anthropology department is more complex and its function is more important than a library club. Yet, we find the collective recognition is more important than function for this stimulus too, thereby ruling out arguments that our first set of results were due to the more interpersonal nature of small reading groups. These results are not specific to the reading group we selected and extend to a group that differs in significant ways. In Study 3 we rule out alternative explanations to ensure the results reflect collective recognition and not some other variable.

Study 3

Study 2 supported the conclusion that collective recognition was more causally central to groups than function. A strength of Study 2 was that collective recognition and function were not directly mentioned, and instead, these variables were assessed through realistic disagreements about various topics. In Study 3 we confirm that the results of Study 2 were about collective recognition rather than some other inference participants were making about the vignettes. We present the two issues (LGBT novels and abortion) from Study 2, but we modified how these issues manifested in the group. In one condition, the members realize they disagree (about LGBT novels or abortion) and they split as a function of their disagreement—which participants can use to infer members no longer collectively recognize themselves as a group. The other conditions decompose this scenario to examine whether something less-than collective recognition explains the results of Study 2.

In one condition (*no knowledge-split*) the members split indirectly because of the disagreement. Their different opinions lead to scheduling conflicts, and so (e.g.) all of the proabortion members meet on one day and all of the prolife members meet on a different day. Therefore, the members no longer interact as a whole, and they split along a contentious fault line, but they probably still collectively recognize themselves as being a group because the issue was not raised. In another condition (*knowledge-no split*) the

members disagree about the issues but decide to remain together; thus, the group is revealed to be heterogeneous but the members presumably continue to collectively recognize themselves as a group. Finally, the members privately disagree about the issue but it never comes up and they never split (*no knowledge-no split*). If collective recognition explains the results of Study 2, then participants should only report that the group terminated in the *knowledge-split* condition, and should report that the group continued existing in all other cases.

Method

Participants. We recruited 200 "hits" on Amazon Mechanical Turk. 186 participants took the survey. Six participants were then excluded because they failed embedded attention checks. The resulting sample was: *M* age = 36.5, 74 women and 105 men. One-hundred and 57 participants were White (145 non-Hispanic, 12 Hispanic), four were Black (three non-Hispanic, one Hispanic), nine were Asian, four were multiracial, one was Native American, and four preferred not to indicate their race.

Design and procedure. Participants judged whether a group's identity ended after some change. The group either split or did not split (split vs. no split). And the group either had knowledge of their disagreements or not (knowledge vs. no knowledge). We randomly assigned condition within-subject.

We used the same group from Study 2, and we used both scenarios: abortion and LGBT issues. There were four conditions with two versions each depending on the issue. In the *split-knowledge* condition, members decided to split because they disagreed. In the *split-no knowledge* condition, members split indirectly because of the disagreement:

> *. . . One day the members split: the prolife members meet Monday and the prochoice members meet Friday. This happened because they had trouble scheduling an overlapping meeting time. The members never discuss their views and they never realize they disagree.*

In the no split-knowledge condition, members disagree but decide not to split. In the no split-no knowledge condition, members neither split nor discussed the issue. Participants responded to the same identity judgment question from Study 2. Therefore, higher ratings indicate a higher belief that the group has terminated.

Results and Discussion

We used a multilevel model. The outcome variable was participants' judgment about whether the group stopped existing: (1–6). The predictors were knowledge (*knowledge* vs. *no knowledge*), splitting (*split* vs. *no split*), and their interaction. There was a main effect of splitting (split), $b = -1.12$, $SE = .14$, $p < .001$, such that groups were rated as having terminated more if they split than if they did not split. There was no main effect of knowledge, $b = .13$, $SE = .14$, $p = .310$. As predicted, there was a significant interaction between knowledge and split, $b = .73$, $SE = .19$, $p < .001$. Participants rated split-knowledge higher than all other conditions (see Figure 2). This was the only condition rated higher than the midpoint, $t(179) = 5.11$, $p < .001$. In all other conditions, ratings were below the midpoint, $p = .022$, $p < .001$, $p < .001$.

In Study 2 we found that when members split because of issues like abortion and reading LGBT novels participants judged that the group had terminated. These were issues that other participant had said were likely to make it hard for the group to maintain a common identity. We interpreted this as reflecting the importance of collective recognition. The results of Study 3 help support this conclusion. If members split on the issue, but maintain collective recognition, then the group persists. Thus, members are a group, and remain a group, because the members collectively recognize themselves as members (Noyes & Dunham, 2017). In combination with Study 2, this suggests that collective recognition—and not

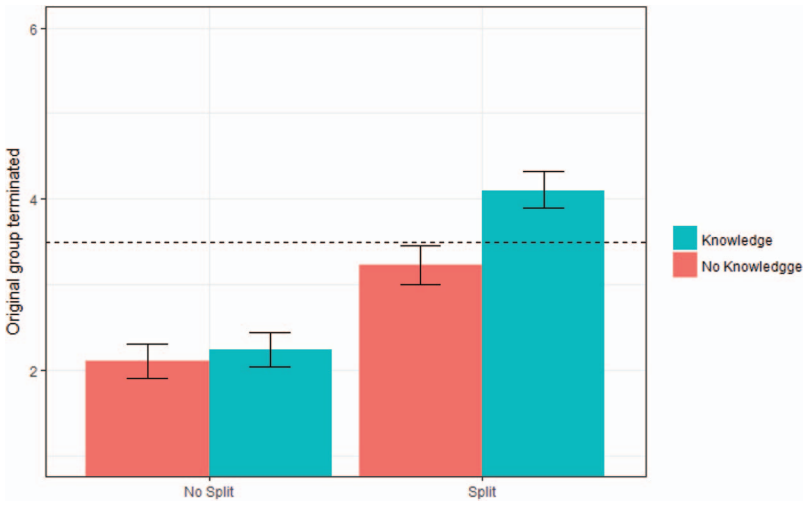

Figure 2. Participant's judgments about whether the group stopped existing; 1 = *totally disagree* to 6 = *totally agree* that the original group no longer exists. Error bars are 95% bootstrapped confidence intervals. See the online article for the color version of this figure.

function—is most central to participants' intuitive beliefs about the causal structure of social groups.

Study 4

One limitation of Studies 2–3 is that we only examined the collective recognition of the members themselves. As we outlined in the introduction though, collective recognition is not always about the members themselves but includes the recognition of the group by relevant third-parties. For example, a relevant outgroup for a reading group on the college campus might be the college administrators. They have the authority to create and regulate student groups. In Study 4 we examined the separate contributions of group members and relevant third parties. Collective recognition theories from philosophy suggest that either of these are sufficient to constitute a group (e.g., Searle, 1995, 2010). Thus, we test whether participants believe that a group persists if either the members of a relevant third-parties recognize the members as a group.

