Cognitive Psychology 129 (2021) 101408

Contents lists available at ScienceDirect

Cognitive Psychology

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

Dichotomous thinking about social groups: Learning about one group can activate opposite beliefs about another group

Hannah J. Kramer [a,*], Deborah Goldfarb [a,b], Sarah M. Tashjian [a,c,d], Kristin Hansen Lagattuta [a]

[a] University of California, Davis, United States
[b] Florida International University, United States
[c] University of California, Los Angeles, United States
[d] California Institute of Technology, United States

ARTICLE INFO

Keywords:
Social cognition
Categories
Generic language
Development
Heuristics

ABSTRACT

Across three studies (N = 607), we examined people's use of a *dichotomizing heuristic*—the inference that characteristics belonging to one group do not apply to another group—when making judgments about novel social groups. Participants learned information about one group (e.g., "Zuttles like apples"), and then made inferences about another group (e.g., "Do Twiggums like apples or hate apples?"). Study 1 acted as a proof of concept: Eight-year-olds and adults (but not 5-year-olds) assumed that the two groups would have opposite characteristics. Learning about the group as a generic whole versus as specific individuals boosted the use of the heuristic. Study 2 and Study 3 (sample sizes, methods, and analyses pre-registered), examined whether the presence or absence of several factors affected the activation and scope of the dichotomizing heuristic in adults. Whereas learning about or treating the groups as separate was necessary for activating dichotomous thinking, intergroup conflict and featuring only two (versus many) groups was not required. Moreover, the heuristic occurred when participants made both binary and scaled decisions. Once triggered, adults applied this cognitive shortcut widely—not only to benign (e.g., liking apples) and novel characteristics (e.g., liking modies), but also to evaluative traits signaling the morals or virtues of a social group (e.g., meanness or intelligence). Adults did not, however, extend the heuristic to the edges of improbability: They failed to dichotomize when doing so would attribute highly unusual preferences (e.g., disliking having fun). Taken together, these studies indicate the presence of a dichotomizing heuristic with broad implications for how people make social group inferences.

1. Introduction

People often expect social groups to differ—for example, they believe that members from one group will hold values, beliefs, preferences, skills, and traits that are distinct from members of another group. Such inferences about between-group differences, whether accurate or not, have been shaped over time through systemic influences like social hierarchies, stereotypes, and prejudice, as

* Corresponding author at: Department of Psychology and Center for Mind and Brain, University of California, 1 Shields Avenue, Davis, CA 95616, United States.
E-mail address: hjkramer@ucdavis.edu (H.J. Kramer).

https://doi.org/10.1016/j.cogpsych.2021.101408
Received 4 August 2020; Received in revised form 19 April 2021; Accepted 22 June 2021
Available online 27 July 2021

well as by directly experienced or observed social interactions and information from others (Richter et al., 2016; Schwarz & Bless, 2007; Tajfel, 1982). In the current study, we examined the inferences individuals make when first learning about a novel social group. In the absence of complex historical and phenomenological evidence, will people assume that different social groups hold opposite characteristics?

Across several domains, people often use heuristics, or cognitive shortcuts, to make judgments because they are quicker than more thorough, careful, and deliberate thinking (Kahneman, 2011; Sherman & Corty, 1984; Sunstein, 2005; Tversky & Kahneman, 1973). Heuristics play a vital role in reducing uncertainty, decreasing cognitive load, and increasing processing efficiency. They can, however, also lead to biased assumptions about social groups (e.g., when forming an impression of a new person from a known social group, people often use stereotypes of that social group as an anchor; when making adjustments away from this anchor they do not go far enough; FeldmanHall & Shenhav, 2019). Here, we investigate whether people apply a *dichotomizing heuristic*—what is true of one group is not true of another group—when forming impressions of novel social groups for which they have limited information.[1] To test this potential heuristic, participants learned about the preferences, abilities, and traits of one group (*characterized group*). Participants then made inferences about an *uncharacterized group's* preferences, abilities, and traits (no information provided). Across three studies, we examined what triggers the dichotomizing heuristic as well as its scope.

Humans assume that individuals who are members of the same group share common features (Gelman, 2003). By three to four years of age, children use category labels to make inferences about the biological and psychological properties of individuals; for example, they infer that two creatures called "fish" will both breathe under water or that two kids categorized as "shy" will both like the same game (Gelman & Markman, 1986; Heyman & Gelman, 2000a, 2000b). Adults also give particular weight to category labels as compared with other features that may signal group membership when drawing inferences (Yamauchi & Markman, 2000). Moreover, preschoolers, school-aged children, and adults tend to extend information learned about the characteristics of one category member to all category members. For example, they predict that individuals from the same social group (e.g., gender or race) will share similar traits, preferences, beliefs, and behaviors (Diesendruck et al., 2015; Kalish, 2012; Krueger & Clement, 1994; Lagattuta & Kramer, 2021; McGarty & Penny, 1988; McGarty & Turner, 1992; Shutts et al., 2013), with adults overestimating stereotypic traits among group members (e.g., all women are caring; Allport, 1954; Devine, 1989; Park & Rothbart, 1982). Thus, when individuals are members of the same category or group, people across a wide age range assume homogeneity: They infer that learning about the characteristics of one member provides valuable evidence about all members and that category labels can be informative about group characteristics.

Categories, however, not only prime beliefs about within-group cohesiveness, but can also exaggerate perceptions of between-group differences, sometimes called *category accentuation* (Eiser, 1971; Tajfel & Wilkes, 1963). For example, 4-year-olds who first sorted faces into two separate categories (mean vs. nice) versus on a scaled continuum (very mean to very nice) more often assumed greater differences in behaviors, preferences, and intentions of meaner versus nicer looking individuals (Master et al., 2012). Moreover, adults evaluate a group with neutral traits more favorably when they simultaneously learn about another group with negative traits compared to when they simultaneously learn about another group with positive traits (Krueger & Rothbart, 1990). Intriguingly, the presence of real-world group differences is not necessary for people to assume dissimilarities; there can be *illusory correlations* (Hamilton & Gifford, 1976). For example, children and adults tend to associate majority groups with more frequent behaviors (e.g., positive behaviors) and minority groups with less frequent behaviors (e.g., negative behaviors), even when the proportions of these actions are equivalent across groups (Johnston & Jacobs, 2003; Lawson & Bower, 2014; Sherman et al., 2009). Thus, when humans learn about multiple categories or groups, they tend to attribute greater between-group differences than warranted by observable or verifiable evidence.

Here, we extend inquiry into people's beliefs about social groups as categories. As reviewed, the typical paradigm has involved (1) telling participants characteristics of one group member and seeing if they extend those characteristics to all group members, or (2) providing information about two separate groups and testing if participants inflate between-group differences. But, what happens if participants only learn the characteristics of *one* group—what assumptions do they make about members of a separate, uncharacterized group? In this uncertain context where no details are provided about the uncharacterized group, do people default to a dichotomizing heuristic—assume that what is true of one group must not be true of the other group? That is, with no descriptions of the uncharacterized group to draw from, will people judge that the two groups have opposite characteristics (evidence of a dichotomizing heuristic), assume the two groups share the same characteristics, or just guess randomly?

Study 1 served as a proof of concept: Our goal was to document whether people assume that they have acquired information about two groups—a characterized group and an uncharacterized group—when they have only learned about one. In this initial study, we further tested whether the language used to discuss the characterized group influences application of the dichotomizing heuristic. That is, we assessed whether use of the dichotomizing heuristic is more pronounced when information about the social category is presented generically (i.e., describing the characterized group as a whole; e.g., "Zuttles like pears") rather than specifically (i.e., describing individuals within a group; e.g., "This Zuttle, X, likes pears"). Generic language encourages children and adults to assume greater homogeneity within categories (Chambers et al., 2008; Gelman et al., 2002) and to create stricter boundaries between categories (Gelman et al., 2010; Goldfarb et al., 2017). Thus, we tested whether participants who learn about the characterized group as a generic whole (versus as specific individuals) more often exhibit the dichotomizing heuristic (assume that the uncharacterized group possesses opposite traits).

A further goal of Study 1 was to identify age-related differences in the application of the dichotomizing heuristic. Extant research

[1] In labeling this a heuristic, we are not arguing that this inference is de facto irrational or rational, only that it is a rule-of-thumb approach that enables decisions under uncertainty with minimal mental effort.

suggests that this cognitive shortcut, if present, would develop with age as individuals continue to explore and form models of the social world. Preschoolers, compared to older children and adults, require more generic input before they generalize group features across all category members (Cimpian & Scott, 2012), and they form less rigid category boundaries between separate social groups (Shutts, 2015). Reliance on heuristics also strengthens not only from early to middle childhood but continues to increase from childhood to adulthood (Jacobs & Potenza, 1991; Lagattuta & Kramer, 2021; Lagattuta & Sayfan, 2013; Lagattuta et al., 2018; Reyna & Ellis, 1994). Based on this combined evidence, we reasoned that young children would less often use a dichotomizing heuristic to make inferences about an uncharacterized group. Therefore, we tested whether older participants would more often endorse inter-group dichotomies than younger children.

Recent evidence to support this new research direction in social group reasoning comes from Moty and Rhodes (2021).[2] Children as young as 4.5 years, and increasing with age (i.e., between 4.5 and 7 years and between 7 and adulthood), judged that characteristics of one group (e.g., skill at baking) do not apply to a different social group. Moreover, learning about groups in generic (e.g., Zarpies are good at baking pizza) versus specific language (e.g., This Zarpie is good at baking pizza) magnified this effect. In these studies, however, the experimenter explicitly told participants that the two groups—Zarpies and Gorps—were "two kinds of people" and they pictured them wearing distinctive clothes (e.g., all green for Zarpies; all yellow for Gorps). Starting from the salient premise that the two groups had different origins and different clothing preferences could have biased participants to expect dissimilarity across all characteristics, facilitating dichotomizing at an early age.

In Studies 1 and 2, we provided a more stringent test by examining whether participants also exhibit the dichotomizing heuristic when the two social groups share a common biological and social origin (e.g., similar to how non-indigenous American colonists separated from the British) and cannot be distinguished visually; thus, making it equivalently plausible to assume either between-group similarities or differences. In Study 2, we also manipulated exposure to a categorization (sorting) task to test whether thinking about the groups as separate entities activates the dichotomizing heuristic. In that study, we further investigated whether intergroup conflict or beneficence shaped the dichotomizing heuristic. In Study 3, we removed the shared biological and social origin of Zuttles and Twiggums, but we provided no distinguishing information beyond the different category labels. In this study, we again tested whether experience separating individuals into different groups (the categorization task) is necessary for activating the heuristic if groups were initially described as distinct. In addition, we measured whether the dichotomizing heuristic was bolstered by learning about the uncharacterized group in the presence of only one other group (i.e., setting up a binary contrast) versus in the presence of many groups (i.e., no clear contrast between one group and another). We also changed our dependent variable from binary to scaled to test the robustness of the effect. Moreover, we assessed the scope and potential limits of the dichotomizing heuristic by analyzing whether participants apply it to benign (e.g., liking apples), novel (e.g., like modies), evaluative (e.g., being smart), and universal characteristics (e.g., disliking getting hurt).

2. Study 1

Study 1 builds upon the novel-group paradigm utilized by Goldfarb et al. (2017; for further examples of novel-group methods, see Killen, 2007; Roberts et al., 2018). In Goldfarb et al. (2017), 5-year-olds, 8-year-olds, and adults learned about a fictitious social group of human-like creatures, Twiggums, who lived in a valley. Some Twiggums left the valley to live in the mountains. There, they formed a new group and renamed themselves Zuttles, a social group whose members still looked indistinguishable from Twiggums. Next, participants learned about Zuttle preferences and abilities either in generic language (e.g., "Zuttles like pears") or in specific language (e.g., "This Zuttle, Dax, likes pears"). Participants then completed a *social categorization task* where they saw a creature paired with three characteristics. Children and adults decided whether that individual was a Zuttle or a Twiggum by making characteristic-to-category judgments. They received no feedback about the correctness of their choice because there was no objective truth given that participants had no information about Twiggums.

