Journal of Experimental Child Psychology 185 (2019) 19–34

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Minimal but meaningful: Probing the limits of randomly assigned social identities

Xin Yang *, Yarrow Dunham

Department of Psychology, Yale University, New Haven, CT 06511, USA

ARTICLE INFO

Article history:
Received 10 December 2018
Revised 26 March 2019

Keywords:
Minimal group
Ingroup bias
Meaningfulness
Essentialism
Prejudice
Discrimination

ABSTRACT

The current studies (total $N = 151$) experimentally manipulated meaningfulness in novel social groups and measured any resulting ingroup biases. Study 1 showed that even when groups were arbitrary and presumptively meaningless, 5- to 8-year-old children developed equally strong ingroup biases as children in more meaningful groups. Study 2 explored the lengths required to effectively reduce ingroup biases by stressing the arbitrariness of the grouping dimension. Even in this case, ingroup bias persisted in resource allocation behavior, although it was attenuated on preference and similarity measures. These results suggest that one needs to go to great lengths to counteract children's tendency to imbue newly encountered social groups with rich affiliative meaning.

Introduction

As children navigate the social world, they quickly notice that humans cluster into categories, which we here refer to as *social groups*. These groups are immensely varied, ranging from rich cultural niches that may exert enduring influences on individuals throughout their lives (e.g., nationality, gender, ethnicity) to more superficial collectives that may form and dissolve readily (e.g., teams picked on a playground, individuals who happen to take the same bus). Intuitively, it seems obvious that people are more attached to groups of the former kind than the latter kind; in simplest terms, we might say that groups vary in meaningfulness and that individuals care more about, and perhaps more strongly

* Corresponding author.
E-mail address: xin.yang@yale.edu (X. Yang).

https://doi.org/10.1016/j.jecp.2019.04.013
0022-0965/

prefer, groups that are more meaningful. The current work presents an experimental test of this possibility.

Much past work has focused on real-world groups such as race, religion, and ethnicity. These groups reflect rich assumptions about people such as shared personality traits, endorsed cultural norms, and long-standing histories of association. Ingroup bias (i.e., a relative preference for ingroups over outgroups) has been found in children for many such groups, including gender (e.g., Yee & Brown, 1994), race/ethnicity (e.g., Aboud, 1988), language (e.g., Kinzler, Dupoux, & Spelke, 2007), nationality (e.g., Barret, 2007), and religion (e.g., Heiphetz, Spelke, & Banaji, 2013). It is intuitive to think that ingroup bias develops because these groups are meaningful to us.

However, research on presumptively less meaningful groups provides us with surprising results on how ingroup biases can arise even under much more minimal conditions. Beginning during the 1970s, Tajfel and colleagues conducted the first "minimal" group study. They initially planned to start with a condition where they did not expect bias to emerge and then add meaning to the groups to induce bias and test various causal models of the origins of ingroup bias. Contrary to their expectations, however, they found that people gave more resources to their ingroup members even when the groups were based on seemingly trivial dimensions such as tendencies to over- or under-estimate dot arrays or taste in abstract art (Tajfel, 1970; Tajfel, Billig, Bundy, & Flament, 1971). In other words, participants showed ingroup biases even in experimentally created and otherwise arbitrary groups (so-called minimal groups). Since then, other studies have replicated and extended their findings using various attitudinal and behavioral measures, including other group induction methods such as random assignment and participant choice (e.g., Brewer & Silver, 1978; Perreault & Bourhis, 1999; for reviews, see Dunham, 2018, and Otten, 2016). In line with the original Tajfel studies, participants have been shown to more positively evaluate their minimal ingroup members, allocate more resources to them, and hold stronger implicit biases toward them. Importantly, these findings indicate that mere social categorization is sufficient, and negative intergroup interactions such as conflict and competition are not necessary to induce intergroup biases. Furthermore, minimal group biases are also found in young children (e.g., Dunham, Baron, & Carey, 2011), appearing by 3 years of age (Richter, Over, & Dunham, 2016), suggesting that these effects do not depend on prolonged experiences with social groups.

These findings lead to the question of how minimal group biases compare with biases in real groups. A reasonable intuition is that, due to their lack of real-world significance, minimal group biases are weaker than real group biases. However, meta-analyses comparing ingroup biases between real and minimal groups in adults have provided mixed evidence on this point, with classic attitude measures supporting the suggestion of stronger bias toward real-world groups (Mullen, Brown, & Smith, 1992) but measures of behavioral discrimination suggesting that, if anything, bias in minimal group contexts is on average stronger than that in real groups (Lane, 2016). However, because these meta-analytic syntheses aggregated across many different kinds of real and minimal groups while not measuring specific dimensions on which various groups vary, they provide only an indirect window into the factors underlying variation in the strength of ingroup bias.

Here we call attention to two (non-exhaustive) factors that could plausibly underlie differences in the strength of ingroup bias: familiarity and meaningfulness. In most cases with real-world groups, familiarity and meaningfulness are likely to be associated in that increased experience with a group provides both increasing familiarity and an increasingly strong foundation for meanings to emerge. That said, they are in principle dissociable. Consider the extent to which newly encountered, and thus equally unfamiliar, groups can vary in meaning (e.g., a group created based on a coin flip vs. a group based on the results of a genetic test implying differential health risk). In what follows, we first talk about how familiarity and meaningfulness each affects the strength of ingroup bias and then propose a way to experimentally test the role of meaningfulness while holding familiarity constant.

Beginning with familiarity, minimal groups are unfamiliar by design in that people have no prior experience or expectation of future interactions with group members. In contrast, people have extensive experience with many real groups, for example, by encountering people who speak the same language or hold the same religious views more frequently than those who differ along these dimensions. There is good reason to think that familiarity tends to result in more ingroup liking. For example, mere familiarity is associated with positive attitudes toward a wide range of objects (Zajonc, 1968), and

some past research on groups has provided more direct empirical evidence that increasing familiarity with an ingroup leads to stronger preferences in favor of that group (e.g., Tajfel & Billig, 1974). Thus, there is a well-established role for familiarity in producing ingroup bias.

Turning to meaningfulness, there is considerably less empirical data directly testing this question. As we noted, many familiar real-world groups are rich in meaning. By contrast, minimal group memberships are generally assigned according to ostensibly unimportant features, in particular features that do not have an established relationship to personal identity such as the tendency to over- or under-estimate dot arrays. Do more meaningful groups lead to more ingroup bias? Evidence is mixed. Some work has shown that when teachers use minimally assigned social groups as a means to organize social activities (thereby providing a cue to social relevance), ingroup bias in children increased (Bigler, Jones, & Lobliner, 1997). However, a more recent study found that minimal groups and groups based on shared interests produced equivalent levels of ingroup bias despite the latter presumably being more meaningful (Sparks, Schinkel, & Moore, 2017). As far as we know, there is no direct experimental test of the relationship between meaningfulness and ingroup bias. Providing such a test was the primary goal of the current work.