We described two groups, a college club (women's literary club) and a nonprofit organization (nonprofit veterinarian organization). We presented four vignettes. In all of the vignettes, the members stopped meeting. The group's basic function continued to be realized but independently by each member. For example, all of the individuals continued to read the same set of books or provide free veterinarian services, but they did so alone. What varied was whether the members were recognized as a group by themselves (*ingroup*) or a relevant third party (*outgroup*). Here, the college administration and the local government were the relevant third parties, respectively. We then asked participants whether the group still existed or not. This helped us refine the nature of collective recognition. It also provided an additional test of the importance of function. Function was kept constant throughout each condition. If the function account is correct, the group should persist in all cases, even when not recognized.

Method

Participants. We recruited 200 "hits" on Amazon Mechanical Turk. One-hundred and 88 participants took the survey. Ten participants were then excluded because they failed embedded attention checks. We did not collect demographic information for this survey.

Design and procedure. Participants were randomly assigned to answer questions about a reading group or a nonprofit organization. Participants read four vignettes presented in randomized order. In all four vignettes, participants learned that members of the group stopped meeting but privately undertook the same activities they had previously performed. For example, the members of the nonprofit organization had previously provided free medical care to animals. After they stopped meeting, the individuals continued to provide these services. What we varied was whether the members of the group continued to recognize themselves as a group (*ingroup* collective recognition), or whether a relevant third party continued to recognize the members as a group (*outgroup* collective recognition). Participants responded to a question about whether the group still existed: for example, "*The Women's Literary Club still exists*," on a 6-point Likert scale ranging from 1= totally disagree to 6 = totally agree. In the *ingroup* condition, participants read (for the nonprofit condition):

> *One day the local government deregisters the nonprofit. The government no longer recognizes the nonprofit or its volunteers. The administration removes funding and resources. The volunteers mutually decide to continue providing free up-to-date health care in their neighborhoods. They stop meeting but they still recognize themselves as members and continue providing their services alone.*

In the *outgroup* condition, participants read (for the nonprofit condition):

> *The volunteers stop meeting. They do not recognize themselves as members anymore. Each individual privately continues to provide free up-to-date medical care to wounded animals in their neighborhoods. The local government keeps the nonprofit registered and continues to allocate funding and resources. The government continues to recognize the volunteer's membership.*

In the *both* vignette, participants received the affirmative parts of these two vignettes, and in the *neither* vignette participants received the negative parts of these two vignettes.

Results and Discussion

To test the effect of collective recognition, we used a multilevel model. The outcome variable was the judgment about whether the group continued to exist (1–6). The predictors were collective recognition by the members themselves (*ingroup*, present vs. absent), collective recognition by a relevant third party (*outgroup,* present vs. absent), and their interaction. Participant was a random effect. There was a main effect of ingroup recognition, $b = 1.56$, $SE = .14$, $p < .001$, and outgroup recognition, $b = 1.67$, $SE = .14$, $p < .001$. There was also a significant interaction between them, $b = -.60$, $SE = .19$, $p = .002$ (see Figure 3). This emerged because each form of collective recognition was sufficient by itself to sustain the existence of the group: participants judged that the group persisted when only ingroup recognition was present, $M = 3.83$, 95% CI [3.61, 4.05], $d = .22$, and when only outgroup recognition was present, $M = 3.96$, 95% CI [3.73, 4.18], $d = .30$.

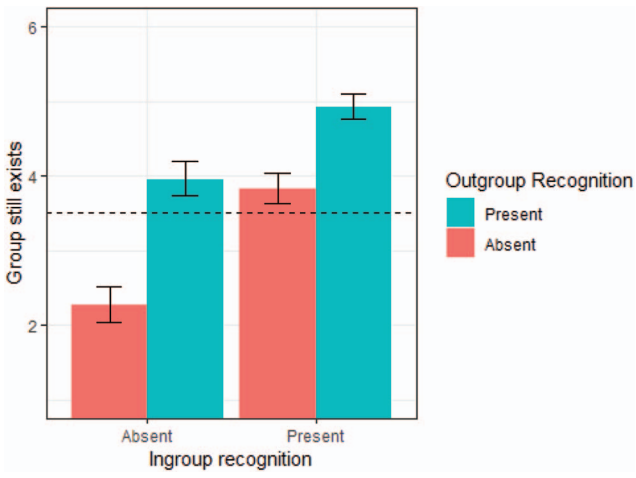

Figure 3. Participant's judgments about whether the group continued to exist; 1 = *totally disagree* to 6 = *totally agree* that the group still exists. Error bars are 95% bootstrapped confidence intervals. See the online article for the color version of this figure.

And thus, though the effect of each on their one was very large (*in* vs. *neither*: $p < .001$, $d = .87$, and *out* vs. *neither*: $p < .001$, $d = .93$), they had a more moderate additive effect (*both* vs. *in*: $p < .001$, $d = .63$, and *both* vs. *out*: $p < .001$, $d = .59$). There was no effect of group: In a model where group was added (nonprofit vs. reading club), the three-way interaction was nonsignificant, $b = .53$, $SE = .39$, $p = .166$, and neither were any of the lower-order group effects (*p*s $> .19$).

This study provided greater insight into collective recognition. We find that a group persists so long as its members or a relevant third-party recognize them as being part of a group—either is sufficient. This makes sense based on the relevant philosophical work because either of these forms of collective recognition can create and sustain social facts (Searle, 1995, 2010). For example, when the individuals are recognized as a group by their university, they have special privileges (i.e., funding, access to campus space) and obligations (i.e., to accept new members). Even if these privileges go unused and these obligations are broken, they still exist and still have social consequences. Likewise, when the group members stop meeting, their promise to each other to continue reading sustains a social fact (their obligation to each other to keep reading women's literature). Ultimately, a group is a set of social facts, normative properties like privileges, duties, and obligations that people have in virtue of being members (Searle, 1995, 2010).

These results provide additional support for the account that collective recognition underlies group continuity. In all four conditions the members continued to realize the basic function of the group—yet, if those members were not recognized by themselves or a third party as being a group, then participants said the group had stopped existing. In contrast, collective recognition had a large influence on judgments about group continuity. Thus, together with Study 2 and 3, there is robust support for the collective recognition account (Noyes & Dunham, 2017).

In the introduction, we suggested that groups might be dual-character. In the same way, participants entertain both a "technical" and "true" sense about scientists, such that a person can be a technical scientist or a true scientist or both, perhaps concepts of groups are also dual-character. For example, an intelligible statement about an institutional group is: "I cannot support Yale University anymore. The true Yale is over." Participants may reason that a group that still technically exists (because it is still collectively recognized over time) does not *truly* exist because it no longer realizes the same basic function. We assess this possibility in Studies 5–8. Function may be important to judgments about group continuity when the second "true" (as opposed to "technical") sense is assessed. Studies 2–4 would reflect the unqualified or technical sense of groups because it did not specify "true" judgments.

