

The Out-Group Homogeneity Effect Across Development: A Cross-Cultural Investigation

Reut Shilo
Bar-Ilan University

Anika Weinsdörfer and Hannes Rakoczy
University of Göttingen

Gil Diesendruck
Bar-Ilan University

The present studies investigated the out-group homogeneity effect in 5- and 8-year-old Israeli and German children ($n = 150$) and adults ($n = 96$). Participants were asked to infer whether a given property (either biological or psychological) was true of an entire group—either the participants' in-group ("Jews" or "Germans") or their out-group ("Arabs" or "Turks"). To that end, participants had to select either a homogenous or a heterogeneous sample of group members. It was found that across ages and countries, participants selected heterogeneous samples less often when inferring the biological properties of out-compared to in-group members. No effect was found regarding psychological properties. These findings have important implications for our understanding of the origins of intergroup bias.

Social psychologists have for long noticed the ease and robustness with which adults develop discriminatory attitudes and behaviors that favor groups they are part of (their "in-groups") and undermine groups they do not belong to (their "out-groups"). Recent developmental studies reveal that such intergroup biases are manifest already by early childhood (Bigler & Liben, 2007; Dunham, Baron, & Carey, 2011; Pauker, Ambady, & Apfelbaum, 2010), and perhaps even infancy (Hamlin, Mahajan, Liberman, & Wynn, 2013; Kinzler, Dupoux, & Spelke, 2007). A crucial undertaking, therefore, is trying to understand the conceptual foundations of such biases.

One such foundation is social essentialism, that is, the idea that people view social categories as reflecting objective partitions of the natural world, defined by inherent and permanent characteristics, and powerful sources for induction (Gelman, 2003). An essentialist construal of certain social categories has been found both in adults (e.g., Haslam & Whelan, 2008) and in children (e.g., Birnbaum, Deeb,

Segall, Ben-Eliyahu, & Diesendruck, 2010; Rhodes & Gelman, 2009). Indeed, a relation between essentialist beliefs and attitudes toward others has been documented both in adults (Keller, 2005; Prentice & Miller, 2007; Williams & Eberhardt, 2008), and children. For instance, Rhodes, Leslie, Saunders, Dunham, and Cimpian (2017) found that experimentally inducing essentialist beliefs about a novel social category led 4.5- to 6-year-olds to share fewer resources with out-group members; and Diesendruck and Menahem (2015) found that experimentally reinforcing an essentialist construal of ethnicity led Israeli Jewish 6-year-olds to draw a Jewish and an Arab character farther apart from each other.

A further—and related—conceptual process postulated to underlie intergroup biases is what social psychologists termed the "out-group homogeneity effect" (Judd & Park, 1988; Simon & Brown, 1987). This effect describes the different conceptualizations of in-group and out-group members that adults hold, whereby in-group members are construed more as unique individuals and out-group members more as homogeneous category exemplars. This effect has been found with a variety of tasks (Boldry, Gaertner, & Quinn, 2007), such as: (a) in-

First authorship is shared by the first two authors.

This research was supported by a Grant from the GIF, the German-Israeli Foundation for Scientific Research and Development, to Gil Diesendruck and Hannes Rakoczy. We want to thank the teachers, parents, and children for their participation.

Correspondence concerning this article should be addressed to Reut Shilo, Department of Psychology, Bar-Ilan University, Ramat-Gan 52900, Israel. Electronic mail may be sent to reutshilo@gmail.com.

group members are viewed as more widely distributed than out-group members across various trait dimensions (Boldry & Gaertner, 2006; Park & Judd, 1990; Simon & Brown, 1987); (b) more subgroups are listed as potentially existing within in-groups than within out-groups (Linville, Fischer, & Yoon, 1996; Park, Ryan, & Judd, 1992); (c) a higher percentage of out-group than of in-group members are thought to possess stereotypic traits (Pickett & Brewer, 2001; Ryan, Judd, & Park, 1996); and (d) better recall of which individual in-group member, compared to an out-group member, is associated with a particular piece of information (Lorenzi-Cioldi, 1993; Ostrom, Carpenter, Sedikides, & Li, 1993; Stewart & Vassar, 2000).