As we alluded to above, familiarity and meaningfulness are generally confounded in real-world groups, making it difficult to assess their independent impacts. In the current research, we explored the role of meaningfulness by manipulating the putative nature of the group while controlling for familiarity. More specifically, we assigned children to unfamiliar groups but varied the assignment procedure such that, if our manipulation was successful, half of the children would associate the group with a higher degree of meaning than the other half. In this way we provide the first experimental test of the link between meaningfulness and ingroup bias.

Meaningfulness can surely be operationalized in many ways. In the current work, we took inspiration from the literature on essentialism (e.g., Gelman, 2004), manipulating the meaningfulness of groups by varying the extent to which our group assignment procedure implied deep and stable differences (as opposed to superficial properties), consistent with a "kind-like" construal of the groups in question. Intuitively, a group defined via deep and stable properties is more meaningful almost by definition than a group defined via shallow and transient properties. We assessed the effectiveness of our manipulation via manipulation check questions (adapted from the essentialism literature) while also using classic intergroup bias measures to assess whether children who were assigned to a more meaningful group showed stronger ingroup biases than those assigned to a less meaningful group.

Essentialist reasoning is complex, for example, incorporating beliefs that category membership is stable, difficult or impossible to change, and inductively rich (Gelman, 2004; Rhodes & Mandalaywala, 2017). Many of our most meaningful social category memberships, including race, gender, and ethnicity, have been associated with essentialist thought (Haslam, Rothschild, & Ernst, 2000; Prentice & Miller, 2007). For these reasons, we expected our manipulation to increase the meaning that children associate with the groups. That said, it is important to also acknowledge a more general debate concerning the relationship between essentialism and prejudice, with some correlational studies reporting a positive relationship (Chen & Ratliff, 2018; Yzerbyt, Rogier, & Fiske, 1998), others reporting a negative relationship (Herek & Capitanio, 1995; Whitley, 1990), and still others reporting no relationship (Haslam, Rothschild, & Ernst, 2002). Experimental work on this topic is similarly mixed (Andreychik & Gill, 2015; Chao, Hong, & Chiu, 2013; Chen & Ratliff, 2018; Diesendruck & Menahem, 2015; Keller, 2005; Rangel & Keller, 2011; Williams & Eberhardt, 2008). One recent developmental study conducted while the current research was under way (Rhodes, Leslie, Saunders, Dunham, & Cimpian, 2017) manipulated essentialist beliefs concerning novel social groups by varying exposure to generic language, a form of input previously linked to essentialism (Rhodes, Leslie, & Tworek, 2012). It reported that essentialist beliefs did not affect attitudinal prejudice but did influence children's resource allocation decisions (i.e., sharing fewer resources with outgroup members), suggesting that different mechanisms underlie intergroup attitudes versus behaviors (cf. Dunham, 2018). By manipulating a component of essentialism in a different way and including a wider range of measures of intergroup bias, we hoped to shed further light on this issue.

To summarize, we assigned children to unfamiliar groups and then provided information about the nature of the group. In one condition we emphasized that groups were based on deep and internal aspects of who the children were, whereas in the other condition we emphasized that groups were

randomly assigned. To confirm effects of the manipulation, we examined children's views of the inductive potential, innateness, stability, and informativeness of the groups, which we took to be important (although non-exhaustive) aspects of meaningfulness. To measure ingroup bias, we used both attitude and behavior items, allowing us to reveal any potential differences in how those measures play out. Finally, we focused on 5- to 8-year-olds, a similar age range as tested in relevant recent studies (e.g., Rhodes et al., 2017), both to compare with that past work and because this is an age range over which ingroup biases in minimal groups are robustly shown (Dunham et al., 2011).

We predicted that children who were randomly assigned to the groups would construe the groups to be less meaningful and hold weaker ingroup biases compared with children in more meaningful groups. If children in both conditions instead showed equivalent levels of ingroup biases despite reporting differences in meaningfulness, it would imply that meaning, at least as manipulated here, is not causally related to ingroup bias. To forecast the actual results, our primary hypothesis was not confirmed in any straightforward way. In Study 1, children acknowledged the difference in meaning associated with the two kinds of groups but held largely equivalent degrees of ingroup bias. Only when we went to extremes to remove meaning from the minimal group in Study 2 were we able to reduce bias on some measures. These findings suggest that children imbue even relatively meaningless groups with rich affiliative meaning.

Study 1

In Study 1, we tested whether assigning children to a presumably "minimal" group would result in weaker intergroup biases compared with assigning them to a presumably more meaningful "maximal" group. To experimentally manipulate the level of meaningfulness, we had children place their hand on a machine that we told them would assign them to one of two groups, but we told them different stories about how the machine worked. Half of the children were told that the machine could look deep inside them to reveal their true category membership, whereas the other half of the children were told that the machine assigned them to a group randomly. Thus, the primary question was whether we would induce stronger intergroup biases in the former condition (the maximal condition) compared with the latter condition (the minimal condition).

Method

Participants

There were 54 participants (M_{age} = 7.11 years, range = 5.1–8.9; 23 boys, 28 girls, and 3 did not report gender; predominantly White; primarily middle-class residents of Connecticut in the northeastern United States), with this sample size selected based on past studies in this area (e.g., Dunham et al., 2011) as well as our lab default policy of running 30 participants per condition when prior information about effect sizes is unknown; in fact, we tested 56 participants, 2 of whom were excluded for not providing responses during the data collection phase. Of the 54 included children, 28 were assigned to the "maximal" condition (M_{age} = 7.09 years, range = 5.1–8.8; 11 boys, 14 girls, and 3 did not report gender) and 26 were assigned to the "minimal" condition (M_{age} = 7.13 years, range = 5.4–8.9; 12 boys and 14 girls). An additional 2 children were recruited but excluded for not providing responses. We acquired informed consent from parents and verbal assent from children prior to participation. All children were tested by a research assistant in a quiet testing room in the lab or at a local museum. The testing phase took 15 to 20 min.

Procedure

Children were introduced to two novel social groups, the Green Group and the Orange Group, and were shown a machine connected to a laptop computer. The machine, actually a pad for digital input intended to be used with a stylus, was a flat matte black surface a little larger than a child's hand (see Fig. S1 in online supplementary material). Children were told that that this machine would decide which group they were going to be in either by "working kind of like flipping a coin" (minimal condition) or by "looking deep inside you and really figuring out what kind of person you are" (maximal

condition). Next, children put one of their hands on the machine. A few seconds later, a series of colored shapes flashed on the screen, finally settling on one color (either green or orange). Children were then assigned to that color group. They put on a color-matched sticker and bracelet as a reminder of their group membership and were asked to report their group assignment.