In one version of the social categorization task (neutral stance), participants sorted creatures without consequence: Nothing happened to those identified as Zuttles. In the other condition (negative stance), a law banned creatures identified as Zuttles from the valley, forced them to be jailed for 100 days, and then required their deportation. While Goldfarb et al. (2017) presented the findings from the categorization task, Study 1 of the current manuscript focuses on a separate measure that participants completed after the social categorization task. This new *uncharacterized group task* asked children and adults to judge whether they thought Twiggums (uncharacterized group) differed from or aligned with Zuttles (characterized group) in their preferences and abilities; that is, they made category-to-characteristic judgments. Of central interest was whether participants would use a dichotomizing heuristic—what is true of one group is not true of the other group—when making inferences about the uncharacterized group. Our paradigm differed from Moty and Rhodes (2021) in that we neither described Zuttles and Twiggums as "two kinds of people" nor showed them as differing in appearance. Rather, to lessen the demand to assume group differences, Zuttles and Twiggums (1) had a common biological and social origin; and (2) looked identical. In addition to these shared features, we also included factors which could suggest that the groups might differ, including elective migration and punitive laws: (1) some Twiggums decided to move to the mountains and call themselves Zuttles; and (2) the Twiggums then made a law banning Zuttles from the valley.

[2] For transparency, although we include Moty and Rhodes (2021) in the Introduction, the methods and hypotheses for the studies reported here were not created to extend that study. We completed all analyses for Study 1 prior to Moty and Rhodes' preregistration. Pilot Study 2 and Study 2 were preregistered and completed prior to Moty and Rhodes (2021) being accepted for publication. Moty and Rhodes (2021) also do not use the term "dichotomizing heuristic" when discussing their findings.

2.1. Method

2.1.1. Participants

Participants included 181 children and adults divided into three age groups: 66 5-year-olds ($M = 5.57$ years, $SD = 0.76$ years, range = 4.20 years to 6.99 years; 34 male, 32 female), 62 8-year-olds ($M = 8.40$ years, $SD = 0.84$ years, range = 7.06 years to 10.48 years; 29 male, 33 female), and 53 adults ($M = 22.61$ years, $SD = 7.47$ years, range = 18.17 years to 60.15 years; 24 male, 29 female). The sample was 14% Asian, 2% Black or African American, 10% Hispanic/Latino, 58% White, and 16% "Other" race or ethnicity.[3] Included participants were fluent in comprehending and speaking English and were typically developing (parent-reported for child participants, self-reported for adults). Children were recruited from a list of previous participants, referral, recruitment at local farmers' markets and schools, listserv emails, and fliers. Adults were recruited from an undergraduate psychology database. Data were collected from October 2013 to October 2014. Participants were the same as those included in Goldfarb et al. (2017) minus nine adults who did not complete the uncharacterized group task. This study was approved by the Internal Review Board at the University of California, Davis: Protocol Number: 448441.

2.1.2. Measures and procedures

This study followed a 3 (age: 5-year-olds, 8-year-olds, adults) × 2 (language: generic, specific) design. Age and language were between-subjects factors.

Narrative and Social Categorization Task. Participants learned from a pre-recorded video (i.e., pictures on a computer screen accompanied by audio) that a group of human-like creatures, Twiggums, lived in a valley. Some of the Twiggums left the valley to live in the mountains. Once there, this group of Twiggums re-named themselves Zuttles. Participants then learned about Zuttles' preferences and abilities in either generic (e.g., "Zuttles like pears.") or specific language (e.g., "This Zuttle, Dax, likes pears."). All featured Zuttle characteristics were benign (i.e., did not indicate any deviance or threat of harm) and non-distinctive (i.e., could be widely shared, even with the participants themselves).

Next, participants completed the social categorization task that required them to determine whether a series of creatures were Zuttles (characterized group) or Twiggums (uncharacterized group) based on three pieces of information about their characteristics (only a subset matched what they learned about Zuttles). In the *negative stance* context, participants were told that the Twiggums had created a law forbidding Zuttles from living in the valley. Thus, those identified as Zuttles would be sent to jail for 100 days and then deported. In the *neutral stance* context, there were no consequences for creatures identified as Zuttles. Context order was counterbalanced. To control for differences in responding due to variation in working memory skills, all participants were given a memory-cue card that displayed the characteristics they had learned about Zuttles. Following research with young children (e.g., Lagattuta, 2008; Lara et al., 2019), we used crossed-out images to signify that a Zuttle/Zuttles hates something or is/are bad at something. For more details on the social categorization task, see Goldfarb et al. (2017).

Uncharacterized Group Task. After the social categorization task, children and adults predicted Twiggums' (uncharacterized group) preferences and abilities. We introduced the task as, "While we talked a lot about Zuttles, I am going to ask you some questions about what you think Twiggums might be like. Just make your best guess. There is no right answer. This is based on what you think." We then asked participants about eight of the characteristics (e.g., "Do you think Twiggums like pears or hate pears?"). The memory-cue sheet remained visible throughout.

We coded each response as *dichotomized* (1; participant said that Twiggums had *the opposite* preference or ability from Zuttles) or *aligned* (0; participant said that Twiggums had *the same* preference or ability as Zuttles). The eight characteristics showed good internal consistency (Cronbach's alpha = 0.87, 95% confidence interval [CI] [0.84, 0.90]; 5-year-olds: Cronbach's alpha = 0.80, 95% CI [0.73, 0.87]; 8-year-olds: Cronbach's alpha = 0.89, 95% CI [0.85, 0.93]; adults: Cronbach's alpha = 0.90, 95% CI [0.86, 0.94]). We calculated the proportion of trials that participants dichotomized Twiggums and Zuttles for analyses (see *Table S1* for individual trial level data). Higher scores reflect a greater frequency of inferring that the uncharacterized group had opposite traits to the characterized group. We removed one 8-year-old from analyses because she both dichotomized and aligned the groups on every trial (e.g., stating that the uncharacterized group both liked pears and hated pears). Although this is a potentially logical response, we excluded her because we could not represent her decisions within our coding scheme. One 5-year-old also followed this pattern for only one characteristic (i.e., stating that the uncharacterized group was both good and bad at playing drums) so we calculated his proportion score out of seven rather than eight trials ($N = 180$).

General Procedure. Participants were tested individually in a quiet room as part of a larger study on social reasoning.[4] Participants completed the social categorization task (as reported in Goldfarb et al., 2017) prior to the uncharacterized group task. Children received $10.00 and adults were given course credit for the approximately one-hour session.

2.2. Results and Discussion

We conducted analyses in RStudio (R Version 3.6.1; R Core Team, 2017). We tested the effects of generic language and age on the proportion of trials that participants dichotomized the two groups (Table 1, Figs. 1 and 2). Language and age were between-subjects

[3] Race and ethnicity data were missing from one 8-year-old participant.

[4] Participants also responded to questions aimed at addressing their essentialist beliefs, working memory, and fairness and procedural justice judgments. These measures will be analyzed in separate manuscripts.

factors. A 2 (language: specific, generic) × 3 (age: 5-, 8-year-olds, adults) univariate analysis of variance (ANOVA) resulted in main effects for language, $F(1, 174) = 7.65$, $p = .006$, $\eta_p^2 = .04$, and age, $F(2, 174) = 8.90$, $p < .001$, $\eta_p^2 = .09$. The Language × Age interaction was not significant, $F(2, 174) = 1.34$, $p = .265$, $\eta_p^2 = .02$.

As anticipated, participants in the generic-language condition dichotomized Zuttles and Twiggums more frequently than those in the specific-language condition ($p = .006$). Still, participants in *both* language conditions judged the two groups to have opposite preferences and skills more often than would be expected by chance (0.50; generic: $t[82] = 6.66$, $p < .001$, $d = 0.73$; specific: $t[96] = 3.70$, $p < .001$, $d = 0.38$). Consistent with our predictions, Tukey's HSD comparisons showed that adults dichotomized the groups more frequently than did 5-year-olds ($p < .001$). Eight-year-olds did not differ from either 5-year-olds ($p = .065$) or adults ($p = .111$). Whereas adults ($t[52] = 7.18$, $p < .001$, $d = 0.99$) and 8-year-olds ($t[60] = 4.43$, $p < .001$, $d = 0.57$) dichotomized the two groups more frequently than would be expected by chance, 5-year-olds responded at chance ($t[65] = 1.69$, $p = .096$, $d = 0.21$).

Our results indicate that by eight years of age, individuals exhibit a dichotomizing heuristic when forming impressions of new social groups. Older children and adults used what they knew about one group to infer that the opposite was true of the other group. Although participants made this inference in both language conditions, learning about the characterized group as a whole (generically versus as specific individuals) boosted the frequency that participants expected dichotomies between social groups. These findings are consistent with Moty and Rhodes (2021) in that generic language magnified children's and adults' expectation of social dichotomies, and older children and adults assume dichotomies more than younger children. Still, it is notable that although the dichotomizing heuristic was strong at the level of the individual (e.g., 81% of 8-year-olds in the generic language condition dichotomized the groups on > 50% of trials), a minority of participants consistently assumed that the two groups would have identical preferences and abilities (see *Table S1*). This provides assurance that aligning the groups is also possible within our paradigm.

3. Study 2

Study 1 confirmed the presence of the dichotomizing heuristic, but questions remain as to the circumstances under which it emerges. Prior research has demonstrated that group divisions must serve a functional purpose before people will use these categories in meaningful ways (Bigler, 1995; Bigler et al., 1997). Perhaps our social categorization task, which preceded the uncharacterized group task, acted as a catalyst for the heuristic because it encouraged participants to imagine that Twiggums and Zuttles were "two different kinds"—the explicit starting premise in Moty and Rhodes (2021). In a pilot study for Study 2 (see Online Supplementary Information), we found preliminary evidence that removing the categorization task reduced the dichotomizing heuristic. As a more rigorous test in Study 2, we manipulated (between subjects) whether participants did or did not complete the social categorization task prior to making inferences about the uncharacterized group. This enabled us to test whether repetition separating individuals into two social groups by making characteristic-to-group inferences (i.e., deciding whether each individual is a Zuttle or a Twiggum based on varying evidence about their preferences and abilities) activates the dichotomizing heuristic.

In Study 1, all participants learned that Twiggums criminalized Zuttles living in the Twiggum Valley (if found, Zuttles would be jailed for 100 days and deported). This evidence of between-group conflict may have boosted the use of the dichotomizing heuristic. Enforced physical and legal boundaries could have signaled to participants that there were perceived incompatibilities between Zuttles and Twiggums. Indeed, intergroup conflict can magnify negative views of outgroup members as well as beliefs about differences (Allport, 1954; Dovidio et al., 2005; Esses et al., 2005). In Study 2, we tested whether intergroup conflict affected adults' tendency to use the dichotomizing heuristic. We manipulated (between subjects) whether the two groups held a negative or positive stance towards each other.

Study 1 focused on categorizing human-like, imaginary creatures. Although this fictional context is a common method in novel group paradigms (see Killen, 2007; Roberts et al., 2017), we sought to test whether participants would apply the dichotomizing heuristic to human social groups. Thus, Study 2 involved novel human groups. As in Study 1, we explicitly stated that the two groups shared the same biological and social origin.

Finally, we examined the scope of the dichotomizing heuristic. Participants in Study 1 may have interpreted the uncharacterized group task as assessing what they "learned" in the social categorization task. If so, then their willingness to expect dichotomies should be confined only to qualities that were featured during the social categorization task. Perhaps, however, the dichotomizing heuristic has even greater reach. In Study 2, we measured whether adults would apply this inference to novel preferences and skills (e.g., if told that Zuttles like "modies" do they infer that Twiggums hate "modies?"). We also examined whether they would even use this heuristic when reasoning about evaluative traits (e.g., if told that Zuttles are "smart" do they infer that Twiggums are "not smart?"). Indeed, the potential real-world negative consequences of the dichotomizing heuristic would be especially severe if people use it when inferring evaluative traits that signal morals or virtues of an uncharacterized group.

Given these multiple changes needed to systematically unpack the conditions that give rise to the dichotomizing heuristic as well as its potential scope and boundaries, we shifted to an exclusive focus on adults for Studies 2 and 3. We return to the necessity of

Table 1
Study 1: Means (SDs) of Proportion of Trials Dichotomized by Age Group and Language Condition.