Importantly, this continuum from viewing others *qua* individuals versus *qua* category exemplars has been extensively discussed in the social psychological literature on adults, especially in regard to its pernicious implications (Brewer, 1988; Fiske & Neuberg, 1990; Hugenberg, Young, Bernstein, & Sacco, 2010; Ostrom et al., 1993). For instance, there is a close relationship between perceived variability and stereotype endorsement (Hewstone & Hamberger, 2000; Park & Hastie, 1987; Ryan et al., 1996), and between construing others as homogeneous category exemplars and dehumanization (for a review see Haslam, 2006). In line with these findings, it has been found that increasing perceived group variability led to a reduction in prejudice and discrimination (Brauer & Er-rafiiy, 2011). In sum, the different conceptualizations of in- and out-groups play a key role in the generation and maintenance of intergroup biases.

From a developmental perspective, there has been some work showing that 5-year-olds view out-group members defined by skin-color as *visually* more alike than in-group members (Aboud, 2003). However, whether children have biased beliefs about fundamental characteristics of in-group and out-group members has rarely been addressed directly, and the evidence that does exist is inconclusive. In particular, Guinote, Mouro, Pereira, and Monteiro (2007) asked Black and White 7- to 9-year-olds to distribute visual targets (schematic drawings of faces, either White or Black) along five levels of four dimensions: ugly–pretty, not nice–nice, poor–rich, and reads badly–reads well. Guinote et al. (2007) found that White children perceived the out-group in a more homogeneous way than the in-group. In contrast, Black children perceived the out-group in a more variable way than the in-group, that is, the opposite of an out-group homogeneity bias. A similar reversal was

documented in Birnbaum et al.'s (2010) study among Israeli children. There, they found that Arab children—part of a minority in Israel—were more likely to draw inferences based on the ethnicity of characters when the characters were Arabs than when they were Jewish, whereas Jewish children's inferential patterns were equivalent regarding Jewish and Arab characters. Given the scant and inconclusive findings on children's conceptual out-group homogeneity effect, and the association of a homogenization of out-groups with negative attitudes toward these groups, the present studies set out to investigate the out-group homogeneity effect in children directly and systematically.

To this end, we assessed children's variability judgments of in- and out-group members through their inductive reasoning strategies. We reasoned that if children assume out-group members to be more similar to one another than in-group members, then the out-group as a whole should be covered by a narrower sample than the in-group. In other words, children should make inductive inferences about the whole group from more restricted information about individual out-group than in-group members. Following research on children's statistical and inductive reasoning (Kushnir, Xu, & Wellman, 2010), we tested this hypothesis by investigating participants' sampling strategies in induction. We selected diversity-based sampling as a means to assess children's homogeneity bias because it has been successfully used with young children, both in regard to animal (Rhodes & Brickman, 2010), and human categories (Noyes & Christie, 2016).

Adults have been shown to take the diversity in samples of evidence into account when drawing inductive inferences (Heit, 2000; Osherson, Smith, Wilkie, López, & Shafir, 1990). Studies with children indicated that the tendency to prefer diverse samples over homogeneous ones in order to draw broad generalizations appears only around 9 years of age (Gutheil & Gelman, 1997; Li, Cao, Li, Li, & Deák, 2009; Lopez, Gelman, Gutheil, & Smith, 1992; Rhodes, Gelman, & Brickman, 2008). Moreover, children's consideration of sample diversity seems to depend on their beliefs about how variable entities of a given domain are with respect to particular target properties. For instance, 7-year-olds could implement diversity-based reasoning when primed to focus on within-category heterogeneity (Rhodes & Brickman, 2010), 5- and 6-year-olds favored diverse samples for generalizations when reasoning about novel animal categories but not about familiar ones (Rhodes & Liebenson, 2015), and 5- to 7-

year-olds preferred diverse samples when having to infer about people's toy preference but not about people's hormones (Noyes & Christie, 2016).