Following the group assignment, children responded to nine different types of questions that were divided into two similar-length blocks with randomization within blocks. The order of the items within each measure was randomized. In general, the questions involved cartoon figures of gender-matched children wearing green or orange T-shirts. In the first block, children completed *preference, similarity, sharing and collaboration, generalization,* and *informative group cues* measures in random order except that the similarly structured *preference* and *similarity* measures were always blocked together. Children then completed the second block, which included a randomly intermixed set of *switched-at-birth, stability and change,* and *open-ended* questions as well as a set of questions concerning likely *future interactions.* Finally, we removed the group markers from the participants, put them out of sight, and asked which group the children had belonged to as a final check on successful encoding of group membership.

Measures

Measures of intergroup biases. For each item in the following three measures, children viewed two individuals, one from each group, on the screen. Ingroup responses were scored as 1, and outgroup responses were scored as 0.

Preference (6 items). Children were asked which person they liked better (3 explicit liking items) and which person they would rather play with (3 playmate preference items).

Similarity (4 items). Children were asked which person was more like them (2 items) and which person they were more like (2 items).

Sharing and collaboration (4 items). Children were asked who they would share with, who they would ask for help, who would share with them, and who would help them.

Measures of meaningfulness and essentialism. Generalization (4 items). Children were told about the hobby of a Green person or an Orange person (e.g., "A girl from the Orange Group loves to work on puzzles as much as she can") and were asked which of two persons (one from each group) shared the same hobby. Within-group generalizations were coded as 1 and were considered the more essentialist response.

Switched-at-birth (2 items). Children were told stories about a baby who was born to Green or Orange parents but was raised by parents in the other group and were asked which group they thought the baby belonged to when he or she grew up (Gelman & Wellman, 1991; Hirschfeld, 1995; Taylor, 1996; Waxman, Medin, & Ross, 2007). This sort of task has frequently been used to assess whether children think that social categories are determined by birth or the environment. Therefore, indicating that the grown-up would still be in the same group as the birth parents was coded as 1 and considered the more essentialist response.

Stability and change (3 items). Children were asked what group they were going to be in when they grew up, whether they could change groups if they really wanted to, and whether they could change groups if they changed stickers. Staying in the same group and not being able to change indicated stability and was coded as 1 and considered the more essentialist response.

Future interactions (4 items). Children were asked about future friends' and neighbors' group memberships (e.g., "When you grow up, do you think you are going to have more Orange friends or Green friends?"; "When you grow up, do you think your neighbors are going to be mostly Orange people, mostly Green people, or both?"). Ingroup guesses were coded as 1. Although not formally entailing essentialism, we considered same-group responses on this measure to reflect a form of stability and general importance of group membership and so considered a score of 1 as more consistent with an essentialist group construal.

For the four measures above, meaningful and essentialist responses were coded as 1, which (assuming acceptable reliability) we intended to sum to a composite score (range = 0–13), although we should note that some past work has found that these measures might not reflect a single construct (Gelman, Ware, & Kleinberg, 2010; for a review, see Rhodes & Mandalaywala, 2017). This score

constituted a manipulation check to see whether we observed higher meaningfulness in the maximal condition.

Additional exploratory measures. Informative group cues (4 items). Children viewed individuals who were not wearing group markers and were asked whether knowing the following (in a randomized order) would help them to determine group membership: (a) if they acted the same, (b) where they came from, (c) what they had inside their bodies, and (d) things they like to do. Children could answer "yes" or "no" to each of the four questions, and answers were dichotomously coded. This was an exploratory measure added to test the factors that children think identify people as belonging to one group or the other. However, because it has not been used in past research, we did not intend to include it in our broader composite.

Open-ended question: "How do you become a Green or an Orange?" (1 item). We included this question to investigate children's spontaneous interpretations of the group manipulation, specifically whether explanations would distinguish the randomly assigned versus discovered group. Answers were coded into categories as described in the Results section below.

Results

Data for both studies and codes to replicate all analyses are available at https://osf.io/d4f86/?view_only=0b69f10065f44c72af2022de4f6b194f. In all models, we included gender, age, and color group as control variables, but because we were not directly investigating these factors, we considered any differences as a function of these variables as exploratory rather than confirmatory.

Meaningfulness and essentialism measures

We first compared the composite score for meaningfulness across conditions (range = 0–13, with higher values indicating more essentialist beliefs), using it as an indicator of whether the maximal condition induced greater essentialist reasoning. Data were analyzed via a linear model predicting the essentialist score as a function of condition (maximal vs. minimal), controlling for participant gender, age (continuous), and color group. We found a significant effect of condition ($\beta = -1.77$, 95% confidence interval [CI] = −3.12, −.42), with children in the maximal condition ($M = 8.85$, $SE = .42$) assigning more meaning to the groups than children in the minimal condition ($M = 7.16$, $SE = .45$), $d = .76$. However, a reliability analysis indicated that the four measures that contributed to the composite did not reliably cohere ($\alpha = .17$, 95% CI = −.18, .52), suggesting that a composite was not appropriate. Thus, we provide a summary of the effect of condition on all four essentialism-related measures when treated separately in Table 1; in brief, the manipulation had a reliable effect on the generalization and stability and change measures but not on switched-at-birth or expectations about future interactions.

Table 1
Results of one-sample *t* tests (two-tailed, chance = .50) for essentialism measures, separated by condition, and two-sample *t* tests for conditional differences in Study 1.

Measure	Maximal	Minimal	Difference			
	M (95% CI)	*M* (95% CI)	M_{Diff}	*t*	*df*	*d*
Generalization	.64 (.51, .77)*	.38 (.23, .52)$^{\sim}$	.27	2.78**	50.57	.76
Switched-at-birth	.59 (.41, .77)	.63 (.46, .81)	−.04	−0.34	50.99	.09
Stability and change	.76 (.65, .86)***	.52 (.41, .63)	.24	3.17**	48.57	.89
Future interactions	.66 (.55, .78)**	.74 (.64, .85)***	−.08	−1.05	51.80	.29

Note. Conditional differences were found in the generalization and stability and change measures. CI, confidence interval.
$^{\sim}$ $p < .10$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Additional exploratory measures

For the informativeness of group cues measure, we observed a strong "yes bias" in children's responses, precluding strong interpretations; we confine further discussion of this measure to supplementary material. For the open-ended question, we coded children's responses into the following types: behavior, essentialism, interaction, markers, machine, and unknown. Unfortunately, the overall rate of codable responses was low, limiting our ability to perform quantitative analysis. Details of coding criteria and general trends are described in the supplementary material; in brief, children in the minimal condition made more references to group markers and the machine's direct role in group assignment, whereas children in the maximal condition mentioned deeper factors such as the likelihood of future interactions with their group. Thus, although this measure should be interpreted cautiously, it suggested that our manipulation was successful.