Study 5

In Study 5 we designed judgments to discriminate between the "technical" and "true" senses of groups in order to test whether two distinct senses existed and thus whether groups were dual-character or not (Knobe et al., 2013). We developed vignettes that we thought would exemplify a group that fits the "technical" but not "true" sense and a group that fits the "true" but not "technical" sense. To do, we used vignettes grounded in mission drift (Ebrahim, Battilana, & Mair, 2014; Jones, 2007). Mission drift is when an organization drifts from its former purpose, and thus no longer realizes its basic function. For example, many nonprofits end up devoting more energy to capital accumulation than aid work over time. We presented vignettes where groups end up drifting from their mission, and thus no longer realize their original function. Then, a new group emerges that does realize the original function. We had participants evaluate whether either group, both, or neither was the "technical" and "true" sense of the group.

Method

Participants. We recruited participants from Amazon Mechanical Turk. Two-hundred people completed the survey. Sixteen of those were excluded for failing embedded comprehension checks. We did not collect demographic information.

Design and procedure. Participants read two vignettes, each depicting a different group. Each vignette described a case where a group drifted from its mission and thus no longer realized its original function. An unrelated group begins which does realize the original function. For one of the vignettes, participants responded to a "technical" judgment, and for one they responded to a "true" judgment. We randomized which group was paired with which judgment.

There were two groups described: a reading club and a nonprofit animal clinic. For each group, participants read an initial brief description. They then read a prompt that described how the group drifted from its original mission, and thus no longer realized its function. A new group begins which realizes the original function instead. This first group is thus descended from the original (*descended*) but no longer functional, whereas the other group is unrelated but realizes the original function (*functional*). Participants then responded to a "technical" or "true" judgment. For technical identity judgments, participants read this (for the reading group):

> *Imagine the Student Budget Committee at Cedar College. All the chartered organizations at Cedar College have a Student Budget Identification Number (SBIN#). Cedar Woman's Literary Club is #74.*
>
> *Each year the Student Budget Committee reviews the existing groups so that it can delete numbers when groups end, renew numbers for continuing groups and provide new numbers for new groups. Each group needs a unique identification number.*
>
> *The Student Budget Committee should technically . . .*

And then they rated the truth value of four separate statements on a 6-point scale: 1= *totally disagree* to 6 = *totally agree*. Here we use the exact wording for the reading group: (a) (*descended*) Give the original #74 to only the current Cedar Woman's Literary Club. (b) (*functional*) Give the original #74 to only the Society for Reading Women. (c) Delete the original #74 and assign new numbers. (d) Give the original #74 to both meetings.

For "true" judgments, participants read this (for the reading group):

> *Imagine a student at Cedar College. This student is different from you in almost every way. Their family was always a member of the Cedar Woman's Literary Club. They want to join the group too. They are thinking about which day they should attend.*

It is really important to them that they join the true Cedar Woman's Literary Club. They want to join the same group their family was a part of.

Ultimately, if you really think about what the original Cedar Woman's Literary Club was, the student should . . .

And then they rated the truth value of four separate statements on a 6 point scale: 1= *totally disagree* to 6 = *totally agree*: (a) (*descended*) Join the current Cedar Woman's Literary Club. (b) (*functional*) Join the current Society for Reading Women. (c) Join neither of the groups. (d) Join either/both of the groups.

Results

First, we looked at whether the ratings differed between technical and true judgments. To do so, we used a 2 (Type: Technical, True) × 4 (Judgment: Descended, Functional, Both, Neither) within-subject ANOVA. This test revealed a significant two-way interaction between type and judgment, $F(3, 549) = 21.7$, $p < .001$, *ges* (generalized eta squared) = .13 (see Figure 4). Participants responded to the two prompts differently and made different patterns of judgments.

Next we examined judgments of the two groups (*descended* vs. *functional*) by focusing on the judgments of single groups (excluding both and neither). There was a two-way interaction between group and judgment type, $b = 1.68$, $SE = .26$, $p < .001$. Participants rated the descended group higher for technical judgment, $b = -.80$, $SE = .18$, $p < .001$, $d = .27$. Participants rated the function group higher for the true judgment, $b = .88$, $SE = .18$, $p < .001$, $d = .29$. This is consistent with the conclusion that "true" judgments differ from "technical" ones. It is also consistent with function being important to group continuity in the "true" sense. The functional group was rated as far more the "true" group than the "technical" group, $t(183) = -9.0$, $p < .001$, $d = .67$. And the functional group was the only group rated above the midpoint, $t(183) = -6.0$, $p < .001$, $d = .44$. This pattern of results generalized across groups (nonprofit vs. reading group): There was a weak three-way interaction when group is included in the model, $b = 1.00$, $SE = .51$, $p = .047$, but the two-way interaction described above was significant for both the reading group ($b = 1.13$, $SE = .35$, $p = .001$) and nonprofit ($b = 2.14$, $SE = .37$, $p < .001$).

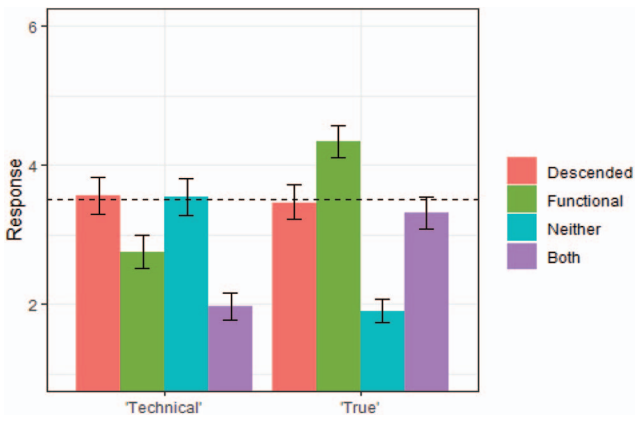

Figure 4. Ratings of the two groups by judgment type; 1 = *totally disagree* to 6 = *totally agree* that the group is the technical or true successor of the original group. Error bars are 95% bootstrapped confidence intervals. See the online article for the color version of this figure.

Next, we examined both and neither judgments. There was again a two-way interaction between both/neither and judgment type, $b = 3.00$, $SE = .22$, $p < .001$. Participants rated both higher for true judgments, $b = 1.41$, $SE = .15$, $p < .001$, $d = .69$, and neither higher for technical judgments, $b = -1.59$, $SE = .16$, $p < .001$, $d = .71$. We take this as a signature of how the different judgments were constructed: The technical judgment demands that a unique identifier is assigned to one and only one group, so by definition neither *should* be higher than both. Indeed, the technical prompt allows for the unique identifier to be deleted, and so the prompt encourages neither judgments. Whereas for the true judgment, it is possible for a participant to join more than a single group, and the prompt suggests the individual would like to join a group, so the prompt allows for and encourages both judgments rather than neither.

Overall, we find support for dual-character groups. Participants relied on collective recognition to make judgments about groups in nonqualified settings (Studies 1–4), and participants rated the descended group (the group presumably most likely to be collectively recognized as the original) as the more "technical" successor. In contrast, participants heavily relied on function to make judgments about the "true" group. This suggests that groups have two senses, the sense in which they are collectively recognized and the sense in which they realize the same function. Participants have a default way of thinking about social groups, as seen in Studies 2–4, where collective recognition underlies groups. The "technical" sense is generally privileged for making default judgments when reasoning about dual-character concepts (Knobe et al., 2013; Leslie, 2015).