The present studies employed a similar methodology to assess the potentially different conceptualizations of in- and out-group members. In two national contexts (Israel and Germany), children from two different age groups (Study 1) as well as young adults (Study 2) were asked to determine whether a given property (either biological or psychological) was true of the entire group—either the participants' in-group or their out-group. To decide whether the property was true of the whole group, participants could choose one of two samples. Each of the samples comprised three individuals from the group in question but differed in terms of diversity, resulting in one homogeneous sample and one heterogeneous sample. We reasoned that if participants viewed out-groups as more homogeneous than in-groups, they should be more inclined to select the homogeneous sample in questions targeting the out-group compared to questions targeting the in-group.

We assessed inferences regarding both biological and psychological properties. The biological properties referred to biological processes and features within the human body; the psychological properties referred to mental processes as well as patterns of behavior. We chose to investigate these two domains for a number of reasons. First, given that these are all internal properties, it is unlikely that children would have had direct experience with them, especially if children's overall degree of contact with out-group members is scant—which, as we note below, was the case. In other words, it was implausible that children's inferences regarding the distribution of such properties would have been driven by direct observation and instead would more likely reflect children's intuitive beliefs about the groups given. Second, we had theoretical reasons to study whether children's beliefs about these two types of properties would differ. Specifically, there is a long-standing theoretical debate as to whether children and adults conceive of "human kinds" in analogy to natural ones (Atran, 1996; Haslam & Whelan, 2008) or according to a "naïve sociology" (Hirschfeld, 1996). In the former case, biological properties might be viewed as proxies for essences, thus distinguishing between in- and out-groups. In the latter case, in turn, psychological properties might be those viewed as core in distinguishing between in- and out-groups (see also Diesendruck & Eldror, 2011). A recent developmental study by Noyes and Christie (2016), indeed found a

difference for these two property domains. Namely, whereas children did not privilege a diverse sample when asked to draw inferences about people's hormones (a biological property), they did so when asked about people's toy preferences (a psychological property). Thus, it was of interest to assess whether children's sampling strategies would vary according to property type.

We investigated the out-group homogeneity effect in Israel and Germany because both share important similarities but also differ in at least two important and interesting respects. Both countries are constituted by a majority that holds Western secular values, both are also characterized by a significant Muslim minority (the majority of Arab Israelis, and Germans with Turkish origins, respectively), and in both, majority children have little contact with the given minority populations. However, Israel and Germany also differ in important ways: In Israel, as opposed to Germany, the majority-minority relation has been shaped by an enduring territorial conflict, rendered by narratives of threats to national survival. A second, related difference is the salience of the minority group in both societies—arguably much higher in Israel than in Germany. Thus, comparing these two countries allowed us to estimate the extent to which the above differences impact the emergence of out-group homogenization already early in development.

Last, we chose to investigate three different age groups: 5- and 8-year-olds, as well as young adults. The children's age groups were selected for several reasons. First, there is robust evidence that by age 5–6, children hold essentialist beliefs about various social categories (Birnbaum et al., 2010; Deeb, Segall, Birnbaum, Ben-Eliyahu, & Diesendruck, 2011; Kinzler & Dautel, 2012; Rhodes & Gelman, 2009). Second, there is also evidence that by age 5–6, children manifest robust implicit and explicit intergroup biases (Benozio & Diesendruck, 2015; Bigler & Liben, 2007; Buttelmann & Böhm, 2014; Dunham et al., 2011). Third, a meta-analysis by Raabe and Beelmann (2011) on prejudice endorsement in different age groups suggests a systematic age difference between middle childhood (5- to 7-year-olds, with a peak in prejudice) and late childhood (8- to 10-year-olds, with a slight decrease in prejudice). Fourth, as mentioned earlier, whereas children in late childhood have been identified to be rather competent in diversity-based inductive reasoning, this competence is much less robust among 5-year-olds (Noyes & Christie, 2016; Rhodes & Brickman, 2010; Rhodes & Liebenson, 2015).