	5-year-olds	8-year-olds	Adults	Across Age
Specific Language	0.55 (0.29)	0.60 (0.34)	0.73 (0.36)	0.63 (0.33)
Generic Language	0.58 (0.36)	0.79 (0.33)	0.90 (0.15)	0.74 (0.33)
Across Language	0.57 (0.32)	0.70 (0.35)	0.80 (0.30)	0.68 (0.34)

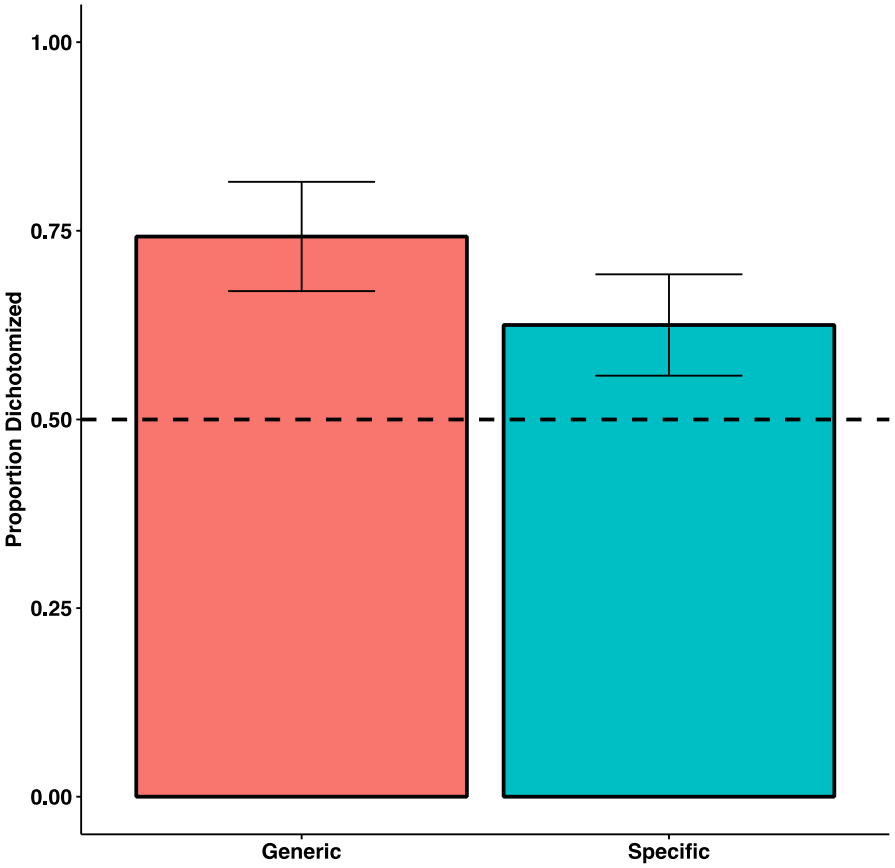

Fig. 1. Study 1: Proportion of trials that participants dichotomized the Zuttles (characterized group) and Twiggums (uncharacterized group) by language condition. Note. Error bars are 95% confidence intervals. Dashed line signifies chance.

understanding the developmental trajectory of the dichotomizing heuristic in the General Discussion.

3.1. Method

3.1.1. Participants

Participants were 214 undergraduate students ($M = 21.13$ years, $SD = 2.43$ years, 105, males, 109 females). Included participants were fluent in comprehending and speaking English and were typically developing (via self-report). The sample was 55% Asian, 2% Black or African American, 17% Hispanic/Latino, 14% White, 8% Multiracial or Multiethnic, and 4% other races or ethnicities. We set the sample size for each cell at 53 participants per cell ($N = 212$) based on Simonsohn's (2015) recommendation of multiplying the original sample ($n = 21$ adults in the generic language condition of Study 1) by 2.5. This sample size (53 participants per condition) gave us $>$ 99% power to detect a $d = 2.76$ (the effect size of the dichotomizing heuristic in the adults in the generic language condition in Study 1) for a one-sample *t*-test (the key test of whether participants in each condition exhibit the dichotomizing heuristic); for the same analysis, it also gives us 80% power to detect a $d > 0.39$. We stopped data collection the day this goal was met. We included three attention checks that appeared randomly during the dichotomizing phases of the study ([1] Please select the number 4; options: 4, 8; [2] Please select the number 3; options: 1, 3; [3] Please select the number 7; options: 7, 4). We excluded four participants who were not 100% accurate. The final sample included 210 participants. Participants received course credit. Data were collected from June 2019 through October 2019. This study was approved by the Internal Review Board at the University of California, Davis: Protocol Number: 1031991. The sample size (including exclusions), method, and planned analyses were pre-registered on Open Science Framework (https://osf.io/nw5jx/?view_only=811245f5693a4b3f8f3112b5a7dec7d3).

3.1.2. Measures and Procedure

This study followed a 2 (categorization: categorization, no-categorization) $\times$ 2 (stance: negative, positive) $\times$ 3 (type: initial, novel, evaluative) design. Categorization was a between-subjects factor that differed in whether participants did (categorization) or did not (no-categorization) complete a social categorization task before the dichotomizing measure. Stance was a between-subjects factor that differed in whether there was conflict (negative) or goodwill (positive) between the two groups. Type was a within-subjects factor where participants responded to three phases of the dichotomizing measure (see details below).

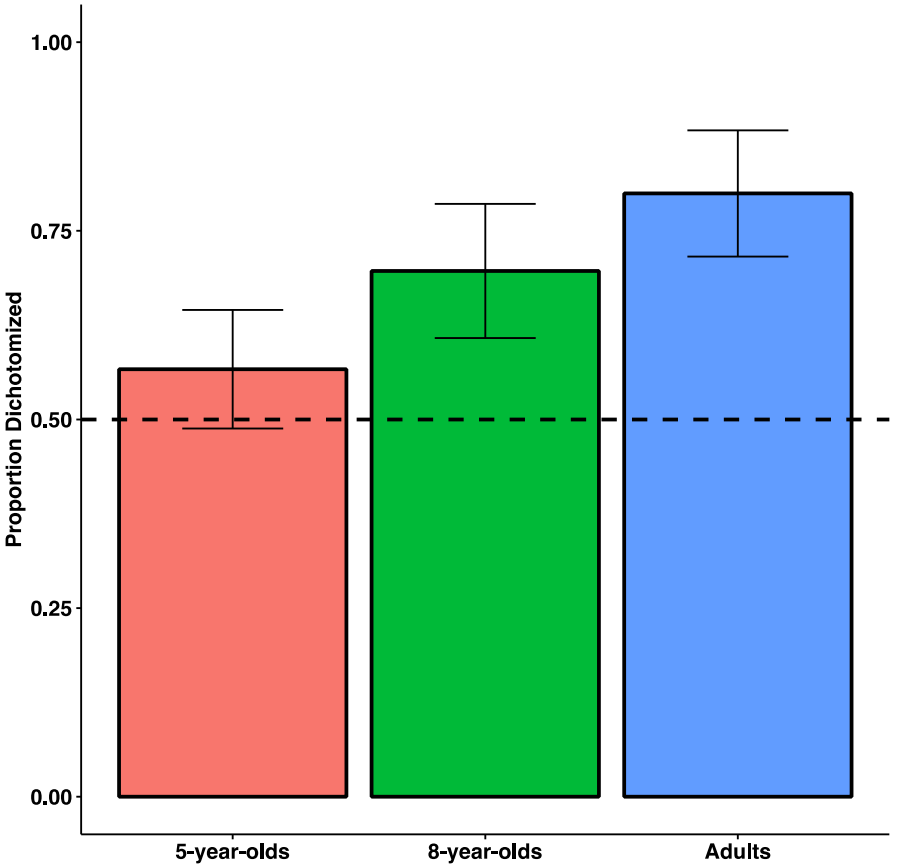

Fig. 2. Study 1: Proportion of trials that participants dichotomized the Zuttles (characterized group) and Twiggums (uncharacterized group) by age group. Note. Error bars are 95% confidence intervals. Dashed line signifies chance.

Narrative and Social Categorization Task. All participants were tested individually in a quiet room at a university laboratory. After completing informed consent and basic demographic information, participants responded to a series of tasks via Qualtrics Survey Software (Qualtrics, Provo, UT). Participants learned about a group of humans called Twiggums who lived on an island. They were also told that some Twiggums left Twiggum Island to form a new group called Zuttles who lived on Zuttle Island. Participants then learned about Zuttle characteristics in generic language (i.e., their preferences and abilities; e.g., Zuttles like pears; Zuttles are bad at playing soccer). These were the same characteristics used in Study 1. All participants were given a piece of paper (to put next to the computer) that listed all of the learned Zuttle characteristics (we used four different blocked orders for this memory card).

Participants in the two categorization conditions then completed the categorization task. As with Study 1, they were told three qualities about each of a series of individuals, and they had to judge whether each person was a Zuttle or a Twiggum. Participants in the Categorization/Negative (Cat/Neg) condition were told that there was a law prohibiting Zuttles from living on Twiggum Island: They sorted individuals to determine who would be deported back to Zuttle Island. In contrast, participants in the Categorization/Positive (Cat/Pos) condition learned that there was a law allowing Zuttles to live on Twiggum Island: They sorted individuals to determine who would be invited to a "welcome" party. This social categorization task, including randomization by evidence type, was identical to Study 1 except that the information was displayed in a written format rather than using a narrated story with pictures. Participants in the no-categorization conditions (No-Categorization/Negative [NoCat/Neg]; No-Categorization/Positive [NoCat/Pos]) read the same information as those in the categorization conditions (i.e., the 12 Zuttle characteristics, the law either prohibiting or welcoming Zuttles to live on Twiggum Island), but they did not complete the social categorization (sorting) task.

Uncharacterized Group Task. All participants then completed three questioning sets where they inferred the traits of the Twiggums. Questions followed a similar format to that of Study 1: Participants were asked to judge whether Twiggums had various preferences, abilities, and traits. These questioning sets followed the same fixed order for all participants.

First, the *initial set* included the 12 characteristics that were featured in the Zuttle familiarization phase and appeared in the categorization task. Thus, this set is a close replication of the uncharacterized group judgment task used in Study 1 (e.g., "Do you think that Twiggums like pears or hate pears?" Answer choices: Twiggums like pears; Twiggums hate pears). Within this questioning set, responses showed excellent internal consistency across conditions (Cronbach's alpha = 0.98, 95% CI [0.97, 0.98]). For analyses, we created a proportion of initial trials dichotomized (see *Table S2* and *Table S5* for individual trial level data).

Second, the *novel set* included 12 preferences and abilities described using novel words presented for the first time (e.g., liking

"blickets"). The use of novel words ensured that participants could not rely on what they had learned during the initial Zuttle familiarization or had applied in the categorization task. Participants responded to four blocks of three questions (presented in random order; e.g., "Here are some foods that you have never heard about. Zuttles like blickets, modies, and kifs. So, what do you think about Twiggums? Do you think Twiggums like blickets or hate blickets?" Answer choices: Twiggums like blickets, Twiggums hate blickets). Within this questioning set, responses showed excellent internal consistency across conditions (Cronbach's alpha = 0.97, 95% CI [0.96, 0.98]). For analyses, we created a proportion of novel trials dichotomized (see *Table S3* and *Table S5* for individual trial level data).

Finally, in the *evaluative set* participants were told 12 evaluative traits about Zuttles and made inferences about Twiggums. Participants learned about 4 positive traits (smart, creative, honest, funny), 4 equivocal traits (shy, confident, superstitious, liberal), and 4 negative traits (mean, moody, messy, lazy). For each trait, participants were randomly assigned to learn that Zuttles either did (two of each valence: positive, equivocal, negative) or did not (two of each valence) possess the trait (e.g., "Zuttles are smart. So, what do you think about Twiggums?" Answers: Twiggums are smart; Twiggums are not smart). Traits were presented in randomized order. Within this questioning set, responses showed excellent internal consistency across conditions (Cronbach's alpha = 0.93, 95% CI [0.92, 0.94]).[5] For analyses, we created a proportion of evaluative trials dichotomized (see *Table S4* and *Table S5* for individual trial level data).