Collective recognition makes sense as the "technical" or default sense because it is more directly relevant for individuating a group and tracking it across changes (Google and Bing both fulfill the same basic function but are not the same group; Google has taken on a life of its own since its conception, expanding its function, even prior to its merger). Having an additional sense of groups is important though because it captures the deeper sense of why a group exists at all—to fulfill a certain function. Having both senses is adaptive and helps one track different forms of information: For example, knowing that a cancer charity's members still recognize themselves as a group, and continue to be recognized by the IRS as a nonprofit, informs you that the members still have certain normative properties: They are obligated to meet certain legal regulations, they have a duty to follow their bylaws and charter, and they have the privilege of being tax exempt, they have a commitment to fulfill their mission statement. Knowing that they only care about money and hardly fund cancer research anymore tells you to cancel your monthly donations.

To strengthen the conclusion that social groups are dual-character we need to confirm that the "true" judgment is measuring what it purports to measure. For example, participants make different sets of judgments in contexts of other linguistic hedges, like "loosely speaking" and "according to experts" even for concepts that are not dual character (Malt, 1990). For example, participants say a whale is "loosely speaking" a fish, but a mammal "according to experts." In order to show that the "true" judgments are not

operating like a "loosely speaking" hedge, we need to test categories that should not be dual character and show that this hedge does not produce different response patterns (Knobe et al., 2013). In Studies 6 and 7 we repeat the basic method of Study 5 with animals and artifacts, which are not dual character (Knobe et al., 2013).

Study 6

We adapted the methods of Study 5 to examine animal populations. An animal population descends over time and loses its original adaptations to an ecological niche, and thus loses its original functional properties. A new animal population adapts to its original ecological niche, and thus takes the function of the original animal population. We reasoned that animal populations should not be dual-character: Animals only have a single sense, their possession of the right essential properties (Gelman, 2003; Knobe et al., 2013). So, we expected participants to rate the descended animal population (which possesses the same essence) higher than the functional animal population for both "true" and "technical" judgments. However, given that participants are willing to say that a whale is "loosely speaking" a fish, certain hedges can elicit functional responses even about animals (Malt, 1990). We thus used Study 6 to determine whether "true" judgments are tapping into something distinct and dual-character-like about social groups, or merely tapping into similarity-like judgments in the way "loosely speaking" does.

Method

Participants. We recruited participants from MTurk. One-hundred and 86 people completed the survey. Thirty-four were excluded for failing embedded attention checks.

Design and procedure. The design of Study 6 is a close analog of Study 5. An original animal population (tigers or frogs) is described. Over time, the population loses its original function (i.e., adaptive behaviors and traits) when it leaves its ecological niche. An unrelated population adapts to that niche, and so takes on the original functional properties of the population. Participants responded to a "technical" judgment and a "true" judgment. The pairing was randomized. The technical and true judgment were worded similarly to Study 5. For the technical judgment, participants had to decide which group to assign a unique identifying number. For true judgment, participants had to decide which population a student should study: For example, "*It is really important to them that they study the true Lampang Tigers. They want to study the same population their family studied."*

Results and Discussion

First, we looked at whether the pattern of judgments differed across type (technical vs. "true"). To do so, we used a 2 (Type: Technical vs. "True") × 4 (Judgment: Descended, Functional, Both, Neither) within-subject ANOVA. This test revealed a significant two-way interaction between type and judgment, $F(3, 453) = 47.05$, $p < .001$, $ges = .09$ (see Figure 5). Participants responded to the two prompts differently and made different patterns of judgments.

Next, we examined judgments of the two populations (descended vs. functional) by focusing on the judgments of single

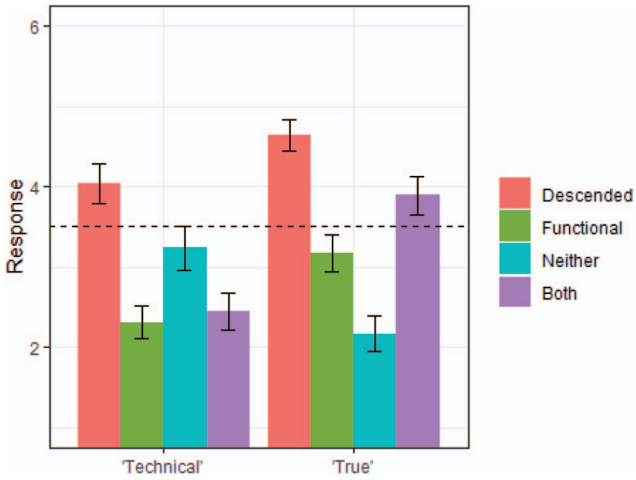

Figure 5. Ratings of the two populations by judgment type; 1 = *totally disagree* to 6 = *totally agree* that the population is the technical or true successor of the original population. Error bars are 95% bootstrapped confidence intervals. See the online article for the color version of this figure.

populations (excluding both and neither). The two-way interaction between population and judgment type was nonsignificant, $b = 0.25$, $SE = .23$, $p = .280$. In both cases, the descended population was rated higher than the functional population, $b = -1.73$, $SE = .16$, $p < .001$, $d = .74$. This effect generalized across animal populations (frog vs. lion): The three-way interaction in a model that included group was nonsignificant, $b = .60$, $SE = .55$, $p = .276$. Next, we used a cross-study comparison. We restricted this analysis to the descended and functional population ratings. We used a multi-level model, which included study (4 vs. 5), judgment type ("technical" vs. "true"), and judgment (descended vs. functional) as predictors. There was a three-way interaction, $b = -1.43$, $SE = .36$, $p < .001$. These results are consistent with animal populations being grounded in essences (Gelman, 2003) and not being dual-character (Knobe et al., 2013). This pattern also supports the conclusion that social groups are dual-character, and that the results of Study 5 were not tapping into something similarity-based like a "loosely speaking" judgment, but instead reflects something more conceptually deep about social groups.

Next, we examined both and neither judgments. There was a two-way interaction between both/neither and judgment type, $b = 2.52$, $SE = .23$, $p < .001$ (see Figure 5) that was similar to Study 5—both judgments were lower than neither for "technical" and higher than neither for "true." As with Study 5, we take this as reflecting the constraints of the different judgment types. Technical judgments demanded low both ratings, whereas "true" judgments invited high both ratings. This contrast suggests that participants were responsive to the different measures and implemented them appropriately for the animal population, but nevertheless reached different conclusions than Study 5.

Study 7

In Study 7 we examined an artifact set with the same logic as Studies 5–6. Artifacts should have a single sense grounded in the creator's intended kind (Bloom, 1996), and should not be dual-

character (Knobe et al., 2013). We examined a set of artifacts that are copied over time (*descended*). People stop using the artifacts for their original function though. A new set of artifacts is made that fulfills their original function (*function*). Therefore, the descended set should belong to the same kind as the original set, because they were all created with the same intended kind in mind.

Method

Participants. One-hundred and 83 people completed the survey. Sixty-three participants were excluded for failing attention checks. We did not collect demographic data.