Fifth, we were interested in assessing the extent to which the hypothesized alternative conceptualizations of groups affect children's reasoning about in- and out-group members even prior to their entrance to schools—arguably a major enculturation factor. If even 5-year-olds evinced a homogenization of out-group members, that would support the idea that such differential conceptualization of in- and out-groups may be part of children's intuitive understanding of the social world, and stress the importance of early interventions to prevent inter-group biases. Finally, given the novelty of the method chosen to assess the homogenization of out-groups, we included an adult group so as to ascertain that the phenomenon we were assessing in children, was analogous to the one found in adults.

Hence, Study 1 examined how 5- and 8-year-olds from the majority group in both countries (Jews, or Germans with German origins, in Israel and Germany, respectively) perceive their own group as compared to the chosen minority group (Arabs, or Germans with Turkish origins, respectively), and Study 2 investigated this question in young adults.

Study 1

Method

Participants

In Israel, participants included 36 children attending kindergartens (33% female; $M_{\text{age}} = 5.6$ years, range = 4.7–6.5 years) and 35 children attending elementary school (63% female; $M_{\text{age}} = 7.9$ years, range = 7.5–8.4). All participants were middle-class secular Jews, recruited from kindergartens and schools in central cities in Israel where Jews are the absolute majority (according to the Israeli Central Bureau of Statistics, 2015, the percentage of Jews in the cities from which children were sampled was 99.9%). Parents received letters containing a consent form; only children with signed consent forms participated in the study. Children received a small reward for their participation. The study was conducted in the children's educational institutions. Data were collected between May 2015 and June 2016.

In Germany, participants included 40 children attending kindergartens (50% female; $M_{\text{age}} = 5.4$, range = 5.0–5.9) and 39 children attending elementary school (48% female; $M_{\text{age}} = 8.2$, range = 7.5–8.9). All children were German native speakers, came from mixed socioeconomic backgrounds and

were recruited via urban kindergartens as well as via a local database of parents who had previously given consent to participate in developmental studies (according to the Department for Statistics and Elections of the City of Göttingen [GÖSIS], in 2016, *Germans without migrant background* constituted 76.8% of Göttingen's population, and *Germans with Turkish migrant background* 1.2%). Part of the study was run in children's kindergartens or schools, part in a laboratory at the University of Göttingen. Children who participated in the laboratory received a small reward. Data were collected between October and December 2016.

Design

The study included two between-subjects variables: country (Israel and Germany) and age group (5- and 8-year-olds). The experimental variables were manipulated within subjects. They were as follows: group membership of the targets (in-group-IG/out-group-OG) and property type (biological/psychological), with two trials of each of the four trial types (i.e., in-group and biological property, in-group and psychological property, out-group and biological property, out-group and psychological property). The dependent variable was the number of heterogeneous samples chosen (0–2 per trial-type, or 0–8 aggregated over all trial-types).

Materials

Children were presented with a PowerPoint Presentation by the experimenter. Each of the eight questions was presented on one slide, followed by a slide depicting the samples children could choose from. The slides had a white background, and were divided in half by a bold black line. On each side, there were three pictures, either of three cartoon characters or of photographs of different locations (see Figure 1). To highlight the separation between the samples, there was a green and a blue box surrounding the three pictures on each side (the sides of the colors were counterbalanced).

Procedure

The experimenter sat with each child separately, and presented a forced-choice sampling-for-induction task similar to that used by Rhodes and Brickman (2010). The task was presented as a game, in which children were asked to pretend that they were scientists “who wanted to find out stuff about



A Question About Outgroup and Biological Property in Israel:

"Imagine That You Are a Scientist and You Want to Find Out Whether Arabs Have 12 Bones in Their Hand.