Intergroup bias measures

For each of the three intergroup bias measures, we fitted a binomial linear mixed effects model predicting bias as a function of condition, controlling for participant gender, age (continuous), and color group (fixed effects), with a random intercept for participant. There was no effect of condition in any of these models ($ps > .61$). Overall, children in both conditions showed equivalently strong ingroup bias (for preference, $M_{Max} = .80$, $SE = .03$; $M_{Min} = .80$, $SE = .03$; for similarity, $M_{Max} = .90$, $SE = .03$; $M_{Min} = .88$, $SE = .03$; for sharing and collaboration, $M_{Max} = .76$, $SE = .04$; $M_{Min} = .74$, $SE = .04$; possible range = 0–1, with higher values indicating more ingroup biases). Although the manipulation failed to influence ingroup bias, it is important to note that children in both conditions tended to show strong ingroup favoritism across all three measures. A summary of results is provided in Table 2.

Relations between measures

A useful and somewhat unusual feature of our design is the large number of within-child measures; Table 3 presents the bivariate correlations among all main measures.

Correlation results between intergroup bias measures and meaningfulness and essentialism measures showed that only the expectations for future interactions measure correlated with the three intergroup bias measures, whereas the other three meaningfulness and essentialism measures did not; Children who expected more future interactions with ingroup members held stronger ingroup biases.

Discussion

Using different cover stories to assign children into maximal and minimal groups, we successfully created groups that varied in meaningfulness, at least in terms of several of the measures. Counter to our prediction, however, despite these differences, children held equally strong ingroup biases across all measures. These results imply that the mere fact of being in a group is sufficient to induce strong ingroup biases, however random and meaningless the group assignment is. Our findings are consistent with past work on the minimal group effect, which shows that both adults and children hold ingroup

Table 2
Results of one-sample *t* tests (two-tailed) for intergroup bias measures, separated by condition, and two-sample *t* tests for conditional differences in Study 1.

Measure	Maximal	Minimal	Difference			
	M (95% CI)	*M* (95% CI)	M_{Diff}	*t*	*df*	*d*
Preference	.80 (.68, .92)***	.80 (.71, .89)***	−.00	−0.05	49.61	.01
Similarity	.90 (.84, .97)***	.88 (.78, .97)***	.03	0.49	46.49	.13
Sharing and collaboration	.76 (.66, .86)***	.74 (.62, .86)***	.02	0.24	48.52	.07

Note. All results were significantly greater than chance (.50) in both conditions, and there were no significant conditional differences. CI, confidence interval.
*** $p < .001$.

Table 3
Correlations among all intergroup bias and meaningfulness/essentialism measures in Study 1.

Measure	1	2	3	4	5	6
Preference	1.00					
Similarity	.57***	1.00				
Sharing and collaboration	.60***	.52***	1.00			
Generalization	.03	.12	.19	1.00		
Switched-at-birth	.09	−.17	.02	.02	1.00	
Stability and change	.22	.01	.19	.08	.21	1.00
Future interactions	.42**	.47***	.38**	.00	−.04	.05

Note. All three intergroup bias measures were reliably intercorrelated ($\alpha = .78$), but meaningfulness/essentialism measures did not ($\alpha = .17$).
** $p < .01$.
*** $p < .001$.

biases across a range of different minimal settings (e.g., groups based on shared painting preferences and the tendency to under- or over-estimate dot arrays).

Our findings make several additional contributions to this literature. First, we document strong ingroup biases even when group assignment is explicitly described as coming about via a random process. Second, and more important, we show that a manipulation that heightens the meaning associated with the minimal group does not increase ingroup bias. This demonstrates that meaningfulness does not inevitably increase ingroup bias (although of course it may well do under some conditions). Why did we observe this pattern? One potential answer is that children imbue minimal groups with such rich affiliative meanings that they are essentially "at ceiling" even prior to our manipulation of meaningfulness. If so, rather than attempting to *increase* meaning in the maximal condition, observing a conditional difference might require *decreasing* meaning in the minimal condition; we pursued this strategy in Study 2, below.

An alternative worry is that children might not have treated the random assignment as truly random. This might occur if they did not fully understand that the group assignment procedure in the minimal condition was random. Although we emphasized that the machine behaved like flipping a coin, children might not share adults' view of randomness (Piaget & Inhelder, 1976). Rather, they might treat random phenomena as deterministic (Banerjee & Bloom, 2015; Metz, 1998), for example, thinking that they could control whether a coin lands on heads or tails or that a coin lands on a certain side for some deep underlying reason. We should note that strong versions of this possibility seem unlikely for several reasons. First, past research has documented that children of this age range do understand randomness in some forms (Kuzmak & Gelman, 1986). Second, we did observe different levels of reported meaningfulness across the two conditions, providing some evidence that children in the minimal condition saw the groups differently than children in the maximal condition. However, we considered this worry enough of a live possibility to address in Study 2, via efforts to further reduce the meaning and importance of the minimal condition, as we describe below.

In summary, although we entered this line of work with the intuition that to differentiate minimal and maximal groups we would need to increase the meaning associated with the maximal condition, the results from Study 1 actually suggest that we need to decrease the meaning of the minimal condition still further. This became the primary goal of Study 2.

Study 2

In Study 1, we found that even in the minimal condition children showed similarly strong ingroup biases as those in the maximal condition, raising the possibility that our minimal condition was not so minimal after all. Therefore, to emphasize the low level of meaningfulness of groups in the minimal condition, and to test whether ingroup biases can be reduced in extremely meaningless groups, we conducted a preregistered Study 2 (preregistration link: https://aspredicted.org/blind.php?x=7gf5ga) with several changes geared toward making this clearer. The purpose of Study 2 was to explore the lengths required to reduce minimal group biases. In brief, we demonstrated randomness using a real

quarter coin to help children understand randomness and, more important, *switched* the groups for children in the minimal condition in order to emphasize that group assignment was arbitrary, unimportant, and very unlikely to be based on deep internal features of the participants.

Method

Participants

There were 97 participants (M_{age} = 7.08 years, range = 5.1–8.9; 48 boys and 49 girls; 40% White, 6% Asian, 5% Black, 3% Hispanic, and the remaining 46% did not report race/ethnicity; primarily middle-class residents of Connecticut). Our preregistered sample size was 80, which was larger than Study 1 because we aimed to increase the power to detect conditional differences. We exceeded our preregistered target of 80 children following lab policy to continue testing on a given study until we have filled in all age and condition cells even if that results in a larger sample size in other cells. However, data were not analyzed until data collection was complete. Of the 97 children, 54 were assigned to the maximal condition (M_{age} = 7.08 years, range = 5.1–8.9; 27 boys and 27 girls) and 43 were assigned to the minimal condition (M_{age} = 7.08 years, range = 5.1–8.8; 21 boys and 22 girls). An additional 10 children were recruited but excluded (7 children failed the final memory check question, 2 children did not complete the study, and 1 child refused to participate). We acquired informed consent from parents and vocal consent from children. All children were tested by the first author or a research assistant in a quiet testing room in the lab, at a local museum, or at a local science center. The testing phase took about 10 min.