Design and procedure. The design of Study 7 is a close analog of Studies 5 and 6. Participants read a vignette about an artifact set (hammer or shovel) that descended over time, such that its design was copied over time. It starts to be conventionally used for a different function though. An unrelated set of artifacts is used for the original function instead. Participants responded to a "technical" judgment and a "true" judgment. The "technical" judgment was again about assigning a unique identifier. The "true" judgment was about buying an object from the same set: "*It is really important to the student that they buy a shovel from the true set of Zaragozan shovels. They want to own a hammer from the same set their family used to own shovels from.*" We randomized the pairing between judgment and set.

Results and Discussion

First, we looked at whether the pattern of judgments differed across the type of identity: technical versus true. We used a 2 (Type: Technical vs. True) × 4 (Judgment: Descended, Function, Both, Neither) within-subject ANOVA. This test revealed a significant two-way interaction between type and judgment, $F(3, 546) = 15.86$, $p < .001$, $ges = .04$ (see Figure 6). Participants responded to the two prompts differently and made different patterns of judgments.

Next, we examined judgments of the two sets (descended vs. functional) by focusing on the judgments of single sets (excluding both and neither). The two-way interaction between set and judgment type was significant in the opposite direction of Study 5, $b = -.78$, $SE = .28$, $p = .006$. In both cases, the descended set was rated higher than the functional set, $b = -1.69$, $SE = .14$, $p < .001$, $d = .72$ – but this was even more pronounced for "true" judgments (see Figure 6). These results generalized across the different artifacts; the three-way interaction in the model with artifact set (shovel vs. hammer) included was nonsignificant, $b = -.31$, $SE = .56$, $p = .57$. Next, we conducted a cross-study comparison between Study 5 and the present study. There was a three-way interaction, $b = -2.45$, $SE = .39$, $p < .001$, between study (5 vs. 7), judgment type (technical vs. true), and population (descended vs. functional). Participants respond differently across the two domains. These results are consistent with artifacts being grounded in the creator's intended kinds (Bloom, 1996) and not being dual-character (Knobe et al., 2013). This adds further support to the discriminant validity of Study 5. Social groups are dual-character. The methods of Study 5 were tapping into something meaningful about concepts of social groups and were not merely tapping into something trivial like similarity ratings.

Next, we examined both and neither judgments. There was a two-way interaction between both/neither and judgment type, $b = 1.38$, $SE = .26$, $p < .001$ (see Figure 6) that was similar to Studies

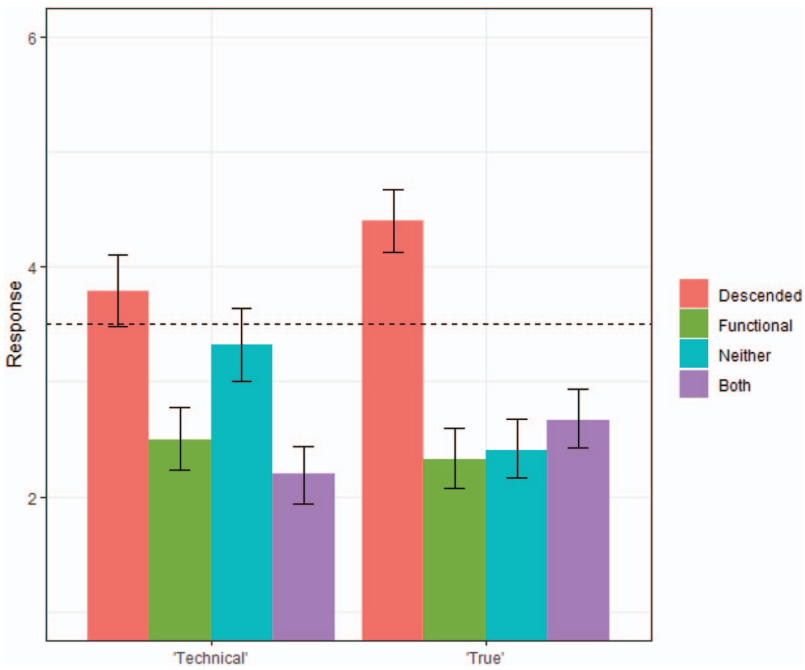

Figure 6. Ratings of the two sets by judgment type; 1 = *totally disagree* to 6 = *totally agree* that the population is the technical or true successor of the original population. Error bars are 95% bootstrapped confidence intervals. See the online article for the color version of this figure.

5 and 6—both judgments were lower than neither for "technical" and higher than neither for "true." As with Study 5 and 6, we take this as reflecting the constraints of the different judgment types. Again, this suggests that participants were responsive to the different measures even if they reached similar conclusions for both judgment types.

Study 8

Studies 5–7 supported the conclusion that social groups are dual-character. There are two senses, the sense in which the group technically still exists (whether it is still collectively recognized, Studies 2–4), and the sense in which the group truly still exist—whether the group realizes the same basic function over time. This dissociation might be part of the cognitive mechanisms underlying people's decisions to leave a group. Researchers have examined real-world schisms in political parties, religious groups, businesses, and nation-states (Bookman, 1994; Dyck, 1997; Sani & Reicher, 1998, 2000). One account suggests that groups split when there is disagreement about the "essential" nature of the group (Sani, 2005; Sani & Reicher, 1998, 2000). A natural reading is that the "essential" nature corresponds to the "true" sense. If participants own ideological commitments influence their beliefs about the "true" nature of the group, then conflicts over policies can lead to ideologically based fracturing. Thus, in Study 8 we investigated whether participants' own ideological commitments would influence their "true" judgments more than their "technical" judgments of social groups in the context of group schisms.

To test the effect of ideology we investigated schisms around politically charged issues. In Study 2 we had investigated two contentious issues that participants believed would lead to a group schism: the morality of abortion and whether to read LGBT books. These issues are both politically charged, such that liberals should tend to be proabortion and pro-LGBT and conservatives should tend to be antiabortion and anti-LGBT—as predicted by a wealth of social scientific data, such as the American National Election Studies.[1] Here, we examined whether political attitudes would lead participants to make different "true" judgments about groups, such that liberals would be more likely to say a pro-LGBT or proabortion group was the "true" successor of the preschism group, and conservatives would be more likely to say an anti-LGBT and antiabortion group was the "true" successor of the preschism group. We hypothesized that "true" judgments would be affected more by ideology because they are grounded in beliefs about the basic function of the group, which should tend to be a value-laden assessment (see Knobe et al., 2013), whereas the "technical" judgments involved collective recognition, which should tend to be value-neutral.

Method

Participants. We apportioned 200 MTurk "hits." One-hundred and 99 participants completed the survey. Forty-four were excluded for failing two attention and memory checks. This resulted in a sample of 155 participants recruited from Amazon Mechanical Turk. Demographic information was not collected.