Which Group Would You Choose in Order to Answer That Question? The Green Group, Which Has Three Arabs Who Live in Jerusalem, or Would You Choose the Blue Group, Which Has One Arab Who Lives in Haifa, One Arab Who Lives in Jerusalem, and One Arab Who Lives in Tel-Aviv?"

A Question About Ingroup and Psychological Property in Germany:

"Imagine That You Are a Scientist and You Want to Find Out Whether Germans Believe That One Can Make a Wish When Seeing a Falling Star.

Which Group Would You Choose in Order to Answer That Question? The Blue Group, Which Has One German Who Is a Teacher, One German Who is a Doctor, and One German Who Is a Policewoman, or Would You Choose the Green Group, Which Has Three Germans Who Are Teachers?"

Figure 1. Examples of items used in Study 1.

people." They were then presented the series of eight questions, with half of them being about the in-group (i.e., four questions, "Jews" in Israel, and "Germans" in Germany) and half about the out-group (i.e., four questions, "Arabs" in Israel, and "Turks" in Germany). Half of the questions were about biological properties (e.g., "Do [in-group] or [out-group members] have blood type F?"), and half about psychological properties (e.g., "Do [in-group] or [out-group members] prefer cucumbers over tomatoes?"). The conventional terms used to refer to the in-group ("Jews"/"Germans") and the out-group ("Arabs"/"Turks") in the two countries were used instead of the terms denoted in the square-brackets. Thus, each child received four types of trials with two tokens of each type. The questions were counterbalanced across participants in all parameters, including order of presentation.

For each question, children had to choose one of the two samples that they thought would help them answer the "research question"—either the homogeneous sample (e.g., three in-group or out-group members who are poor) or the

heterogeneous sample (e.g., one in-group or out-group member who is rich, one in-group or out-group member who is average, and one in-group or out-group member who is poor). So for instance, in one trial, the child saw a white screen, and on it appeared a question in writing that the experimenter read to the child: "Imagine that you are a scientist and you want to find out whether Arabs have 12 bones in their hand." In this example, the question appeared in Israel and was about out-group biological property. Then, as can be seen in Figure 1, the child saw the next slide, in which the two types of samples were displayed. In order to make the distinction between the two types of samples easier for children, each was depicted within a separate and colored frame—one green and one blue—and the experimenter referred to the two samples by mentioning their frame color. For instance, the experimenter asked:

Which group would you choose in order to answer that question? The green group, which has three Arabs who live in Jerusalem, or would

you choose the blue group, which has one Arab who lives in Haifa, one Arab who lives in Jerusalem, and one Arab who lives in Tel-Aviv?

The child could either answer verbally or point at the chosen sample. Importantly, both colors referred to the same group membership in a given question: either the in-group or the out-group. After children chose one of the two samples, the experimenter showed a new slide, and presented to children the next “research question” (see Figure 1, Table 1 for a full list of the properties, and Table S1 for a full list of the different samples.)

Results and Discussion

The mean numbers of trials in which children chose the heterogeneous sample as a function of the four conditions are depicted in Figure 2. Analyses were conducted using a repeated-measures analysis of variance (ANOVA), with country, age group, and gender as between-subjects variables, and group membership and property type as within-subjects variables. Gender was included in the analysis since previous research found that gender may play a role in intergroup biases (Benozio & Diesendruck, 2015). Overall, there was a significant effect of age group, such that 8-year-olds selected more heterogeneous samples ($M = 1.40$, $SD = 0.45$) than 5-year-olds ($M = 1.18$, $SD = 0.47$), $F(1, 142) = 8.45$, $p = .004$, $\eta_p^2 = .06$. This is consistent with the finding that the strategy to select more diverse samples for drawing inferences becomes more common with age (Lopez et al., 1992; Rhodes, Gelman, et al., 2008). Interestingly, there was no main effect of country, $F(1, 142) = 1.46$, $p = .23$, $\eta_p^2 = .01$, or gender, $F(1, 142) = 0.65$, $p = .42$, $\eta_p^2 = .01$. There was a significant three-way interaction among gender, age group, and country. To identify the source of this interaction, we split the data by gender, and