Procedure

The procedure was identical to that in Study 1 except for the following changes. First, to help children understand randomness, we demonstrated it using a coin (a quarter) as we described less meaningful groups to children in both conditions: "When we flip a coin, we sometimes get heads and sometimes get tails. So if we just put people in one group if we get heads [demonstrate with the coin] and put them in the other group if we get tails [demonstrate with the coin], this wouldn't really tell you very much about what people are like."

Second, to further indicate that the group assignment was totally arbitrary and meaningless in the minimal condition, after the machine assigned the children to a color group, the experimenter told them that the lab had run out of materials of that color and that they would instead be assigned to the other color group. Then we gave them stickers and bracelets that matched the other color.

Thus, by providing more emphasis of chance processes as well as actually overriding the machine's assignment to put the children in the other group, we hoped to emphasize the shallowness of the grouping procedure. The maximal condition was the same as that described in Study 1 except for minor changes to make the wording parallel across conditions given the changes to the minimal condition. As in Study 1, we asked children which group they were in just after the machine picked the color for them (but before the group switch in the minimal condition), and all children answered correctly.

After the final group assignment (the original group assignment in the maximal condition and the group switching in the minimal condition), children completed two sets of measures, one focusing on intergroup biases, including *preference, similarity,* and *resource allocation* and the other focusing on group meaningfulness and essentialism, including *generalization, switched-at-birth,* and *stability and change* items modeled on Study 1. The six types of questions followed a randomized order. In addition, the items within each type of question were randomized. Unlike Study 1, we did not organize questions into two blocks because this study was much shorter after dropping some measures. As in Study 1, the questions also involved cartoon figures of gender-matched children wearing green or orange T-shirts. Finally, we removed the group markers and put them out of sight from the participants and asked which group the children ended up being in for the game as a final check on successful encoding of group membership. All but 7 children answered this item successfully, and we excluded those children who failed to answer this item from data analyses because a failure to encode their group membership meant that we had no substantive prediction concerning them.

Measures

Measures of intergroup biases. Preference (4 items). This measure was identical to that in Study 1, with the only changes being in the number of items (the number of explicit liking and playmate preference items was 2 instead of 3). Responses were forced choice, with 1 indicating an ingroup response.

Similarity (4 items). We added concrete example properties in the questions (e.g., instead of asking "Who is more like you," we asked "Who *likes the same kind of snacks* as you?"). Responses were again forced choice, with 1 indicating an ingroup response.

Resource allocation (3 items). Children completed 3 trials, where they were asked to distribute one, three, or five stickers between an ingroup child and an outgroup child by putting the stickers into green and orange boxes. They could not keep any stickers for themselves or leave any stickers on the table. We calculated the average number of stickers given to ingroups across 3 trials (range = 0–3).

Measures of meaningfulness and essentialism. Generalization (4 items). This measure was identical to that in Study 1.

Switched-at-birth (2 items). We adapted this task from the inheritance measure in Rhodes et al. (2017), using specific properties rather than category membership. For example, the birth parents "always clap their hands when they are happy," whereas the foster parents "always laugh out aloud when they are happy." Then we asked whether the child, once grown up, would exhibit the birth parents' or foster parents' property. We then asked whether it was possible that the person would also exhibit the other property, a question previously used to explore discreteness of group boundaries (Rhodes et al., 2017, Study 3). Indicating that the grown-up would share the property only with the birth parents was coded as 2, indicating once in the two questions that the grown-up would share the property with the birth parents was coded as 1, and indicating that the grown-up would share the property only with the foster parents was coded as 0.

Stability and change (4 items). As in Study 1, children were asked what group they were going to be in when they grew up and whether they could change groups if they changed stickers. We also asked whether they could still stay in the same group if they lost their sticker and bracelet and whether the machine could give them the other color if they tried it again. Similarly, staying in the same group and not being able to change indicated stability and was coded as 1. One should note that in the minimal condition children were assigned to a different group than the one the machine had showed them; thus, saying that they would stay in the group that they were assigned to (but not what the machine showed them) was taken to indicate stability.

Results

Meaningfulness and essentialism measures

Following our preregistered analysis plan, we created a composite score for meaningfulness (range = 0–12, with higher values indicating more meaningfulness and essentialist beliefs). However,

Table 4
Results of one-sample *t* tests (two-tailed, chance = .50 for the generalization and stability and change measures; chance = 1.00 for the switch-at-birth task) for meaningfulness/essentialism measures, separated by condition, and two-sample *t* tests for conditional differences in Study 2.

Measure	Maximal	Minimal	Difference			
	M (95% CI)	*M* (95% CI)	M_{Diff}	*t*	*df*	*d*
Generalization	.72 (.64, .80)***	.81 (.73, .90)***	−.10	−1.68~	93.38	.34
Switched-at-birth	.94 (.86, 1.03)	1.02 (.93, 1.12)	−.08	−1.26	89.44	.26
Stability and change	.57 (.50, .64)~	.41 (.34, .47)**	.16	3.39**	94.85	.69

Note. Conditional differences were found in the stability and change measure. Whereas children in both conditions scored high on the generalization measure, children in both conditions were at chance for the switched-at-birth task. CI, confidence interval.
~ $p < .10$.
** $p < .01$.
*** $p < .001$.

a reliability analysis indicated that the individual items did not cohere, calling our use of a composite into question (α = −.01, 95% CI = −.35, .34). In Table 4, we summarize the effect of condition on all three of these measures analyzed separately; in a nutshell, the manipulation reliably affected the stability and change measure but not generalization (which in fact trended in the opposite direction) or switched-at-birth. In the interest of full disclosure, we should note that if we do use the composite from our analysis plan, the effect of condition is *not* significant (β = −.11, 95% CI = −.80, .57); children in the minimal condition (M = 6.93, SE = .23) thought of the groups as similarly meaningful as children in the maximal condition (M = 7.04, SE = .25), d = .06.