Design and procedure. Participants read two vignettes from Study 2 involving a reading club. One of these vignettes described a group disagreeing about whether they should read LGBT novels or not and then fracturing into two groups, one that was pro-LGBT novels and one that was anti-LGBT novels. Another vignette was structurally similar but involved a disagreement about the morality of abortion that led to a schism. Thus, in each case, the group schisms along a liberal-conservative divide, and so one group was always on the liberal side of an issue and one group was always on the conservative side of an issue. We then had participants make either a "technical" judgment or a "true" judgment about whether the liberal-leaning group, the conservative-leaning group, both groups, or neither group was the successor of the preschism group. Participants read both the LGBT schism vignette and the abortion schism vignette. For each vignette, they either made a "technical" judgment or a "true" judgment, the pairing between vignette and judgment was randomized.

Like Study 5, the "technical" judgment involved the decision of allocating a unique numerical identifier. For example, for the abortion schism vignette, participants rated on a 6-point Likert scale ranging from 1 = *totally disagree* to 6 = *totally agree*, whether a student budget committee should "*give the original #74 to only the prochoice meeting*," "*give the original #74 to only the prolife meeting*," "*delete the original #74 and assign new numbers*," or "*give the original #74 to both meetings*." The "true" judgment involved the decision of which group a student should join who wanted to join the same group as their family. For example, for the abortion schism vignette, participants rated on the same 1–6 scale, whether a student should "*joint the prochoice meeting*," "*join the prolife meeting*," "*join neither of the meetings,*" or "*join either/both of the meetings.*"

After making these judgments, participants responded to a political attitude survey: We asked them their political orientation: *very conservative* to *very liberal*, 5 points. We asked them their agreement to five statements about abortion attitudes and five statements about LGBT attitudes. For example, "I oppose gay marriage" and "Abortion is immoral." These statements were rated on a 6-point Likert scale: 1= *totally disagree* to 6 = *totally agree*. All 11 items were highly correlated, $\alpha = .91$, 95% CI [.90, .92]. So, we collapsed these measures into a single conservative to liberal attitude measure.

Results and Discussion

First, we looked at whether the pattern of judgments was different between technical and true judgments. We used a 2 (Type: Technical vs. True) × 4 (Judgment: Liberal-Leaning Group, Conservative-Leaning Group, Both, Neither) within-subject ANOVA. This test revealed a significant two-way interaction between type and judgment, $F(3, 462) = 44.17$, $p < .001$, $ges = .08$. Participants responded to the two prompts differently and made different patterns of judgments.

[1] In the 2016 Time Series Study, which includes data from over 3,000 individuals, people higher in conservativism (1–7, *extremely liberal* to *extremely conservative*) expressed less warm feelings towards gay people (1–100 feeling thermometer, standardized beta = −.47, $SE = .02$, $p < .001$) and expressed more rejection of the legality of abortion (1–4, 1 = *by law, abortion should never be permitted*, standardized beta = −.48, $SE = .02$, $p < .001$). These analyses exclude nonscale responses like "don't know" or refusal to answer. Analyses were conducted as linear models, standardizing each variable, and treating the variables as continuous.

We examined how participants' attitudes influenced their ratings of the groups, and whether attitudes had different influences on true and technical identity. We used a multilevel model. We restricted this analysis to the rating of the singular groups (liberal-leaning vs. conservative-leaning) and excluded the both and neither ratings. Included in the model was also judgment type (technical vs. true), participants' political attitudes, and the interaction terms. This analysis revealed a significant three-way interaction, $b = 3.19$, $SE = .93$, $p < 001$.

Participants' political attitudes influenced both types of judgments: For technical judgments, there was a two-way interaction, $b = 2.85$, $SE = .65$, $p < .001$ (see Figure 7). Conservatism predicted lower ratings for the liberal-leaning group, $b = -1.36$, $SE = .49$, $p = .006$, and nonsignificantly higher ratings for the conservative-leaning group, $b = .50$, $SE = .50$, $p = .312$. For true judgments, there was also a two-way interaction, $b = 4.07$, $SE = .65$, $p < .001$. Conservatism predicted lower ratings of liberal-leaning groups, $b = -2.58$, $SE = .58$, $p < .001$, and higher ratings of conservative groups, $b = 2.48$, $SE = .56$, $p < .001$. Participants' political attitudes influenced true judgments more than technical judgments. True judgments were more than three times as affected by participants' attitudes: On technical judgments, liberals (+0.5 *SD* on attitude measure) rated the liberal Group 0.34 scale points higher than the conservative group. Conservatives (−0.5 *SD* on attitude measure) rated the conservative group 0.40 scale points higher than the liberal group. On true judgments, liberals rated the liberal group 1.35 scale points higher. Conservatives rated the conservative group 0.91 scale points higher. Participants believed that the group conforming to their attitudes was the true successor, $M = 3.86$, 95% CI [3.56, 4.16], $t(129) = 2.39$, $p = .018$. They did not believe the group conforming to their attitudes was the technical successor, $M = 2.23$, 95% CI [1.96, 2.50], $t(129) = -9.44$, $p < .001$.

These results further support the dual-character nature of social groups. Participants agreed that neither of the break-away groups was the "technical" successor of the preschism group, which is consistent with the results of Study 2. In that study, participants believed the members no longer collectively recognized themselves as members of the group, so participants said the group stopped existing. When thinking about the "true" successor of the preschism group though, participants rated the group consistent with their own ideology as being the "true" group. Thus, once again, participants reached very different conclusions about the "technical" and "true" sense of social groups.

Critically, the group was never defined around liberal or conservative policy issues. Nor were participants members of the groups. Thus, these effects are plausibly stronger in real-world settings in which people have stakes in the group and their ideological commitments are directly relevant. These results support a plausible mechanism contributing to real-world schisms. People with different ideological commitments will disagree about which policies would sustain or continue the "true" nature of their group. This disagreement will motivate them to break off when conflict arises around how the group should continue, or when there are dramatic changes in policies. This mechanism accords with social psychological models of real-world schisms (Sani, 2005; Sani & Reicher, 1998, 2000).

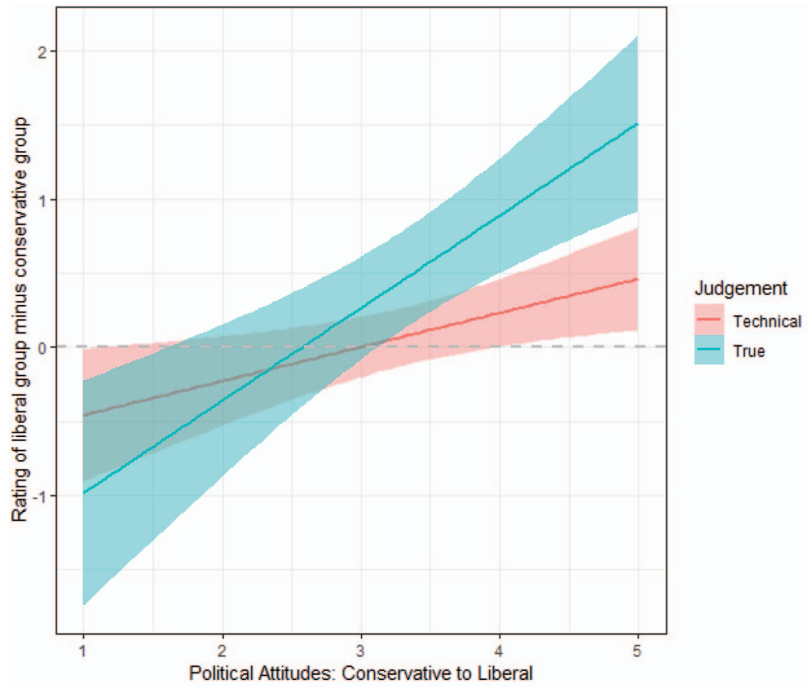

Figure 7. Ratings of the liberal-leaning group minus conservative-leaning group as a function of participants' political attitudes. The figure was made using geom_smooth from ggplot2, the colored region surrounding the line is the 95% confidence interval. See the online article for the color version of this figure.