3.2. Results and Discussion

We examined whether the proportion of trials that participants dichotomized the characterized and uncharacterized groups differed by categorization condition (categorization vs. no categorization task), stance (negative vs. positive intergroup relations), and judgment type (initial vs. novel vs. evaluative traits). Analyses were conducted in RStudio (version 3.6.1; R Core Team, 2019). See Online Supplemental Materials for relations between the social categorization task and the uncharacterized group task.[6]

We ran a 2 (categorization: categorization, no categorization) × 2 (stance: negative, positive) × 3 (type: initial, novel, evaluative) mixed-effects ANOVA on the proportion of trials that participants dichotomized the groups. Categorization and stance were between-subjects factors and type was a within-subjects factor. This analysis resulted in a main effect for categorization, $F(1, 206) = 195.48$, $p < .001$, $\eta_p^2 = .49$, and type, $F(2, 412) = 6.09$, $p = .002$, $\eta_p^2 = .03$, qualified by Categorization × Type, $F(2, 412) = 17.51$, $p < .001$, $\eta_p^2 = .08$, and Categorization × Stance × Type interactions, $F(2, 412) = 3.50$, $p = .031$, $\eta_p^2 = .02$.

As shown in Table 2 and Fig. 3, participants in the categorization conditions dichotomized the groups more often than those in the no-categorization conditions for each of the judgment types (initial, novel, evaluative) and regardless of intergroup stance ($ps < .001$). Although stance had no effect in the categorization conditions (Cat/Neg = Cat/Pos for initial, novel, and evaluative traits, $ps > .647$), negative intergroup relations boosted dichotomizing responses in the no-categorization conditions for the initial (NoCat/Neg > NoCat/Pos, $p = .009$) and novel questioning sets (NoCat/Neg > NoCat/Pos, $p = .008$). Participants in the categorization conditions (Cat/Neg, Cat/Pos) dichotomized the two groups more frequently than would be expected by chance for all judgment types (see chance analyses below), but they did so more frequently for initial > novel > evaluative ($ps < .045$; but for Cat/Pos: initial = novel, $p = .064$). In contrast, participants in the no-categorization conditions dichotomized the groups at comparably low rates and below chance (see below) regardless of judgment type ($ps > .230$; but for NoCat/Pos: evaluative > initial, novel, $ps < .018$).

Participants who completed the social categorization measure prior to the uncharacterized group task (Cat/Neg, Cat/Pos) assumed that the two groups had opposite traits (compared to chance [0.50]) across all judgment types (initial, novel, evaluative; Cat/Neg: $t[51]s > 3.18$, $ps < .003$, $ds > 0.44$; Cat/Pos: $t[51]s > 2.90$, $ps < .005$, $ds > 0.40$). In contrast, participants in the no-categorization conditions (NoCat/Neg, NoCat/Pos) dichotomized the groups below chance across all judgment types (NoCat/Neg: $t[53]s > 3.74$, $ps < .001$, $ds > 0.51$; NoCat/Pos: $t[51]s > 7.25$, $ps < .001$, $ds > 1.01$). These effects also consistently occurred at the individual level (see *Tables S2, S3, S4*). For example, whereas 85% of participants in the Cat/Neg and Cat/Pos conditions dichotomized the groups on more than 50% of trials, only 24% of participants in the NoCat/Neg and 10% of participants in NoCat/Pos conditions followed this pattern.

Study 2 clarified the factors that lead people to use a dichotomizing heuristic when forming impressions of social groups: Adults only displayed this rule-of-thumb shortcut at high frequency when they perceived the two groups as separate via making characteristic-to-category inferences during the social categorization task. In contrast, animosity between the groups was not necessary. Beyond illustrating what catalyzes this cognitive shortcut, we also documented its expansive range. That is, we discovered that adults not only assumed that the uncharacterized group (Twiggums) differed from the characterized group (Zuttles) when reasoning about the initial set of preferences and abilities that they applied during the social categorization task (e.g., whether a group likes pears), but also when considering novel preferences and abilities (e.g., whether a group likes modies) as well as evaluative traits (e.g.,

[5] After responding to the key measures reported in this manuscript (i.e., the social categorization task and the uncharacterized group judgment task), participants answered some exploratory pilot questions (i.e., how similar participants viewed themselves to Zuttles; how similar participants viewed themselves to Twiggums; how similar participants viewed Zuttles and Twiggums; and what they thought some of the novel words meant).

[6] Although it was not a primary goal to replicate the findings from Goldfarb et al. (2017), the design of the current study afforded this opportunity. Conceptually replicating Goldfarb et al. (2017), participants categorized individuals at comparable rates regardless of the consequence (i.e., deportation vs. invitation to a "welcome" party). Also replicating Goldfarb et al. (2017), we found that participants were more likely to categorize individuals as Zuttles the more evidence matched what they learned about Zuttles, and adults were willing to categorize an unknown individual as a Zuttle at above-chance levels as long as at least one piece of evidence matched what they had learned about Zuttles. See Online Supplemental Material for analytic details.

Table 2
Study 2: Means (SDs) of Proportion of Trials Dichotomized.

	Categorization- Negative (n = 52)	Categorization- Positive (n = 52)	No-Categorization-Negative (n = 54)	No-Categorization- Positive (n = 52)	All Conditions
Initial	0.83 (0.29)	0.85 (0.31)	0.28 (0.39)	0.11 (0.22)	0.51 (0.45)
Novel	0.74 (0.36)	0.77 (0.34)	0.30 (0.38)	0.14 (0.23)	0.49 (0.43)
Evaluative	0.66 (0.35)	0.64 (0.35)	0.26 (0.30)	0.23 (0.26)	0.44 (0.37)
All	0.74 (0.29)	0.75 (0.28)	0.28 (0.32)	0.16 (0.18)	0.48 (0.38)

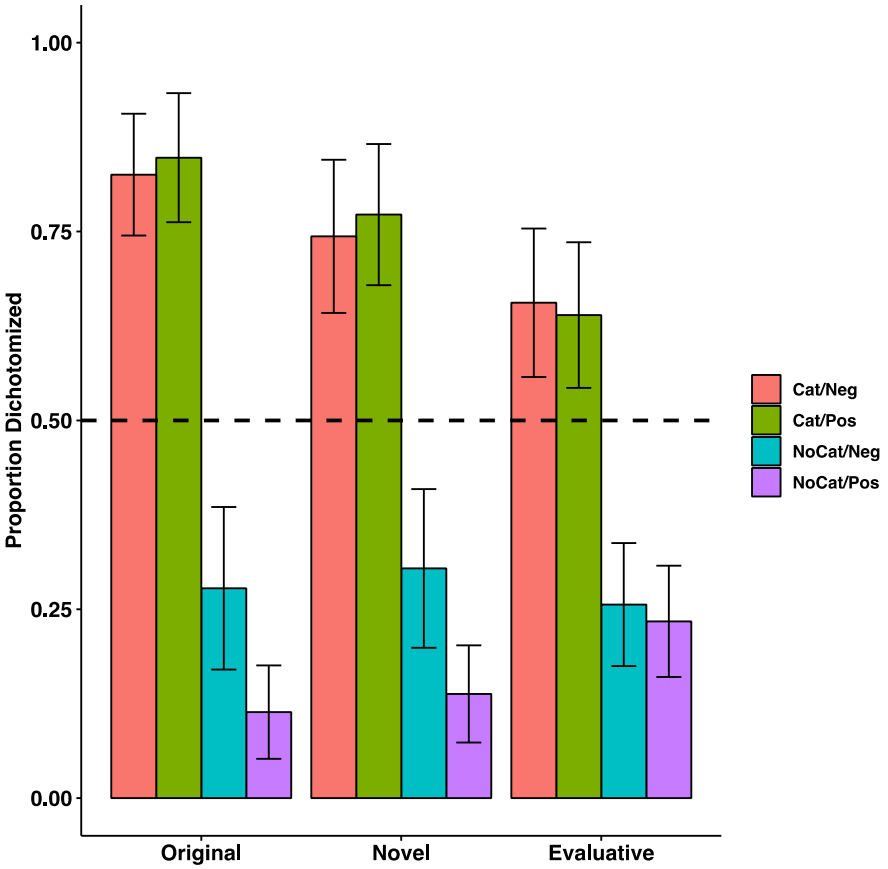

Fig. 3. Study 2: Proportion of trials that participants dichotomized the Zuttles (characterized group) and Twiggums (uncharacterized group) by type (Initial, Novel, Evaluative), categorization (categorization, no-categorization), and stance conditions (positive, negative). Note. Error bars are 95% confidence intervals. Dashed line signifies chance. Cat/Neg = Categorization/Negative; Cat/Pos = Categorization/Positive; NoCat/Neg = No-Categorization/Negative; NoCat/Pos = No-Categorization/Positive.

whether a group is lazy).

4. Study 3

Study 2 demonstrated that repetitively attempting to sort individuals into separate social groups based on evidence about each individual's characteristics was necessary for adults to show the dichotomizing heuristic when later asked to make inferences about the preferences and skills of an uncharacterized group. It is possible, however, that the categorization task would have carried less weight if participants had not been told that Zuttles and Twiggums shared an identical biological and social origin as well as looked the same. Thus, in Study 3, we tested whether making the social groups functionally separate from the start would be sufficient to elicit the dichotomizing heuristic, even in the absence of the categorization task.

Study 3 also addressed two additional methodological limitations. In Studies 1 and 2 as well as in Moty and Rhodes (2021), the narratives highlighted binary contrasts: There were two groups (i.e., Zuttles vs. Twiggums) and participants made binary choices (e.g., They judged whether Twiggums were either good or bad at basketball). Perhaps these binary cues augmented the dichotomizing heuristic. Designing the prior studies in these ways was sensible because we were interested in capturing what participants assumed

about two groups when forced to choose. From a methodological and theoretical perspective, however, it is important to examine how these factors influence the dichotomizing heuristic. In Study 3, therefore, we allowed participants to make scaled (as opposed to binary) judgments, and we manipulated between subjects whether participants learned that there were many groups versus only two groups who lived in the region.

Because we found that participants in Study 2 were willing to dichotomize groups even when reasoning about evaluative traits (e.g., whether a social group was smart or lazy), we tested a new potential boundary condition. That is, in Study 3 we measured adults' use of the dichotomizing heuristic when the characterized group (Zuttles) had preferences that one would expect most humans to have (e.g., Zuttles dislike getting sick; Zuttles like feeling happy). We will refer to these characteristics as *universal*. Here, to infer that the uncharacterized group (Twiggums) had opposite characteristics would require participants to judge that this group held extremely atypical preferences (e.g., they like getting sick; they dislike feeling happy).

We also investigated a possible outcome related to the dichotomizing heuristic. In particular, we were curious about whether the tendency to dichotomize is related to beliefs about broader intergroup relations. To do this, we explored whether individuals who assumed greater differences between Zuttles and Twiggums also expected that the two groups lived more segregated (versus integrated) lives in their social communities. We included questions about whether Zuttles and Twiggums lived in the same neighborhoods, went to the same schools, were friends, were romantically involved with each other, and went to the same doctors. Assessing beliefs about the Twiggum/Zuttle society is also of interest because it speaks to the contexts under which people might utilize the dichotomizing heuristic. For example, do adults only apply the dichotomizing heuristic when they assume that the groups live separate lives or do they still make this cognitive shortcut even when they think that a society is well integrated?

4.1. Method

4.1.1. Participants

Participants were 212 undergraduate students ($M = 20.20$ years, $SD = 2.25$ years, 43 males, 162 females, 1 gender non-binary).[7] Included participants were fluent in comprehending and speaking English and were typically developing (via self-report). The sample was 39% East Asian (e.g., Chinese, Korean, Japanese), 1% Native Hawaiian or Other Pacific Islander, 11% South Asian (e.g., Indian, Pakistani), 12% South East Asian (e.g., Thai, Vietnamese, Filipino), 8% White, and 10% another race or ethnicity or multiracial.[8] We set the sample size at 53 participants per cell ($N = 212$) to remain consistent with Study 2. In addition, 53 participants per condition gives us 82% power to detect a $d = 0.40$ or greater (the smallest effect detected in Study 2) for a one-sample *t*-test. We stopped data collection the day this goal was met. We included five attention checks throughout the study (e.g., Please select the number 3; options: 1, 2, 3, 4, 5). We excluded 6 participants who missed two or more of these checks. The final sample included 206 participants. Participants received course credit.

Data were collected in February 2021. This study was approved by the Internal Review Board at the University of California, Davis: Protocol Number: 1031991. The sample size (including exclusions), method, and planned analyses were pre-registered on the Open Science Framework (https://osf.io/73gdh/?view_only=0889aaedfd3b4da6a414f3440d937116).