analyzed the interaction between country and age group. These analyses revealed a significant interaction between country and age group among boys, $F(1, 73) = 4.81$, $p = .03$, $\eta_p^2 = .06$, such that in Israel there was an increase in the sum of heterogeneous sample choices from kindergarten ($M = 3.96$, $SD = 1.99$) to elementary school ($M = 5.77$, $SD = 1.64$), whereas in Germany there was not (kindergarten: $M = 5.25$, $SD = 1.71$; elementary school: $M = 5.15$, $SD = 1.98$). Among girls, the two-way interaction between country and age group was not significant, $F(1, 69) = 1.04$, $p = .31$, $\eta_p^2 = .02$.

Most importantly, the overall repeated-measures ANOVA—that is, without the split by gender—revealed an interaction between group membership and property type, $F(1, 142) = 6.01$, $p = .015$, $\eta_p^2 = .04$. Namely, when asked about psychological properties, there was no difference in sampling strategies for out-group ($M = 1.31$, $SD = 0.70$) and in-group ($M = 1.25$, $SD = 0.72$), paired- $t(149) = -0.92$, $p = .36$, $d = 0.08$. In turn, when asked about biological properties, children chose the heterogeneous sample less often in questions about out-groups ($M = 1.18$, $SD = 0.71$) as compared to questions about in-groups ($M = 1.37$, $SD = 0.69$), paired- $t(149) = 2.63$, $p = .009$, $d = 0.22$.

To ascertain that the above reported parametric differences did not derive from particular individuals, we conducted Wilcoxon signed-rank tests assessing the distribution of children vis-à-vis their pattern of responses on biological and psychological properties for in-groups and out-groups. These analyses corroborated the parametric analysis. When asked about psychological properties, the distribution of children choosing heterogeneous or homogeneous samples did not differ for out-groups and in-groups, $T = 496$, $p = .348$, $r = -.054$. In turn, when asked about biological properties, children chose the heterogeneous sample less often for out-groups than for in-groups, $T = 197$, $p = .01$, $r = -.15$.

In a final analysis, we examined whether children in both age groups are capable of diversity-based sampling. To this end, we looked separately at 5- and 8-year-olds’ overall (out of eight trials) tendency to choose heterogeneous samples, and compared those to chance performance (chance = 4). We found that both 5- and 8-year-olds chose heterogeneous samples more often than would be expected by chance, 5-year-olds: $M = 4.66$, $SD = 1.87$, $t(75) = 3.06$, $p = .003$, $d = 0.35$; 8-year-olds: $M = 5.57$, $SD = 1.81$, $t(73) = 7.43$, $p < .001$, $d = 0.86$.

Table 1
List of Properties Used in Study 1

Biological	Psychological
Have 12 bones in their hand	<i>Israel</i> : believe that a butterfly that enters the house brings good luck <i>Germany</i> : believe that one can make a wish when seeing a falling star
Have a part in their head called “Amygdala”	Want to be doctors the most when they grow up
Have blood type F	Believe there are aliens
Have three parts in their hearts	Prefer cucumber over tomato