General Discussion

We investigated how participants construe social groups. We considered whether participants believe groups are grounded in collective recognition or grounded in function. We found that participants' default concept of groups is grounded in collective recognition (Studies 1–4): A group persists over time because its members or a relevant third-party continue to collectively recognize the members as a group. Function was not irrelevant to participants' concepts of groups though. Instead, groups were construed in a dual-character manner. When thinking about the *true* group, and whether the group *truly* still exists, participants relied on function—that is, whether the group still realizes its basic function. For example, whether a nonprofit organization still realizes its mission statement, or whether it had drifted off course.

These findings have multiple implications for theories of conceptual representation. We extend the proposal that social groups are grounded in "mutual intentions" (Noyes & Dunham, 2017; which is roughly analogous to collective recognition), in important ways. We show that this account generalizes not only to membership judgments but to judgments about how entire groups persist over time. Our account is consistent with the conclusion that participants' default concept of groups is grounded in what we call collective recognition—but we also show that collective recognition is not restricted to ingroup members (which Noyes & Dunham, 2017 assumed) but can include relevant third parties. Even more critically, we show that collective recognition operates in conjunction with an additional sense of groups (the "true" sense) grounded in function and we provide an account of how and why these two senses are related as they are. That is, collective recognition provides more proximate (and precise) conditions for deciding whether individuals constitute a group and tracking a single group over time. Therefore, collective recognition is used for default judgments of group existence. Yet, function provides a deeper sense of why groups exist, and so underlies an additional *true* judgment.

Institutional entities are not limited to social groups. There are also institutional objects like money. Both of these kinds of institutional entities involve collective recognition. Dollar bills are only money when they are collectively recognized as money, and individuals only become a group when they are collectively recognized as a group (Noyes & Dunham, 2017; Noyes et al., 2018). More work is needed to characterize concepts of institutional entities. Other domains, like natural kinds and artifacts, have been extensively studied (see Gelman, 2004, and Carrara & Mingardo, 2013, for reviews of these domains, respectively). We converge with recent work suggesting that concepts of institutional entities involve distinct intuitive theories that cannot be predicted from natural kinds or artifacts (Noyes & Dunham, 2017; Noyes et al., 2018). Neither of these accounts incorporate collective recognition. Thus, for a complete picture of ordinary concepts, institutional entities need to be studied more comprehensively, and they need to be considered as a distinct conceptual domain that cannot be subsumed into natural kinds or artifacts.

These results have implications for social psychological accounts of real-world group schisms. One model suggests that groups split when ideologically distinct subgroups disagree about the "essential" nature of the group (Sani, 2005; Sani & Reicher, 1998, 2000). Our account, grounded in third-party judgments about novel groups, fits with this model. Participants do in fact entertain an additional "true" sense of groups, which is distinct from their judgments about whether the group actually exists or not. Moreover, participants' ideologies influence their judgments about the "true" group. This ideological influence could lead to situations where a subgroup believes that if the other subgroup gets their way, the "true" group will have ended. This could prompt them to leave the group, believing their break-away group will be the "true" group instead.

These results have real-world implications. As described above, we show how participants' concepts can lead to a self-fulfilling prophecy: People that disagree about the "true" nature of the group can bring about the end of the "technical" group when they decide to split. Many other real-world decisions about groups could also be influenced by participants' concepts of groups. For example, a decision about how to apportion funding can be considered as a judgment about which group is "technically" the same as the original group, but it can also be considered a judgment about which group is "truly" the same as the original. Similarly, whether one blames a group (e.g., Democratic Party) for a prior action (e.g., supporting slavery) could be influenced by whether one considers blame as belonging to the "technical" successor of the group or the "true" successor of the group. Subtle reframing could influence the decision that people make.

Social groups are an important part of the ordinary lives of human beings. People cooperate to form these entities in order to solve certain problems, and thus to realize certain functions. We show that the cooperative act itself, of coming together to recognize individuals as part of a greater collective, is the basic causal framework people use to reason about groups. But people also track whether groups succeed in realizing their functions. For example, when the United States was founded there was the formal act. The founding members ratified the constitution and recognized themselves as a new nation, and other nations followed suit in recognizing them as well. There was also an aspirational act though. There was a shared vision for the United States, exemplified by the six goals laid out in the constitutions' preamble.

> *We the People of the United States, in order to form a more perfect Union, establish Justice, insure domestic Tranquility, provide for the common defense, promote the general Welfare, and secure the Blessings of Liberty to ourselves and our Posterity, do ordain and establish this Constitution for the United States of America.*

According to participants' intuitive theories of social groups, the United States will persist so long as people continue to recognize it. This is a somewhat minimal standard. People can recognize the nation across dramatic changes in its policies. But, the United States considered as a nation that realizes a basic function, enacting the six goals laid out in the constitution, is more fragile. The "true" United States may falter if it strays too far from these values, such as whether it drifts into fascism. Or, indeed, the "true" United States may still not exist. As we have shown above, this conclusion will depend on one's own values.

Context

Philosophically, institutional entities like money, social roles, and social groups are a distinct domain of entities that require different forms of explanation than natural kinds or artifacts (Guala, 2016; Searle, 2010). Theories of conceptual representation have not typically considered these entities, focusing instead on natural kinds and

artifacts and using those two cases to make connections made between philosophy and cognitive psychology (see Bloom, 1998 and Gelman, 2003). This neglect of a ubiquitous conceptual domain in daily life motivated the authors and collaborators to understand people's intuitive theories of institutional entities (e.g., Noyes & Dunham, 2017; Noyes et al., 2018). This specific article emerged out of this larger research program. The specific research questions presented here were inspired by three primary sources: First, we were inspired by the contrast between collective recognition-based approaches (e.g., Searle, 2010) and functionalist approaches (Guala, 2016) in the relevant philosophical literature. Second, we were inspired by dual character accounts (Knobe et al., 2013) and their relationship to institutional entities. Third, we were inspired by research exploring how participants think about the identity of individual objects and persons (e.g., Rips et al., 2006). This motivated a major extension of Noyes and Dunham (2017), to examine collective recognition and function, whether groups are dual character, and to consider groups as individual entities.