4.1.2. Measures and Procedure

This study followed a 2 (size: two, many) × 2 (categorization: categorization, no-categorization) × 4 (type: initial, novel, evaluative, universal) design. Size was a between-subjects factors that varied in whether participants learned that two groups lived on the island (Two) or whether many groups lived on the island (Many). Categorization was a between-subjects factor that differed in whether participants did (Cat) or did not (NoCat) complete a social categorization task before the dichotomizing measure. Type was a within-subject factor where participants engaged in four phases of the dichotomizing task which each inquired about different types of characteristics (initial, novel, evaluative, and universal).

Narrative and Social Categorization Task. All participants were tested online via Qualtrics Survey Software (Qualtrics, Provo, UT) due to the restrictions on in-person testing caused by the Covid-19 pandemic. After completing informed consent and basic demographic information, participants responded to the main tasks. All participants learned about an island where groups of humans lived. Half of the participants learned that two groups of humans lived on the island (Two-Group Condition: "Imagine that there are two groups of humans who live on an island. One of the groups is called Zuttles. The other group is called Twiggums"). In the other condition, participants learned that many groups of humans lived on the island (Many-Group Condition: "Imagine that there are many groups of humans who live on an island. One of the groups is called Zuttles."). We included a comprehension check ("So, how many groups of humans live on the island?" Choices: 1, 2, many). Participants were not able to move on until they answered this question correctly. In both conditions, we avoided using any "difference" language so as to not cue that the groups differed beyond having separate category labels.

Next, all participants learned about the 12 Zuttle characteristics (i.e., their preferences and abilities as in Studies 1 and 2; e.g., Zuttles like pears). Half of the participants then completed the categorization task. In the two-group version (Two-Group/Categorization [Two/Cat]), the categorization task was the same as Study 2. For participants in the many-groups version (Many-Group/

[7] Three additional participants consented to the study but did not complete it (i.e., they were never assigned to a condition). One of these participants provided some basic demographic information, but did not respond to any of the other variables. The other two participants, consented but then responded to no additional variables.

[8] Between Study 2 and Study 3 we updated our basic demographics to be more specific.

Categorization [Many/Cat]), the only difference was that participants were asked to determine whether each individual was a Zuttle or "someone from another group." Participants in the no-categorization condition did not complete the categorization task (Two-Group/No-Categorization [Two/NoCat]; Many-Group/No-Categorization [Many/NoCat]).

Uncharacterized Group Task. All participants then completed four questioning sets where they made inferences about the preferences, skills, and traits of the uncharacterized group, the Twiggums. Participants in the Two Group Condition (Two/Cat, Two/NoCat) were introduced to the task in the same way as participants in Study 2. Participants in the Many Group Condition (Many/Cat, Many/NoCat) were told that they would be answering questions about one of the other groups, Twiggums. These were similar questions to Study 2 except that participants made judgments about Twiggums' preferences, abilities, and traits using a 5-point scale rather than making a binary choice. These questioning sets followed the same fixed order for all participants.

First, participants responded to the *initial set* of the 12 characteristics that were included in the Zuttle familiarization narrative (all conditions) and also appeared in the categorization task (Two/Cat, Many/Cat). On each trial, participants were reminded of the Zuttles' characteristic, and then they made a judgment about Twiggums (e.g., "Remember…Zuttles like pears. So, what do you think about Twiggums?", 5-point scale: ranging from Twiggums DISLIKE pears to Twiggums LIKE pears, midpoint = Twiggums are NEUTRAL in this preference).

Second, participants responded to the *novel set* of 12 preferences and abilities described using novel words (e.g., being good at playing daxes). Participants responded to four blocks of three questions (presented in random order; e.g., "Here are some activities that you have never heard about. Zuttles are good at playing lups, daxes, and nafs. So, what do you think about Twiggums?", 5-point scale: Twiggums are BAD at playing lups to Twiggums are GOOD at playing lups, midpoint = Twiggums are NEUTRAL in this skill).

Third, in the *evaluative set* participants were told 12 evaluative traits about Zuttles and made inferences about Twiggums. Participants learned about 4 positive traits (smart, creative, honest, funny), 4 equivocal traits (shy, confident, superstitious, liberal), and 4 negative traits (mean, moody, messy, lazy). For each trait, participants were randomly assigned to learn that Zuttles either did (two of each valence: positive, equivocal, negative) or did not (two of each valence) possess the trait (e.g., "Zuttles are smart. So, what do you think about Twiggums?", 5-point scale: Twiggums ARE NOT smart to Twiggums ARE smart, midpoint = Twiggums are NEUTRAL in this trait). Traits were presented in randomized order.

Fourth, in the *universal set* participants responded to 12 preference trials that involved Zuttles holding universal attitudes (e.g., Zuttles dislike getting sick). Participants learned that Zuttles had 6 universal dislikes (e.g., Zuttles dislike the sound of nails on a chalkboard) and 6 universal likes (e.g., Zuttles like having fun). Participants then made judgments about Twiggums on a 5-point scale (e.g., "So, what do you think about Twiggums?", 5-point scale: Twiggums DISLIKE getting sick to Twiggums LIKE getting sick, midpoint = Twiggums are NEUTRAL in this preference). These items were presented in a random order.

We used a 5-point scale for several reasons. It mirrored prior studies in that participants had to make a judgment about Twiggums' characteristics. It also allowed us keep the same training (i.e., the items that Zuttles liked/disliked and were good/bad at),[9] limiting the number of changes we made from prior studies (maximizing what we could learn in this study). Participants could pick the scale midpoint instead of being forced to pick one side of the scale creating a more stringent test of the dichotomizing heuristic. We labeled the midpoint "neutral in this preference/skill/trait" to limit ambiguity about the scale values. Moreover, dichotomizing the groups on this scale required participants to switch sides of the scale depending on the trial. That is, for half of the trials, dichotomizing required participants to pick a "1" or a "2" on the scale whereas for the other half of trials, dichotomizing would lead participants to pick a "4" or a "5."

Coding and scoring. For all questioning sets, the items were rescored so that higher scores indicated a greater difference between the characteristics of Zuttles and the characteristics of Twiggums. For all questioning sets, responses showed excellent internal consistency (initial: Cronbach's alpha = 0.91, 95% CI [0.89, 0.93]; novel: Cronbach's alpha = 0.95, 95% CI [0.93, 0.96]; evaluative: Cronbach's alpha = 0.91, 95% CI [0.89, 0.93]; universal: Cronbach's alpha = 0.94, 95% CI [0.93, 0.95]). We also rescored the judgments to approximate a dichotomous choice. That is, participants received a score of 1 for every trial that they chose a 4 or 5 (i.e., chose the opposite side of the scale for Twiggums from what they were told about Zuttles) and a score of 0 for every trial that they chose a 1, 2, or 3 on the scale (i.e., chose the same side of the scale for Twiggums as they were told about Zuttles or chose the neutral). See *Tables S5, S6, S7, and S8,* individual level trial data. Including the neutral scale point as a *failure* to dichotomize provides a more conservative test of the dichotomizing heuristic in that we required participants to actually endorse that the uncharacterized group (Twiggums) held *opposite* characteristics.

Proximity Task. At the end of the study, participants responded to questions examining their perceptions of the island where Zuttles and Twiggums live. We were interested in whether participants believed that Zuttles and Twiggums lived segregated or integrated lives on the island. Participants responded to a series of questions that required a binary response: (1) When you picture the island where Zuttles and Twiggums live, Zuttles likely… (a) Only live in neighborhoods with Zuttles or (b) Live in neighborhoods with both Zuttles and Twiggums; (2) When you picture the island where Zuttles and Twiggums live, Zuttle children likely… (a) Only go to school with Zuttles or (b) Go to school with both Zuttles and Twiggums; (3) When you picture the island where Zuttles and Twiggums live, Zuttles are likely… (a) Only friends with Zuttles or (b) Friends with both Zuttles and Twiggums; (4) When you picture the island where Zuttles and Twiggums live, Zuttles are likely… (a) Only romantically involved with Zuttles or (b) Romantically involved with Zuttles or Twiggums; (5) When you picture the island where Zuttles live, Zuttles likely… (a) Only go to doctors and hospitals used by Zuttles or (b) Go to doctors and hospitals used by both Zuttles and Twiggums. Participants viewed questions in a random order

[9] The only exception to this was that in the current study participants learned that Zuttles either liked or disliked certain foods rather than learning that they either liked or hated certain foods.

(response options were also randomized). These items showed good internal consistency (Cronbach's alpha = 0.85, 95% CI [0.82, 0.88]). For analyses, we created a score of the proportion of trials that participants judged that the two groups lived segregated lives.

4.2. Results and discussion

Results are divided into two sections. First, we assessed participants' responses to the uncharacterized group task, both their scaled ratings and their rescored responses that approximated a dichotomous choice. Second, we explored the data from the proximity task. Analyses were conducted in RStudio (version 3.6.1; R Core Team, 2019). See Online Supplemental Materials for analyses on the social categorization task alone.

4.2.1. Uncharacterized group task

We ran a 2 (size: two groups, many groups) × 2 (categorization: categorization, no categorization) × 4 (type: initial, novel, evaluative, universal) mixed-effects ANOVA on the 5-point ratings (the higher the score the greater the difference between Twiggums and Zuttles). Size and categorization were between-subjects' factors and type was a within-subjects' factor. This analysis resulted in a main effect for categorization, $F(1, 202) = 14.41$, $p < .001$, $\eta_p^2 = .07$, and type, $F(3, 606) = 121.25$, $p < .001$, $\eta_p^2 = .38$ (see Table 3). There were no other significant effects (including any for group size), $Fs < 3.42$, $ps > .066$, $\eta_p^2 s < .02$. Thus, conceptually replicating and extending Study 2, participants in the categorization conditions assumed greater differences between the groups than did participants in the no categorization conditions ($p < .001$). In addition, participants expected greater differences between Twiggums and Zuttles when making inferences about initial, novel, and evaluative characteristics as compared to universal characteristics ($ps < .001$; all other comparisons: $ps > .382$).

We used the rescored data (see Coding and Scoring) to examine whether participants dichotomized the groups more frequently than would be expected by chance. We ran a two-tailed *t*-test compared to chance (0.40). Chance is 0.40 because on any given trial there was a 40% possibility of choosing a 4 (0.20) or a 5 (0.20) at random (0.20 + 0.20 = 0.40).[10] Participants in both conditions dichotomized the groups more often than would be expected by chance for initial (e.g., Zuttles like apples), novel (e.g., Zuttles like modies), and evaluative items (e.g., Zuttles are smart; Cat: $Ms > 0.55$, $t[102]s > 4.26$, $ps < .001$, $ds > 0.42$; NoCat: $Ms > 0.48$, $t[102]s > 2.30$, $ps < .023$, $ds > 0.23$; Fig. 4). Thus, the categorization task was unnecessary to create the heuristic when the two groups were labeled as distinct from the start and had no stated common origin. We also found that participants dichotomized Twiggums' and Zuttles' initial, novel, and evaluative traits above chance irrespective of whether they learned that there were two or many groups on the island (Two Groups: $Ms > 0.56$, $t[103]s > 3.95$, $ps < .001$, $ds > 0.39$; Many Groups: $Ms > 0.48$, $t[101]s > 2.26$, $ps < .026$, $ds > 0.22$). Therefore, we found no evidence that use of the dichotomizing heuristic is dependent on there being an explicit contrast between two groups.

We did, however, discover an important and reasonable boundary condition for the dichotomizing heuristic. Participants in the categorization, no categorization, two groups, and many groups conditions assumed that the groups would be aligned in universal preferences (e.g., they both like feeling happy; $Ms < 0.32$, $|t|s > 2.07$, $ps < .041$, $ds > 0.20$). Said another way, participants were unwilling to view Twiggums as opposite to Zuttles when doing so required that Twiggums had highly atypical preferences (e.g., Twiggums like getting gum stuck in their hair).[11]

4.2.2. Proximity task

We conducted a 2 (size) × 2 (categorization) ANOVA on the proportion of trials that participants assumed that the groups lived segregated lives. All effects were null, $Fs < 1.62$, $ps > .205$, $\eta_p^2 s < .01$. Participants expected that Twiggums' and Zuttles' lives were more integrated than segregated ($M = 0.36$, 95% CI [0.31, 0.42], $t[205] = -5.17$, $p < .001$, $d = 0.36$).[12] This suggests that participants were willing to use a dichotomizing heuristic even when they viewed the societies as more integrated than segregated. We also examined correlations between dichotomizing ratings (i.e., the 5-point scale) and beliefs about segregation. Controlling for size and categorization conditions, the greater the difference that participants expected between Twiggums and Zuttles (collapsing across judgment type), the more frequently they expected the two groups to live segregated lives ($r[202] = 0.23$, 95% CI [0.10, 0.36], $p < .001$).[13]

[10] We had planned to conduct these t-tests at the Categorization x Size level, but because this interaction was not significant, it was more appropriate to examine the level of the main effects. Table 3 shows the means and standard deviations by each condition.