Intergroup bias measures

We begin with the confirmatory models described in our analysis plan. We found a marginally significant effect of condition on the preference measure (β = −.66, 95% CI = −1.39, .03); children in the maximal condition showed somewhat stronger ingroup bias than those in the minimal condition (M_{Max} = .68, SE = .03; M_{Min} = .56, SE = .04; possible range = 0–1, with higher values indicating more ingroup biases). Moreover, the effect of condition was significant on the similarity measure (β = −.89, 95% CI = −1.59, −.24); children in the maximal condition also showed stronger ingroup bias than those in the minimal condition (M_{Max} = .73, SE = .03; M_{Min} = .57, SE = .04; possible range = 0–1, with higher values indicating more ingroup biases). However, there was no effect of condition on the resource allocation measure (β = −.14, 95% CI = −.43, .15). Children in the two conditions behaved similarly (M_{Max} = 1.85, SE = .10; M_{Min} = 1.71, SE = .11; possible range = 0–3, with higher values indicating more ingroup biases). In addition, there was an effect of age on this measure (β = −.18, 95% CI = −.31, −.05) in that older children were fairer when allocating resources between ingroup and outgroup members. A summary of results is provided in Table 5.

Relations between measures

As in Study 1, Table 6 presents the bivariate correlations among all main measures.

Among all intergroup bias measures and meaningfulness measures, we found that the generalization and stability and change measures each correlated with preference and similarity measures; children who made more within-group generalizations and thought that groups were stable and less likely to change held stronger ingroup biases. Due to the failure of the essentialism composite to cohere, we did not use it as a predictor of ingroup bias, but we should note that this represented a deviation from our preregistered analysis plan.

Discussion

In Study 2, after going to greater lengths to reduce the meaning of the minimal condition (i.e., emphasizing the randomness of the procedure and then switching the children's initially assigned

Table 5
Results of one-sample t tests (two-tailed, chance = .50 for the preference and similarity measures; chance = 1.50 for the resource allocation measure) for intergroup bias measures, separated by condition, and two-sample t tests for conditional differences in Study 2.

Measure	Maximal	Minimal	Difference			
	M (95% CI)	M (95% CI)	M_{Diff}	t	df	d
Preference	.68 (.59, .76)***	.56 (.46, .66)	.11	1.70~	88.31	.35
Similarity	.73 (.66, .80)***	.57 (.47, .67)	.16	2.59*	79.12	.53
Resource allocation	1.85 (1.71, 2.00)***	1.71 (1.54, 1.88)*	.14	1.26	88.78	.26

Note. In the maximal condition all results were significantly greater than chance, whereas in the minimal condition only the result for resource allocation was significantly greater than chance. Conditional differences were found on the preference (p = .09) and similarity (p = .01) measures, but not on the resource allocation measure (p = .21). CI, confidence interval.
~ p < .10.
* p < .05.
*** p < .001.

Table 6
Correlations among all intergroup bias measures and meaningfulness/essentialism measures in Study 2.

Measure	1	2	3	4	5
Preference	1.00				
Similarity	.42***	1.00			
Resource allocation	.45***	.31**	1.00		
Generalization	.30**	.21*	.03	1.00	
Switched-at-birth	−.05	−.17	−.05	−.06	1.00
Stability and change	.20*	.37***	.14	.17	−.10

Note. All intergroup bias measures cohered ($\alpha = .62$), but meaningfulness/essentialism measures did not cohere ($\alpha = -.01$).
* $p < .05$.
** $p < .01$.
*** $p < .001$.

group), we observed reduced levels of ingroup biases on two of the three measures, similarity and preference, although the difference between conditions was only marginal in one of those two measures (i.e., preference). Importantly, children in the minimal condition did not show ingroup bias on these two measures. On the resource allocation measure, however, we did not observe a conditional difference; children in both conditions showed a similar degree of ingroup favoritism, giving more stickers to their ingroup. Taken together, these results suggest that meaningfulness of groups does matter for ingroup bias, but only when we go further to remove the meanings that children seem to spontaneously attribute to even minimally assigned groups. Thus, observing an effect of meaningfulness requires carefully stripping away the meanings associated with a group.

Critically, we did see evidence that our manipulation of meaningfulness was effective; children in the minimal condition believed that the groups were less stable and more prone to change than their peers in the maximal condition. However, we should note that on other measures children in both conditions indicated that the groups were similarly meaningful, and we did not observe differences in meaningfulness on the composite measure we specified in our analysis plan. Therefore, it seems that the minimal condition is not wholly meaningless despite the fact that we used explicitly random assignment and switched the group assignment for trivial reasons (running out of that color).

Some interesting patterns of relationships across measures also emerged. Most striking, we found that children who made more within-group generalizations and saw group membership as more stable had stronger ingroup biases. These results provide some evidence for a meaningfulness–bias link. However, as we return to in the General Discussion, results were not always consistent across studies and measures and so should be interpreted cautiously.

Why did bias on the resource allocation measure persist even in the minimal condition? A possible interpretation is that, even in unimportant and arbitrary groups, cooperating more with ingroup members is a useful default strategy, perhaps because individuals expect their ingroups to cooperate with them. One influential perspective on this comes from Yamagishi and Kiyonari (2000), who argued that groups serve as a container of expectations for generalized ingroup reciprocity. This is also termed "depersonalized trust," that is, the expectation that members of the same group are likely to cooperate at higher rates (Brewer, 2008), and our results suggest that it may have a different psychological foundation than other aspects of ingroup bias (for a discussion, see Dunham, 2018).

General discussion

These two studies experimentally manipulated the meaningfulness of groups and examined the relationship between levels of meaningfulness and magnitudes of ingroup bias. Unexpectedly, even in groups that we intended to be random and meaningless, 5- to 8-year-old children showed equally strong ingroup biases compared with those in groups that we intended to be more meaningful in Study 1. It was only after the group assignment process was made even more explicitly arbitrary and unimportant in Study 2 that ingroup biases on some measures were successfully reduced. These results suggest that putting children into two contrasting groups, even if relatively unimportant, is

sufficient to induce strong ingroup biases and that such biases do not necessarily become stronger when the groups are defined in richer terms.

Nonetheless, in Study 2 we showed that decreased levels of meaningfulness led to weaker ingroup biases on attitude measures. This result is consistent with some previous findings that essentialism is related to prejudice (Yzerbyt et al., 1998) and that manipulating people's essentialist beliefs about race leads to changes in racial biases (Andreychik & Gill, 2015; Chao et al., 2013; Chen & Ratliff, 2018; Diesendruck & Menahem, 2015; Keller, 2005; Rangel & Keller, 2011; Williams & Eberhardt, 2008). However, to successfully reduce ingroup biases, we needed to go to great lengths, emphasizing the randomness of group assignment with training and switching groups to further show how unimportant the groups were. And even after these efforts to make the minimal group less meaningful, we still saw similar levels of bias on the resource allocation measure. We conclude that it takes surprisingly little to elicit strong group attachments and assumptions of rich meaning in even randomly assigned social identities. Taken together, these results, at least within our experimental context, suggest that meaningfulness matters for ingroup bias at the lower (if not the lowest) end of the meaningfulness spectrum. Beyond that, children may spontaneously imbue groups with sufficient meanings that manipulations layered on top of that have little further effect on ingroup bias.