References

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20,* 351–368. http://dx.doi.org/10.1093/pan/mpr057

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition, 60,* 1–29. http://dx.doi.org/10.1016/0010-0277(95)00699-0

Bloom, P. (1998). Theories of artifact categorization. *Cognition, 66,* 87–93.

Bookman, M. Z. (1994). War and peace: The divergent breakups of Yugoslavia and Czechoslovakia. *Journal of Peace Research, 31,* 175–187. http://dx.doi.org/10.1177/0022343394031002005

Burnham, M. J., Le, Y. K., & Piedmont, R. L. (2018). Who is MTurk? Personal characteristics and sample consistency of these online workers. *Mental Health, Religion & Culture, 21,* 1–11. http://dx.doi.org/10.1080/13674676.2018.1486394

Carrara, M., & Mingardo, D. (2013). Artifact categorization: Trends and problems. *Review of Philosophy and Psychology, 4,* 351–373. http://dx.doi.org/10.1007/s13164-013-0151-6

Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences, 37,* 461–480.

Del Pinal, G., & Reuter, K. (2017). Dual character concepts in social cognition: Commitments and the normative dimension of conceptual representation. *Cognitive Science, 41,* 477–501. http://dx.doi.org/10.1111/cogs.12456

Dyck, B. (1997). Exploring organizational family trees: A multigenerational approach for studying organizational births. *Journal of Management Inquiry, 6,* 222–233. http://dx.doi.org/10.1177/105649269763007

Ebrahim, A., Battilana, J., & Mair, J. (2014). The governance of social enterprises: Mission drift and accountability challenges in hybrid organizations. *Research in Organizational Behavior, 34,* 81–100. http://dx.doi.org/10.1016/j.riob.2014.09.001

Gelman, S. A. (2003). *Oxford series in cognitive development. The Essential Child: Origins of Essentialism in Everyday Thought*. New York, NY: Oxford University Press.

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8,* 404–409. http://dx.doi.org/10.1016/j.tics.2004.07.001

Gelman, S. A. (2013). Artifacts and essentialism. *Review of Philosophy and Psychology, 4,* 449–463.

Gilbert, M. (2013). *Joint commitment: How we make the social world*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199970148.001.0001

Guala, F. (2016). *Understanding institutions: The science and philosophy of living together*. Princeton, NJ: Princeton University Press.

Gutheil, G., Bloom, P., Valderrama, N., & Freedman, R. (2004). The role of historical intuitions in children's and adults' naming of artifacts. *Cognition, 91,* 23–42. http://dx.doi.org/10.1016/S0010-0277(03)00165-3

Gutheil, G., Gelman, S. A., Klein, E., Michos, K., & Kelaita, K. (2008). Preschoolers' use of spatiotemporal history, appearance, and proper name in determining individual identity. *Cognition, 107,* 366–380. http://dx.doi.org/10.1016/j.cognition.2007.07.014

Gutheil, G., & Rosengren, K. S. (1996). A rose by any other name: Preschoolers' understanding of individual identity across name and appearance changes. *British Journal of Developmental Psychology, 14,* 477–498. http://dx.doi.org/10.1111/j.2044-835X.1996.tb00719.x

Huff, C., & Tingley, D. (2015). "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics*. Advance online publication. http://dx.doi.org/10.1177/2053168015604648

Jones, M. B. (2007). The multiple sources of mission drift. *Nonprofit and Voluntary Sector Quarterly, 36,* 299–307. http://dx.doi.org/10.1177/0899764007300385

Keil, F. C. (1989). *Concepts, kinds and development*. Cambridge, MA: Bradford Books/MIT Press.

Knobe, J., Prasada, S., & Newman, G. E. (2013). Dual character concepts and the normative dimension of conceptual representation. *Cognition, 127,* 242–257. http://dx.doi.org/10.1016/j.cognition.2013.01.005

Leslie, S. J. (2015). "Hillary Clinton is the only man in the Obama administration": Dual character concepts, generics, and gender. *Analytic Philosophy, 56,* 111–141. http://dx.doi.org/10.1111/phib.12063

Mallon, R. (2016). *The construction of human kinds*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198755678.001.0001

Malt, B. C. (1990). Features and beliefs in the mental representation of categories. *Journal of Memory and Language, 29,* 289–315. http://dx.doi.org/10.1016/0749-596X(90)90002-H

Malt, B. C., & Sloman, S. A. (2007). Category essence or essentially pragmatic? Creator's intention in naming and what's really what. *Cognition, 105,* 615–648.

Millikan, R. G. (1998). A common structure for concepts of individuals, stuffs, and real kinds: More mama, more milk, and more mouse. *Behavioral and Brain Sciences, 21,* 55–65.

Molouki, S., & Bartels, D. M. (2017). Personal change and the continuity of the self. *Cognitive Psychology, 93,* 1–17. http://dx.doi.org/10.1016/j.cogpsych.2016.11.006

Noyes, A., & Dunham, Y. (2017). Mutual intentions as a causal framework for social groups. *Cognition, 162,* 133–142. http://dx.doi.org/10.1016/j.cognition.2017.02.007

Noyes, A., Keil, F. C., & Dunham, Y. (2018). The emerging causal understanding of institutional objects. *Cognition, 170,* 83–87. http://dx.doi.org/10.1016/j.cognition.2017.09.008

Rabbie, J. M., & Horwitz, M. (1988). Categories versus groups as explanatory concepts in intergroup relations. *European Journal of Social Psychology, 18,* 117–123. http://dx.doi.org/10.1002/ejsp.2420180204

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21–59). Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511529863.004

Rips, L. J., Blok, S., & Newman, G. (2006). Tracing the identity of objects. *Psychological Review, 113,* 1–30. http://dx.doi.org/10.1037/0033-295X.113.1.1

Rosenberg, A. (2016). *The philosophy of social science*. Boulder, CO: Westview Press.

Sani, F. (2005). When subgroups secede: Extending and refining the social psychological model of schism in groups. *Personality and Social Psy-*

chology Bulletin, 31, 1074–1086. http://dx.doi.org/10.1177/0146167204274092

Sani, F., & Reicher, S. (1998). When consensus fails: An analysis of the schism within the Italian Communist Party (1991). *European Journal of Social Psychology, 28,* 623–645. http://dx.doi.org/10.1002/(SICI)1099-0992(199807/08)28:4<623::AID-EJSP885>3.0.CO;2-G

Sani, F., & Reicher, S. (2000). Contested identities and schisms in groups: Opposing the ordination of women as priests in the Church of England. *British Journal of Social Psychology, 39,* 95–112. http://dx.doi.org/10.1348/014466600164354

Searle, J. (1995). *The social construction of reality*. New York, NY: The Free Press.

Searle, J. (2010). *Making the social world: The structure of human civilization*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:osobl/9780195396171.001.0001

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition, 131,* 159–171. http://dx.doi.org/10.1016/j.cognition.2013.12.005

Received September 5, 2018
Revision received July 30, 2019
Accepted October 5, 2019 ■