[11] We also conducted one-sample two-tailed t-tests on the proportion of trials that participants selected a 2, 3, 4, or 5 on the scale (i.e., selected a characteristic that differed from Zuttles) compared to chance (0.80). For initial, novel, and evaluative characteristics, participants in the categorization, in the two group, and in the many group conditions avoided selecting "1″ on the scale (i.e., saying that Twiggums were the same as Zuttles; $Ms > 0.85$, $ts > 2.18$, $ps < 0.032$, $ds > 0.22$), but participants in the no-categorization condition did not avoid selecting "1" more frequently than would be expected by chance for initial, novel, and evaluative characteristics ($Ms < 0.82$, $ts < 0.65$, $ps > 0.515$, $ds < 0.06$). In contrast, participants in the both categorization conditions and in the both group conditions chose "1" more frequently than would be expected chance for universal items ($Ms < 0.66$, $ts > 3.74$, $ps < 0.001$, $ds > 0.37$).

[12] We had planned to run this analysis by condition, but the lack of effects in the ANOVA led us to collapse across conditions (see *Table 3* for means and standard deviations by condition).

[13] This same relation held when we examined each of the individual judgments separately (initial: $r[2\ 0\ 2] = 0.21$,95% CI [0.07, 0.34], $p = .003$; novel: $r[2\ 0\ 2] = 0.19$, 95% CI [0.05, 0.32], $p = .007$; evaluative: $r[2\ 0\ 2] = 0.16$, 95% CI [0.03, 0.29], $p = .019$; universal: $r[2\ 0\ 2] = 0.17$, 95% CI [0.04, 0.30], $p = .013$).

Table 3
Study 3 Means (SDs) of Dichotomizing Ratings, Proportion of Trials Dichotomized, and Proximity Questions.

	Two-Categorization (n = 53)	Two-No-Categorization (n = 51)	Many-Categorization (n = 50)	Many-No-Categorization (n = 52)	All Conditions
	Dichotomizing Ratings				
Initial	4.01 (0.86)	3.48 (1.06)	3.69 (0.90)	3.37 (1.04)	3.64 (0.99)
Novel	3.89 (1.03)	3.42 (1.21)	3.52 (0.93)	3.33 (1.12)	3.54 (1.09)
Evaluative	3.96 (0.94)	3.29 (1.11)	3.43 (0.79)	3.40 (0.96)	3.52 (0.99)
Universal	2.95 (1.36)	2.03 (1.05)	2.58 (1.15)	2.23 (0.97)	2.45 (1.19)
All	3.70 (0.82)	3.05 (0.92)	3.30 (0.74)	3.07 (0.80)	3.28 (0.86)
	Proportion of Trials Dichotomized				
Initial	0.66 (0.33)	0.54 (0.35)	0.56 (0.37)	0.50 (0.33)	0.56 (0.35)
Novel	0.62 (0.40)	0.49 (0.40)	0.52 (0.39)	0.48 (0.35)	0.53 (0.39)
Evaluative	0.64 (0.36)	0.47 (0.33)	0.46 (0.35)	0.49 (0.33)	0.52 (0.35)
Universal	0.37 (0.39)	0.18 (0.28)	0.28 (0.34)	0.19 (0.25)	0.26 (0.33)
All	0.57 (0.31)	0.42 (0.29)	0.45 (0.30)	0.41 (0.26)	0.47 (0.29)
	Proximity Questions				
Neighborhoods	0.38 (0.49)	0.40 (0.49)	0.37 (0.49)	0.46 (0.50)	0.40 (0.49)
School	0.42 (0.50)	0.24 (0.43)	0.27 (0.45)	0.37 (0.49)	0.33 (0.47)
Friends	0.38 (0.49)	0.34 (0.48)	0.38 (0.49)	0.38 (0.49)	0.37 (0.48)
Romantic	0.45 (0.50)	0.32 (0.47)	0.35 (0.48)	0.40 (0.50)	0.38 (0.49)
Healthcare	0.38 (0.49)	0.26 (0.44)	0.35 (0.48)	0.35 (0.48)	0.33 (0.47)
All	0.40 (0.42)	0.31 (0.35)	0.35 (0.39)	0.39 (0.37)	0.36 (0.38)

Note. Dichotomizing ratings range from 1 to 5 (higher scores indicate that participants expect greater differences between Zuttles and Twiggums. Proportion of Trials Dichotomized range from 0 to 1 (higher scores are indicative of dichotomizing; on these trials, chance is 0.40). Proximity Questions range from 0 to 1 (higher scores are indicative of beliefs that the groups are more segregated).

5. General discussion

Individuals apply a *dichotomizing heuristic* when forming impressions of social groups: They use what they know about one group to infer that the opposite is true about an uncharacterized group. Although participants exhibited this shortcut across the generic and specific language conditions in Study 1, learning about the characterized group as a generic whole boosted its frequency. In Studies 2 and 3 we discovered that conceptualizing the groups as distinct entities (either by engaging in a social categorization task or by initially being told that they are separate groups) is a prerequisite for eliciting the shortcut among adults. Once triggered, however, adults applied the dichotomizing heuristic when making category-to-characteristic inferences about a novel group's benign, novel, and evaluative traits. Studies 2 and 3 also demonstrated that intergroup conflict (i.e., present versus absent) or the number of groups who inhabit the geographical area (i.e., two versus many groups) did not measurably shape use of the heuristic. Moreover, adults even exhibited the cognitive shortcut when making scaled rather than binary judgments, and their beliefs did not appear to be fully driven by an assumption that the two groups lived segregated lives. Still, participants exhibited a rational boundary: They did not apply the dichotomizing heuristic when doing so would require them to say that the uncharacterized group held highly unusual preferences (e.g., liking to be sick; disliking to be happy). Below, we expand on these findings and propose suggestions for future research.

5.1. Documenting a dichotomizing heuristic

When deciding what a new social group is like, a person should ideally take time to make thoughtful evaluations, check for consistencies and inconsistencies in the kinds of characteristics group members express, and avoid making unfounded generalities. Although such practices could help foster fair, accurate impressions, they are time and resource intensive, and as a result, people often use shortcuts. For example, the *availability heuristic*—the ease with which something comes to mind is viewed as a cue to its prevalence (Tversky & Kahneman, 1973)—leads adults to infer that repeatedly observed traits (e.g., laziness), even if only paired with one group member, describe the entire group (Rothbart et al., 1978). Children and adults also expect bidirectional individual-to-group relations: They extend traits from an individual to the group as well as from the group to an individual (Disendruck et al., 2015; Gelman et al., 2010; Rhodes & Gelman, 2008; Sherman et al., 2009).

Here, we discovered another way that individuals streamline social judgments: They use a dichotomizing heuristic. The heuristic operated in the *absence* of direct contact with, prior knowledge of, or learned "facts" about the new group in question. Instead, it only required first conceptualizing the uncharacterized group as separate and distinct from the characterized group. Even without making group differences salient (as in Moty & Rhodes, 2021), the dichotomizing heuristic emerged across several contexts. Participants exhibited the effect when they learned about the characterized group as a generic whole and as individual group members (Study 1), when there was and was not intergroup conflict (Study 2), when they learned about the characterized group in the context of just one other group or in the context of many other groups (Study 3), and when participants made binary (Study 1 and Study 2) as well as scaled judgments (Study 3). Moreover, adults not only applied it to features that they initially learned about Zuttles that were benign (e.g., liking apples), but they also extended it to novel preferences and skills (e.g., liking modies), and even to evaluative traits that signaled the morals and virtues of a group (e.g., being mean). Notably, among these manipulations, the only context where they did not

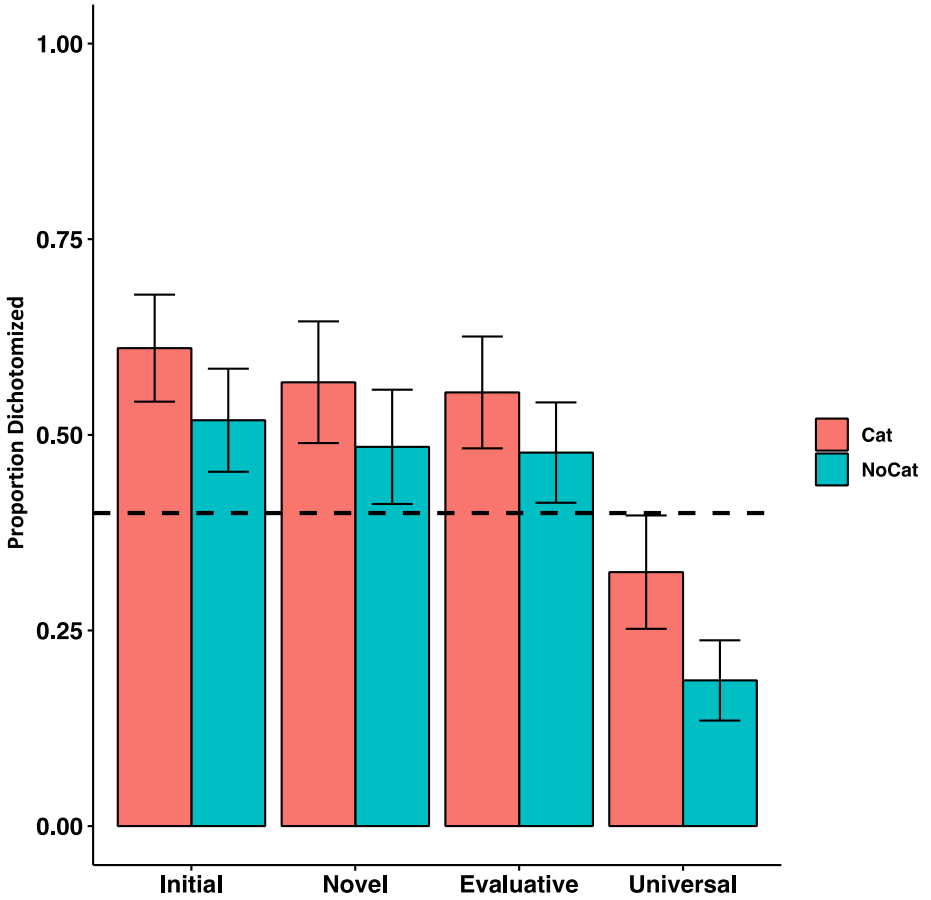

Fig. 4. Study 3: Proportion of trials that participants dichotomized the Zuttles (characterized group) and Twiggums (uncharacterized group) by type (Initial, Novel, Evaluative, Universal) and categorization condition (Cat, NoCat). *Note.* Error bars are 95% confidence intervals. Dashed line signifies chance (0.40). Cat = Categorization; NoCat = No-Categorization.

use the heuristic was when doing so would render Twiggums as having extremely atypical human preferences (e.g., liking spiders crawling on them; disliking winning money).

Several aspects of human cognition likely contribute to the dichotomizing heuristic. Social categorization is powerful: Partitioning individuals into groups can change people's motivations and cognitions (Hugenberg et al., 2010; Young & Hugenberg, 2010; Qian et al., 2019). For example, category labels cue between-group differences in properties, characteristics, and traits (Eiser, 1971; Hamilton & Gifford, 1976; Johnston & Jacobs, 2003; Krueger & Rothbart, 1990; Lawson & Bower, 2014; Master et al., 2012; Tajfel & Wilkes, 1963). Children and adults also utilize speech pragmatics to infer meaning from both what the speaker says and what they leave out (i.e., implied contrasts; Bohn & Frank, 2019; Diesendruck & Markson, 2001; Grice, 1975; Halberda, 2006; Markman et al., 2003; Papfragou & Musolino, 2003). These beliefs about group differences, combined with language pragmatics, likely work in tandem to facilitate the dichotomizing heuristic. In other words, people have general beliefs that distinct categories differ on at least some factors. Given this starting assumption, people may then search for the features that differentiate groups (Sherman et al., 2009). When a speaker provides information about one group (e.g., Zuttles like apples), the listener may infer that the decision to exclude information about the second group was purposeful (i.e., this must be the key differentiating information). Said another way, the listener may assume that if the information was relevant to both groups then the speaker would have said as much.