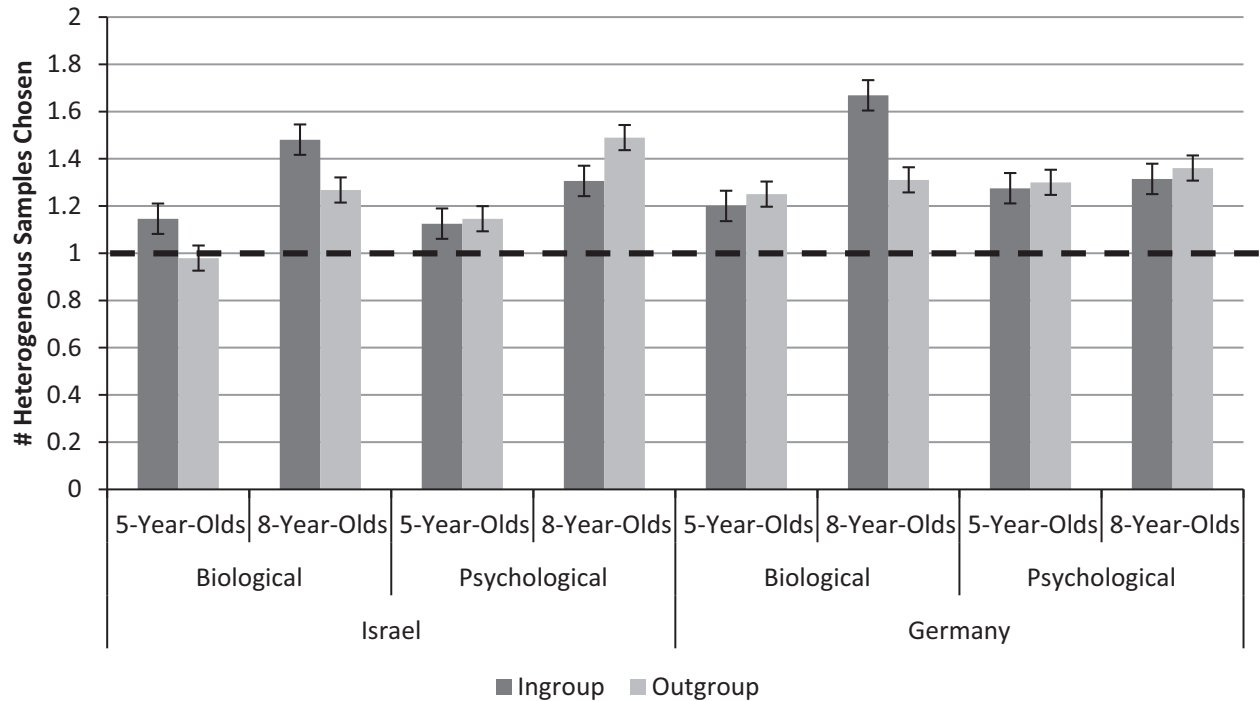


Figure 2. Mean numbers of heterogeneous samples chosen by children (range = 0–2), in Study 1.

Note. Sample choices depicted as a function of the between-subjects factors country (Israel/Germany) and age group (5-/8-year-olds), as well as the within-subjects factors property type (biological/psychological) and group membership (in-group/out-group). Error bars indicate standard errors. The dashed line marks chance level.

In sum, we see that 5- and 8-year-olds are able to utilize a diversity-based sampling strategy in some cases. More importantly, we found that regarding biological properties, but not psychological ones, children deploy this strategy more often when drawing inferences about in-groups than about out-groups, and that they do so even before entering elementary school. In order to widen the developmental perspective, Study 2 examined whether this intergroup bias is present in adulthood.

Study 2

Method

Participants

In Israel, 56 Bar-Ilan University students (64% female; $M_{\text{age}} = 24.4$ years, range = 19.8–34.5) participated between June and November 2016. Participants received either class credit or payment (approximately 3 Euros) for their participation. In Germany, 40 students of the University of Göttingen were recruited (50% female; $M_{\text{age}} = 24.8$ years, range = 20–33) in January 2017. As an expense allowance, participants received a bar of chocolate.

All participants signed a consent form prior to participation.