The reduced levels of ingroup biases on attitudes but not resource allocations in Study 2 could be interpreted as an attitude–behavior dissociation. This dissociation connects to past literature that documents the differences between affective and behavioral components of intergroup biases and the increased difficulty in changing behaviors compared with changing attitudes (for reviews, see Hewstone, Rubin, & Willis, 2002, and Schütz & Six, 1996). In particular, there is emerging evidence that evaluative outcomes in minimal group paradigms might systematically diverge from more cooperation or resource-based measures, a possibility that this work, as well as other recent related work, supports (Rhodes et al., 2017; for further review, see Dunham, 2018). In any case, a similar gap between attitude and behavior has been found in other fields, such as in the health domain concerning consumption, and surrounding issues such as pollution and climate change, where people might claim certain attitudes and have good intentions but fail to regularly take attitude-consistent actions (e.g., Juvan & Dolnicar, 2014; Kollmuss & Agyeman, 2010).

To our knowledge, this is the first research that experimentally manipulated meaningfulness of groups and examined ingroup biases. Our result that children imbue newly encountered randomly assigned groups with rich affiliative meanings and develop strong ingroup biases helps to illuminate the strong minimal group effect in past studies. For example, Sparks et al. (2017) showed that minimal group memberships are as strong as shared interests in affecting children's friendship preferences, perception of similarity, and sharing behavior. In addition, Bigler et al. (1997) found that randomly assigning children into groups that are then functionally used is as powerful in inducing ingroup biases as biologically based groups that are functionally used (but one should note that in their control condition, when there was no functional use, biases were weaker). Taken together, these findings demonstrate that a presumably meaningless cue could strongly influence biases, and perhaps this is because children imbue it with rich meanings as indicated by the current research.

In the current studies, we manipulated meaningfulness of groups using strategies inspired by the essentialism literature. But meaningfulness can be defined and manipulated in many other ways, and it is possible that other dimensions of meaning would lead to different outcomes. Nonetheless, we believe that our definition of "meaningful" as denoting groups that reflect deep and important distinctions between people (it tells "what kind of person you are") accords with many everyday uses and can be readily mapped onto real-world groups such as religion and ethnicity. Other strategies could include directly stating that the groups are deep and meaningful, explicitly listing the (deep and meaningful or shallow and meaningless) properties on which group assignment is based or employing the type of functional use paradigms developed by Bigler and colleagues (e.g., Bigler et al., 1997).

It is worth noting that in Study 2 the level of biases in the maximal condition also appeared to be weaker than that in the maximal condition in Study 1. Given this, we directly compared the two common measures, preference and similarity, across the two maximal conditions in our studies. We did not find evidence of reduction on the preference measure ($p = .10$). However, there was a significant decline in bias from Study 1 to Study 2 on the similarity measure ($p < .001$), possibly due to our modification of that measure (to make questions about specific shared preferences rather than general

similarity). Although it is unclear why this modification led to reduced biases from Study 1 to Study 2, the fact that we still found a significant conditional difference (maximal vs. minimal conditions) on this measure in Study 2 provides further evidence that the manipulation of meaningfulness was effective. More generally, we should note that the level of meaningfulness attributed to the groups also appeared to decline in both the maximal and minimal conditions from Study 1 to Study 2. Although it remains speculative, one possibility is that the training of randomness led children in both conditions to perceive the groups as less meaningful overall.

Our results are somewhat at odds with a recent study (Rhodes et al., 2017) that also employed novel groups and manipulated something like essentialism before measuring ingroup bias on attitudes and resource allocation measures. Whereas those authors reported effects of the manipulation on resource allocation but not on attitudes, we observed no effects in Study 1 and the opposite pattern of results in Study 2. However, many aspects of these studies differed. For example, those authors manipulated essentialism via generic language rather than via direct description, as we employed here. It is possible that these two approaches have different consequences and that in particular generic language has a larger effect on the perception of discrete category boundaries, as argued by Rhodes et al. (2017). A more comprehensive comparison of different strategies for manipulating meaningfulness and/or essentialism would be a useful avenue for future work.

It is important to acknowledge that we needed to diverge from our preregistered analysis plan in Study 2 because our meaningfulness measures did not hang together into a coherent composite. We should note that these measures are commonly employed in the study of essentialism with children, although it remains unclear and debated in the field whether different strands of essentialist beliefs such as nativism (e.g., as measured by the switched-at-birth task), inductive potential (e.g., as measured by the generalization measure), and boundary intensification (e.g., as indicated by the follow-up questions of switched-at-birth) actually constitute a single construct. Similar to the current studies, other past work has also found that individual measures do not hang together (Gelman et al., 2010). The lack of coherence across multiple studies could serve as a warning for the frequent use of composite scores when measuring essentialism. Importantly, instead of measuring a single construct, these measures might each capture an empirically distinct aspect of essentialism (Rhodes & Mandalaywala, 2017), raising the possibility that each might have different relations with intergroup bias. This empirical observation dovetails with a conceptual point by Gelman (2004), namely that we do not yet know whether there is a single essentialist "instance" or whether essentialism is a mixture of different sets of tendencies such as nativism, inductive potential, boundary salience, and the like. Especially for young children, these beliefs appear to be even more dissociable than had previously been assumed (Gelman, Heyman, & Legare, 2007). Because in the current studies we manipulated one part of essentialism, namely the belief that categories reflect deep and stable properties as opposed to shallow and superficial properties of category members, it is unclear how other components of essentialist belief do or do not relate to one another or to the magnitude of ingroup bias. Future studies on essentialism should investigate the nature of each measure to explore whether they are tapping different aspects of essentialism or more general meaningfulness.

In closing, the current research demonstrates that the emergence of ingroup bias does not depend on the presence of rich and deep meanings associated with the groups; instead, mere social categorization following explicitly random assignment is sufficient to produce strong ingroup favoritism. In part, this is because children imbue even these randomly imposed social identities with a surprising degree of meaning. But perhaps even more striking, directly manipulating the meaningfulness of the groups does not have an effect on the degree of ingroup bias except when the experimenter actively engages in behavior that undermines the relevance of the group (in our case by switching the children's group membership for a trivial reason). These findings have critical relevance for the broader project of understanding the early emergence of prejudice and discrimination and bolster the contention that mere self-categorization into a group is sufficient to induce robust ingroup biases (Dunham, 2018). Given these findings, intervention efforts might best focus on factors such as norms against discriminatory behavior and exclusion (Killen, 2007), which can lead children to override the tendencies toward prejudicial evaluation and behavior that were so starkly revealed here.