Importantly, however, data from Study 3 revealed that the dichotomizing heuristic cannot be reduced to pragmatics alone. In contrast to inferring that Twiggums had opposite characteristics to Zuttles when judging benign, novel, and evaluative traits, participants predicted that Twiggums and Zuttles would *share* preferences considered universal to humans (e.g., they expected both Twiggums and Zuttles to like feeling happy and to dislike getting hurt). This evidence highlights a clear boundary condition for the heuristic—adults are unlikely to apply it when doing so would lead them to make highly improbable assumptions about a social group. Potentially, then, adults' use of the dichotomizing heuristic may be influenced by their beliefs about how common or uncommon a particular characteristic is in a population. Still, many of the characteristics participants consistently dichotomized for initial and evaluative trait sets could also be viewed as widespread (e.g., lots of people like apples). Thus, base rates may only affect the dichotomizing heuristic in extreme boundary cases such as those tested in Study 3. Future studies are needed that manipulate base rates and test dichotomous thinking about social groups. It also would be important to try to disentangle whether the effect arises from

inferences about groups in general, versus inferences based on intentionally-communicated information about groups (see Schwarz, 1994).

5.2. Catalysts of the dichotomizing heuristic

Generic language is pervasive: Its ubiquity crosses the boundaries of language (Chierchia, 1998) and context, including parent–child conversations (Gelman et al., 2008) and news reports (Dukes & Gaither, 2017). Previous studies have shown that hearing about groups in generic versus specific language encourages children and adults to assume greater homogeneity within a group (Rhodes et al., 2012). Here, we show that generic language also permeates concepts of uncharacterized groups (see also Moty & Rhodes, 2021). That is, generic language (versus specific language) inflated the assumption that two groups must have opposite preferences and abilities despite information only being provided about one of the groups. This type of speech may have made the descriptions of the characterized group appear more exclusive; generic language is deemed more appropriate when a characteristic is distinctive (i.e., characteristic is specific to that group; e.g., lions have manes) versus non-distinctive (i.e., characteristic is widespread across groups; e.g., lions are males; Cimpian et al., 2010; Leslie, 2008). Although dampened, participants who learned about individual Zuttles in specific language condition also dichotomized the two groups. Thus, while generics appears to further cement the use of the shortcut, discussing group members as individuals does not eradicate unsupported inferences about unknown groups (see also Goldfarb et al., 2017).

Across studies, we learned that participants apply the dichotomizing heuristic if they conceptualize the groups as distinct and separate, and there are multiple avenues for forming this impression. When two groups derived from the same biological and social origin and looked identical (Studies 1 and 2), a brief experience of attempting to categorize individuals into two groups based only on information about the preferences and skills of one group led participants to later judge that the two groups must have opposite traits. Another approach is to tell participants that there are two or more social groups (no origin information), describe characteristics about one group, and then ask them to infer what the other group is like (Study 3). This strategy produced the dichotomizing effect even though we avoided language that would cue differences beyond the groups having separate labels (i.e., we didn't use leading words such as "different," "distinct," or "separate" when discussing the groups and provided no evidence that they had distinguishing physical features). The least subtle pathway to elicit the heuristic is to describe groups are "different kinds" and show that they look different too—enabling children as young as 4.5 years of age to dichotomize groups (Moty & Rhodes, 2021). Indeed, our developmental findings from Study 1 underscore the interpretation that perceiving groups as distinct is a necessary trigger of the dichotomizing heuristic. Consistent with prior work indicating weaker category boundaries and less consistent characteristic-to-category judgments in children 5 years and younger compared to older children and adults (Goldfarb et al., 2017; Lagattuta & Kramer, 2021; Rhodes & Gelman, 2008; Riggs et al., 2014; Shutts, 2015; Sloutsky et al., 2015), the social categorization task may not have sufficiently induced 5-year-olds to treat Zuttles and Twiggums as distinct groups. Young children likely need, as in Moty and Rhodes (2021), to be unambiguously told and shown that members of one group differ from members of another group for them to apply this cognitive shortcut.

Studies 2 and 3 further identified factors that failed to influence adults' use of the dichotomizing heuristic. First, we did not find strong support that intergroup conflict (versus goodwill between groups) mattered: Adults dichotomized groups at equivalent rates regardless of whether Twiggums prohibited versus welcomed Zuttles to their community. Still, there may be contexts in which negative intergroup relations could elicit the shortcut that groups must hold opposite characteristics. For example, in the current research, we used novel groups to which the children and adults did not belong so that neither existing knowledge nor in-group bias (Killen, 2007; Rhodes et al., 2012; Roberts et al., 2017) could shape judgments. Potentially, if participants were members of one of the groups, intergroup animosity could boost dichotomizing rates. Second, Study 3 showed that the number of social groups featured in the learning context (i.e., two versus many) did not affect participants' use of the dichotomizing heuristic. Both participants in the two-group condition and those in the many-group condition dichotomized the groups, suggesting that the heuristic was not simply produced by an assumption that we (as the researchers) had intentionally set up an explicit contrast between the two groups.

5.3. Limitations and future directions

The current studies only tested the dichotomizing heuristic using an experimental paradigm featuring novel groups. We think it is likely that people also apply this cognitive shortcut in real-world settings. For example, the dichotomizing heuristic could contribute to why some individuals erroneously assume that the Black Lives Matter movement means that other lives do not matter. Indeed, beliefs about the dichotomous nature of social groups could materialize in opposition towards organizations and institutions that support or give voice to marginalized groups. Such assumptions may be especially likely when people also hold zero-sum beliefs about social groups (e.g., as one group's situation improves that must mean that another group's has worsened; Norton & Sommers, 2011). Another example comes from our own experience seeing this shortcut play out: An author's son took issue with his sister's "Girls are strong" t-shirt and shouted, "Boys aren't weak!"[14] We could also imagine how this form of binary thinking might contribute to the growing political divide in the United States. For example, even before a Democrat learns the contents of a Republican-supported bill, they may

[14] Statements like, "Girls are strong" could also be problematic for reasons beyond the dichotomizing heuristic because if people have a stereotype about a group then these sorts of statements might be read as the speaker lacking confidence because people do not state the obvious (see Chestnut & Markman, 2018; Chestnut et al., 2021).

assume that if Republicans like the bill, then they (as a Democrat) will not like it. These ideas are currently speculative and will require rigorous experimental testing to confirm or refute them.

Additional aspects of the dichotomizing heuristic should be explored in future research. For example, our paradigm oversimplified social group membership: Individuals were either Zuttles or Twiggums. In reality, people belong to multiple social groups. Although we did introduce the presence of multiple groups in Study 3, it also will be important to test how intersectionality influences the dichotomizing heuristic. Indeed, manipulating the saliency of an individual's different social group memberships (see Chiao et al., 2006; Gaither et al., 2014) might alter their use of the dichotomizing heuristic. We also look forward to exploring sources of individual differences in use of the dichotomizing heuristic as well as potential connections between dichotomizing and other outcomes. For example, in Study 3, we found that individuals who assumed greater differences between the two groups also expected them to live in more segregated communities. We imagine that dichotomizing groups may have several other negative downstream consequences. For example, it may be used for justifying social exclusion in an intergroup context. That is, when deciding whom to befriend 7- to 16-year-olds weight shared interests over ethnic group affiliation when provided with both pieces of information (Hitti & Killen, 2015; McGlothlin & Killen, 2005). Because people presume that preferences and abilities do not cross group lines, members of outgroups may be prematurely excluded: Shared preferences cannot be prioritized if they remain undiscovered.

Finally, our contributions to understanding the developmental trajectory of the dichotomizing heuristic are limited. We found that children younger than 8 did not assume that groups have opposite preferences and abilities. We primarily used these developmental findings to bolster the argument that people must perceive of the groups as separate before they will dichotomize their characteristics (i.e., younger children form weaker category boundaries). We discontinued our investigation of children's use of the heuristic in Studies 2 and 3 given the need to more strongly establish the triggers, scope, and boundaries of the phenomenon in adults. We direct interested readers to Moty and Rhodes (2021) who showed that when groups are portrayed as saliently distinct (e.g., they have different category labels and can be distinguished by their clothes) children as young as 4.5 years will assume between-group differences. Still, from that work it is unclear how far children will generalize the heuristic (i.e., the characteristics they used were all benign—would young children also generalize to novel, evaluative, or universal traits?). Further inquiry into the development of this cognitive shortcut is needed.

5.4. Conclusion

Individuals apply a dichotomizing heuristic when learning about social groups: They assume that what is true of one group is not true of another. In addition to documenting this cognitive shortcut, we also elucidated the circumstances under which it arises. Whereas conceptualizing the two groups as distinct is a necessary trigger of the dichotomizing heuristic and generic language heightens its use, intergroup conflict, learning about only two groups, and binary (versus scaled) judgments are not required. Once activated, adults apply this rule-of-thumb broadly, even to evaluative traits, such as morals, intelligence, and work ethic; refraining only when dichotomizing would lead them to assign highly unusual traits to the uncharacterized group (e.g., disliking feeling happy). Taken together, individuals can think that they know about two social groups when they only have learned about one.

Acknowledgement

This research was funded by an internal grant to KHL. While working on this manuscript, HJK was supported by the Predoctoral Training Consortium in Affective Science from the National Institute of Mental Health (201302291), DG was supported by the National Science Foundation (2031043), and SMT was supported by a National Science Foundation Graduate Research Fellowship (2016207607). DG and SMT were affiliated with the University of California, Davis when we collected the data for Study 1. We thank the individuals who participated. We also thank the Mind-Emotion Development Lab. Portions of this research were presented at the Biennial Meeting of the Cognitive Development Society in Portland, Oregon (2017), and at the Society for Personality and Social Psychology in New Orleans, Louisiana (2020).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cogpsych.2021.101408.

References

Allport, G. W. (1954). *The nature of prejudice*. Oxford: Addison-Wesley.

Bigler, R. S. (1995). The role of classification skill in moderating environmental influences on children's gender stereotyping: A study of the functional use of gender in the classroom. *Child Development, 66*, 1072–1087. https://doi.org/10.1111/j.1467-8624.1995.tb00923.x.

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. https://doi.org/10.1111/j.1467-8624.1997.tb01956.x.

Bohn, M., & Frank, M. C. (2019). The pervasive role of pragmatics in early language. *Annual Review of Developmental Psychology*. https://doi.org/10.1146/annurev-devpsych-121318-0850537.

Chambers, C. G., Graham, S. A., & Turner, J. N. (2008). When hearsay trumps evidence: How generic language guides preschoolers' inferences about unfamiliar things. *Language and Cognitive Processes, 23*, 749–766. https://doi.org/10.1080/01690960701786111.