Acknowledgments

We gratefully acknowledge the support of the John Templeton Foundation (56036). We thank the children and their parents who participated in the studies; Yale Peabody Museum of Natural History and Connecticut Science Center for participating in this research; Catalina Canon and Haley Hegefeld for assistance in data collection; Frank Keil, Julian Jara-Ettinger, Ashley Jordan, and Alexander Noyes for feedback on earlier versions of the manuscript; and members of the Social Cognitive Developmental Lab at Yale University for helpful discussions.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2019.04.013.

References

Aboud, F. E. (1988). *Children and prejudice*. Oxford, UK: Basil Blackwell.
Andreychik, M. R., & Gill, M. J. (2015). Do natural kind beliefs about social groups contribute to prejudice? Distinguishing biosomatic essentialism from bio-behavioral essentialism, and both of these from entitativity. *Group Processes & Intergroup Relations, 18*, 454–474.
Banerjee, K., & Bloom, P. (2015). "Everything happens for a reason": Children's beliefs about purpose in life events. *Child Development, 86*, 503–518.
Barret, M. (2007). *Children's knowledge, beliefs and feelings about nations and national groups*. Hove, UK: Psychology Press.
Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development, 68*, 530–543.
Brewer, M. B. (2008). Depersonalized trust and ingroup cooperation. In J. I. Krueger (Ed.), *Rationality and social responsibility (Modern Pioneers in Psychological Science* (pp. 215–232). New York: Psychology Press.
Brewer, M. B., & Silver, M. (1978). Ingroup bias a function of task characteristics. *European Journal of Social Psychology, 8*, 393–400.
Chao, M. M., Hong, Y. Y., & Chiu, C. Y. (2013). Essentializing race: Its implications on racial categorization. *Journal of Personality and Social Psychology, 104*, 619–634.
Chen, J. M., & Ratliff, K. A. (2018). Psychological essentialism predicts intergroup bias. *Social Cognition, 36*, 301–323.
Diesendruck, G., & Menahem, R. (2015). Essentialism promotes children's inter-ethnic bias. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01180.
Dunham, Y. (2018). Mere membership. *Trends in Cognitive Sciences, 22*, 780–793.
Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development, 82*, 793–811.
Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*, 404–409.
Gelman, S. A., Heyman, G. D., & Legare, C. H. (2007). Developmental changes in the coherence of essentialist beliefs about psychological characteristics. *Child Development, 78*, 757–774.
Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology, 61*, 273–301.
Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition, 38*, 213–244.
Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39*, 113–127.
Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology, 41*, 87–100.
Heiphetz, L., Spelke, E. S., & Banaji, M. R. (2013). Patterns of implicit and explicit attitudes in children and adults: Tests in the domain of religion. *Journal of Experimental Psychology: General, 142*, 864–879.
Herek, G. M., & Capitanio, J. P. (1995). Black heterosexuals' attitudes toward lesbians and gay men in the United States. *Journal of Sex Research, 32*, 95–105.
Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53*, 575–604.
Hirschfeld, L. (1995). Do children have a theory of race? *Cognition, 54*, 209–252.
Juvan, E., & Dolnicar, S. (2014). The attitude–behaviour gap in sustainable tourism. *Annals of Tourism Research, 48*, 76–95.
Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology, 88*, 686–702.
Killen, M. (2007). Children's social and moral reasoning about exclusion. *Current Directions in Psychological Science, 16*, 32–36.
Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 12577–12580.
Kollmuss, A., & Agyeman, J. (2010). Mind the gap: Why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental Education Research, 8*, 239–260.
Kuzmak, S. D., & Gelman, R. (1986). Young children's understanding of random phenomena. *Child Development, 57*, 559–566.
Lane, T. (2016). Discrimination in the laboratory: A meta-analysis of economics experiments. *European Economic Review, 90*, 375–402.

Metz, K. E. (1998). Emergent understanding and attribution of randomness: Comparative analysis of the reasoning of primary grade children and undergraduates. *Cognition and Instruction, 16*, 285–365.
Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22*, 103–122.
Otten, S. (2016). The minimal group paradigm and its maximal impact in research on social categorization. *Current Opinion in Psychology, 11*, 85–89.
Perreault, S., & Bourhis, R. Y. (1999). Ethnocentrism, social identification, and discrimination. *Personality and Social Psychology Bulletin, 25*, 92–103.
Piaget, J., & Inhelder, B. (1976). The origin of the idea of chance in children (P. B. Leake, H. D. Fishbein, & L. Leake, Trans.). London: Norton.
Prentice, D. A., & Miller, D. T. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science, 16*, 202–206.
Rangel, U., & Keller, J. (2011). Essentialism goes social: Belief in social determinism as a component of psychological essentialism. *Journal of Personality and Social Psychology, 100*, 1056–1078.
Rhodes, M., Leslie, S.-J., Saunders, K., Dunham, Y., & Cimpian, A. (2017). How does social essentialism affect the development of inter-group relations? *Developmental Science, 21*, e12509.
Rhodes, M., Leslie, S.-J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 13526–13531.
Rhodes, M., & Mandalaywala, T. M. (2017). The development and developmental consequences of social essentialism. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*, e1437.
Richter, N., Over, H., & Dunham, Y. (2016). The effects of minimal group membership on young preschoolers' social preferences, estimates of similarity, and behavioral attribution. *Collabra, 2*(1), 1–8.
Schütz, H., & Six, B. (1996). How strong is the relationship between prejudice and discrimination? A meta-analytic answer. *International Journal of Intercultural Relations, 20*, 441–462.
Sparks, E., Schinkel, M. G., & Moore, C. (2017). Affiliation affects generosity in young children: The roles of minimal group membership and shared interests. *Journal of Experimental Child Psychology, 159*, 242–262.
Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*(5), 96–102.
Tajfel, H., & Billig, M. (1974). Familiarity and categorization in intergroup behavior. *Journal of Experimental Social Psychology, 10*, 159–170.
Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*, 149–178.
Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development, 67*, 1555–1571.
Waxman, S., Medin, D., & Ross, N. (2007). Folk-biological reasoning from a cross-cultural developmental perspective: Early essentialist notions are shaped by cultural beliefs. *Developmental Psychology, 43*, 294–308.
Whitley, B. E. (1990). The relationship of heterosexuals' attributions for the causes of homosexuality to attitudes toward lesbians and gay men. *Personality and Social Psychology Bulletin, 16*, 369–377.
Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology, 94*, 1033–1047.
Yamagishi, T., & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly, 63*, 116–132.
Yee, M., & Brown, R. (1994). The development of gender differentiation in young children. *British Journal of Social Psychology, 33*, 183–196.
Yzerbyt, V. Y., Rogier, A., & Fiske, S. T. (1998). Group entitativity and social attribution: On translating situational constraints into stereotypes. *Personality and Social Psychology Bulletin, 24*, 1089–1103.
Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2), 1–27.