Chiao, J. Y., Heck, H. E., Nakayama, K., & Ambady, N. (2006). Priming race in biracial observers affects visual search for black and white faces. *Psychological Science, 17*, 387–392. https://doi.org/10.1111/j.1467-9280.2006.01717.x.
Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics, 6*, 339–405. https://doi.org/10.1023/A:100832421.
Cimpian, A., Brandone, A. C., & Gelman, S. A. (2010). Generic statements require little evidence for acceptance but have powerful implications. *Cognitive Science, 34*, 1452–1482. https://doi.org/10.1111/j.1551-6709.2010.01126.x.
Cimpian, A., & Scott, R. M. (2012). Children expect generic knowledge to be widely shared. *Cognition, 123*, 419–433. https://doi.org/10.1016/j.cognition.2012.02.003.
Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–18. https://doi.org/10.10137/0022-3514.56.1.5.
Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology, 37*, 630–641. https://doi.org/10.1037/0012-1649.37.5.630.
Diesendruck, G., Salzer, S., Kushnir, T., & Xu, F. (2015). When choices are not personal: The effect of statistical and social cues on children's inferences about the scope of preferences. *Journal of Cognition and Development, 16*, 370–380. https://doi.org/10.1080/15428372.2013.848870.
Dovidio, J. F., Gaertner, S. L., Hodson, G., Houlette, M. A., & Johnson, K. M. (2005). Social inclusion and exclusion: Recategorization and the perception of intergroup boundaries. In D. Abrams, M. A. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion* (pp. 245–264). Psychology Press.
Dukes, K. N., & Gaither, S. E. (2017). Black racial stereotypes and victim blaming: Implications for media coverage and criminal proceedings in cases of police violence against racial and ethnic minorities. *Journal of Social Issues, 73*, 789–807. https://doi.org/10.1111/josi.12248.
Eiser, J. R. (1971). Enhancement of contrast in absolute judgment of attitude statements. *Journal of Personality and Social Psychology, 17*, 1–10. https://doi.org/10.1037/h0030455.
Esses, V. M., Jackson, L. M., Dovidio, J. F., & Hodson, G. (2005). Instrumental relations among groups: Group competition, conflict, and prejudice. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice* (pp. 227–243). Blackwell Publishing Ltd.
FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behavior, 3*, 426–435. https://doi.org/10.1038/s41562-019-0590-x.
Gaither, S. E., Chen, E. E., Corriveau, K. H., Harris, P. L., Ambady, N., & Sommers, S. R. (2014). Monoracial and biracial children: Effects of racial identity saliency on social learning and social preferences. *Child Development, 85*, 2299–2316. https://doi.org/10.1111/cdev.12266.
Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday though*. Oxford University Press.
Gelman, S. A., Goetz, P. J., Sarnecka, B. W., & Flukes, J. (2008). Generic language in parent-child conversations. *Language Learning and Development, 4*, 1–31. https://doi.org/10.1080/15474400701542625.
Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183–209. https://doi.org/10.1016/0010-0277(86)90034-X.
Gelman, S. A., Star, J. R., & Flukes, J. E. (2002). Children's use of generics in inductive inferences. *Journal of Cognition and Development, 3*, 179–199. https://doi.org/10.1207/S15327647JCD0302_3.
Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology, 61*, 273–301. https://doi.org/10.1016/j.cogpsych.2010.06.001.
Goldfarb, D., Lagattuta, K. H., Kramer, H. J., Kennedy, K., & Tashjian, S. M. (2017). When your kind cannot live here: How generic language and criminal sanctions shape social categorization. *Psychological Science, 28*, 1597–1609. https://doi.org/10.1177/0956797617714827.
Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and Semantics* (Volume 3, pp. 41–58). New York: Academic Press.
Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology, 53*, 310–344. https://doi.org/10.1016/j.cogpsych.2006.04.003.
Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in intergroup perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology, 12*, 392–407. https://doi.org/10.1016/S0022-1031(76)80006-6.
Heyman, G. D., & Gelman, S. A. (2000a). Preschool children's use of novel predicates to make inductive inferences about people. *Cognitive Development, 15*, 263–280. https://doi.org/10.1016/S0885-2014(00)00028-9.
Heyman, G. D., & Gelman, S. A. (2000b). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology, 77*, 1–19. https://doi.org/10.1006/jecp.1999.2555.
Hitti, A., & Killen, M. (2015). Expectations about ethnic peer group inclusivity: The role of shared interests, group norms, and stereotypes. *Child Development, 86*, 1522–1537. https://doi.org/10.1111/cdev.12393.
Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review, 117*, 1168–1187. https://doi.org/10.1037/a0020463.
Jacobs, J. E., & Potenza, M. (1991). The use of judgment heuristics to make social and object decisions: A developmental perspective. *Child Development, 62*, 166–178. https://doi.org/10.1111/j.1467-8624.1991.tb01522.x.
Johnston, K. E., & Jacobs, J. E. (2003). Children's illusory correlations: The role of attentional bias in group impression formation. *Journal of Cognition and Development, 4*, 129–160. https://doi.org/10.1207/S15327647JCD0402_01.
Kahneman, D. (2011). Thinking fast and slow. *FSG Adult.*
Kalish, C. (2012). Generalizing norms and preferences within social categories and individuals. *Developmental Psychology, 48*, 1133–1143. https://doi.org/10.1037/a0026344.
Killen, M. (2007). Children's social and moral reasoning about exclusion. *Current Directions in Psychological Science, 16*, 32–36. https://doi.org/10.1111/j.1467-8721.2007.00470.x.
Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology, 67*, 596–610. https://doi.org/10.1037/0022-3514.67.4.596.
Krueger, J., & Rothbart, M. (1990). Contrast and accentuation effects in category learning. *Journal of Personality and Social Psychology, 59*, 651–663. https://doi.org/10.1037/0022-3514.59.4.651.
Lagattuta, K. H. (2008). Young children's knowledge about the influence of thoughts on emotions in rule situations. *Developmental Science, 11*, 809–818. https://doi.org/10.1111/j.1467-7687.2008.00727.x.
Lagattuta, K. H., & Kramer, H. J. (2021). Advanced emotion understanding: Children's and adults' knowledge that minds generalize from prior emotional events. *Emotion, 21*, 1–16. https://doi.org/10.1037/emo0000694.
Lagattuta, K. H., & Sayfan, L. (2013). Not all past events are equal: Biased attention and emerging heuristics in children's past-to-future forecasting. *Child Development, 84*, 2094–2111. https://doi.org/10.1111/cdev.12082.
Lagattuta, K. H., Tashjian, S. M., & Kramer, H. J. (2018). Does the past shape anticipation for the future? Contributions of age and executive function to advanced theory of mind. *Zeitschrift für Psychologie, 226*, 122–133. https://doi.org/10.1027/2151-2604/a000328.
Lara, K. H., Lagattuta, K. H., & Kramer, H. J. (2019). Is there a downside to anticipating the upside? Children's and adults' reasoning about how prior expectations shape future emotions. *Child Development, 90*, 1170–1184.
Lawson, C. A., & Bower, C. A. (2014). Illusory correlations in preschoolers. *Cognitive Development, 31*, 22–34. https://doi.org/10.1016/j.cogdev.2014.02.005.
Leslie, S. J. (2008). Generics: Cognition and acquisition. *Philosophical Review, 117*, 1–47. https://doi.org/10.1215/00318108-2007-023.
Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology, 47*, 241–275. https://doi.org/10.1016/S0010-0285(03)00034-3.
Master, A., Markman, E. M., & Dweck, C. S. (2012). Thinking in categories or along a continuum: Consequences for children's social judgments. *Child Development, 83*, 1145–1163. https://doi.org/10.1111/j.1467-8624.2012.01774.x.
McGarty, C., & Penny, R. E. C. (1988). Categorization, accentuation and social judgment. *British Journal of Social Psychology, 27*, 147–157. https://doi.org/10.1111/j.2044-8309.1988.tb00813.x.
McGarty, C., & Turner, J. C. (1992). The effects of categorization on social judgment. *British Journal of Social Psychology, 31*, 253–268.

McGlothlin, H., & Killen, M. (2005). Children's perceptions of intergroup and intragroup similarity and the role of social experience. *Journal of Applied Developmental Psychology, 26*, 680–698. https://doi.org/10.1016/j.appdev.2005.08.008.
Moty, K., & Rhodes, M. (2021). The unintended consequences of the things we say: What generics communicate to children about unmentioned categories. *Psychological Science, 32*, 189–203. https://doi.org/10.31234/osf.io/zkjyr.
Norton, M. I., & Sommers, S. R. (2011). Whites see racism as a zero-sum game that they are losing. *Perspectives in Psychological Science, 6*, 215–218. https://doi.org/10.1177/1745691611406922.
Papfragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics—pragmatics interface. *Cognition, 86*, 253–282. https://doi.org/10.1016/S0010-0277(02)00179-8.
Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology, 42*, 1051–1068. https://doi.org/10.1037/0022-3514.42.6.1051.
Qian, M. K., Quinn, P. C., Heyman, G. D., Pascalis, O., Fu, G., & Lee, K. (2019). A long-term effect of perceptual individuation training on reducing implicit racial bias in preschool children. *Child Development, 90*, e290–e305. https://doi.org/10.1111/cdev.12971.
R Core Team (2017). R: A language and environment for statistical computing. R Foundation for statistical computing Vienna, Austria. URL https://www.R-project.org/.
Reyna, V. F., & Ellis, S. C. (1994). Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological Science, 5*, 275–279. https://doi.org/10.1111/j.1467-9280.1994.tb00625.x.
Rhodes, M., & Gelman, S. A. (2008). Categories influence predictions about individual consistency. *Child Development, 79*, 1270–1287. https://doi.org/10.1111/j.1467-8624.2008.01188.x.
Rhodes, M., Leslie, S. J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences, USA, 109*, 13526–13531. https://doi.org/10.1073/pnas.1208951109.
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*, 8. https://doi.org/10.1525/collabra.44.
Riggs, A. E., Kalish, C. W., & Alibali, M. W. (2014). When you've seen one, have you seen them all? *Children's memory for general and specific learning episodes Developmental Psychology, 50*, 1653–1659. https://doi.org/10.1037/a0036130.
Roberts, S. O., Gelman, S. A., & Ho, A. K. (2017). So it is, so it shall be: Group regularities license children's prescriptive judgments. *Cognitive Science, 41*, 576–600. https://doi.org/10.1111/cogs.12443.
Roberts, S. O., Guo, C., Ho, A. K., & Gelman, S. A. (2018). Children's descriptive-to-prescriptive tendency replicates (and varies) cross-culturally: Evidence from China. *Journal of Experimental Child Psychology, 165*, 148–160. https://doi.org/10.1016/j.jecp.2017.03.018.
Rothbart, M., Fulero, S., Jensen, C., Howard, J., & Birrell, P. (1978). From individual to group impressions: Availability heuristics in stereotype formation. *Journal of Experimental Social Psychology, 14*, 237–255. https://doi.org/10.1016/0022-1031(78)90013-6.
Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. *Advances in Experimental Social Psychology, 26*, 123–162. https://doi.org/10.1016/S0065-2601(08)60153-7.
Schwarz, N., & Bless, H. (2007). Mental construal process: The inclusion/exclusion model. In D. A. Stapel, & J. Suls (Eds.), *Assimilation and contrast in social psychology* (pp. 119–141). Philadelphia, PA: Psychology Press.
Sherman, S.J., & Corty (1984). Cognitive heuristics. In R.S. Wyer, J.R. & T.K. Srull (Eds.). Handbook of social cognition Vol. 1 (pp. 189-286). Lawrence Erlbaum Associates Publishers.
Sherman, J. W., Kruschke, J. K., Sherman, S. J., Percy, E. J., Petrocelli, J. V., & Conrey, F. R. (2009). Attentional processing in stereotype formation: A common model for category accentuation and illusory correlation. *Journal of Personality and Social Psychology, 96*, 305–323. https://doi.org/10.1037/a0013778.
Shutts, K. (2015). Young children's preferences: Gender, race, and social status. *Child Development Perspectives, 9*, 262–266. https://doi.org/10.1111/cdep.12154.
Shutts, K., Pemberton, C. K., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development, 14*, 35–62. https://doi.org/10.1080/15248372.2011.638686.
Simonsohn, U. (2015). Small telescopes: Detectability and the evaluative of replication results. *Psychological Science, 26*, 559–569. https://doi.org/10.1177/0956797614567341.
Sloutsky, V. M., Deng, W. S., Fisher, A. V., & Koos, H. (2015). Conceptual influences on induction: A case for late onset. *Cognitive Psychology, 82*, 1–31. https://doi.org/10.1016/j.cogpsych.2015.08.005.
Sunstein, C. R. (2005). *Moral heuristics. Behavioral Brain Sciences, 28*, 531–542. https://doi.org/10.1017/S0140525X05000099.
Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology, 33*, 1–39. https://doi.org/10.1146/annurev.ps.33.020182.000245.
Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgment. *British Journal of Psychology, 54*, 101–114. https://doi.org/10.1111/j.2044-8295.1963.tb00865.x.
Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232. https://doi.org/10.1016/0010-0285(73)90033-9.
Yamauchi, T., & Markman, A. B. (2000). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 776–795. https://doi.org/10.1037//0278-7393.26.3.776.
Young, S. G., & Hugenberg, K. (2010). Mere social categorization modulates identification of facial expressions of emotion. *Journal of Personality and Social Psychology, 99*, 964–977. https://doi.org/10.1037/a0020400.