Design

Tested in the same procedure as children, a pilot study on adults yielded a ceiling effect with adults selecting the heterogeneous samples exclusively. Therefore, the adults' design was modified in two ways. First, to decrease the possibility of adults answering based on social desirability, we manipulated group membership between subjects, with each participant receiving questions either about the in-group ("Jews" or "Germans") or about the out-group ("Arabs" or "Turks"). Second, we introduced a new within-subjects factor, the variable of "cost," such that on half of the trials (8) choosing the heterogeneous sample was more expensive than choosing the homogenous sample, and on the other half (8) vice versa. Similar to the children, adults were asked to imagine themselves as new researchers in a university, who had an undefined limited "research budget" that they should use most efficiently for trying to find the answers to a few "research questions." Then, in half of the trials the homogenous sample was presented as more

expensive, whereas in the other half the heterogeneous sample was declared as being more expensive. We hoped that this new factor would require adult participants to be more judicious in their decisions about which sample to choose. To make all trials most comparable, we kept the difference between the expensive and the cheap sample constant across all trials, namely, US\$130 or 130€, in the Israel and German samples, respectively. For example, in one trial the heterogeneous sample's price was 750 (US\$ or €, depending on country) and the homogenous sample's price was 620, and in another trial the heterogeneous sample's price was 160 and the homogenous sample's price was 290. Prices ranged between 150–950. As in the children's task, property type varied within subjects. Thus, each participant received four types of trials with four tokens each. In the resulting total of 16 questions, half of the questions were about psychological properties (e.g., "Do in-group or out-group members spend more than two and a half hours a day using their smartphones?") and half about biological properties (e.g., "Do in-group or out-group members tend to develop gluten intolerance?", see Figure 3 for an example, and Table 2 for a list of all properties). The questions were counterbalanced across participants in all parameters, including order of presentation. The dependent variable was

the number of heterogeneous sample choices the participant made (0–8 for each group membership by property type trial, 0–16 in total, see Table S2 for a full list of the different samples).

Materials

The PowerPoint slides developed for adults had the same format as those used in Study 1. Each "research question" was followed by a test slide depicting two samples of three group members each, differing in terms of diversity, and now also in terms of costs. As for the instructions, there was a difference between countries: in Israel, the instructions were given orally, whereas in Germany, the instructions appeared in writing at the beginning of the PowerPoint Presentation.

Procedure

In Israel, the instructions and the coding of answers were done face to face, so that the experimenter sat with the participant throughout the entire study. Before the beginning of the study, participants filled in a paper-pencil consent form. In Germany, participants were recruited at the campus of the University of Göttingen and brought to a quiet room with six laptops installed. After being greeted by the experimenter, the participants were seated next to one laptop each and asked to fill in a paper-pencil consent form before starting to read all task-related instructions on the computer by themselves. The instructions were given in the same PowerPoint presentation as the test slides. To ensure participants could navigate only in the intended direction, a script written in the programming language Python was activated simultaneously to the PowerPoint presentation. The same script automatically recorded participants' sample choices.

As a cover story, adults were asked to pretend that they are new researchers at a university. To be "good researchers," they were told they had to answer as many "research questions" as possible before running out of budget, while at the same time, they had to conduct high quality studies by receiving correct answers. The participants were then presented with the 16 "research questions."

Results and Discussion

The mean numbers of trials in which participants chose the heterogeneous sample as a function of conditions are depicted in Figure 4. Analyses were

Table 2
List of Properties Used in Study 2

Biological	Psychological
Lack of vitamin A60	Believe that a widow at a wedding brings bad luck
Be at increased risk of colorectal cancer	<i>In Israel:</i> were fans of Brazil in the soccer World-Cup <i>In Germany:</i> were fans of Spain during the soccer World-Cup in South Africa
Have the gen APRST	Believe in global warming
Tend to develop gluten intolerance	Their favorite European (Israel) or Italian (Germany) food is pizza
Have a tendency to high blood pressure	Spend more than 2.5 hr on their smartphones
Have a pH level of 7.6	Donate to charity more than three times a year
Have a large amount of melanin in their eye	Have a tendency to develop major depression
Compared to other OECD-countries, have a small amount of calcium in their bones	What the Happy Index among group members is

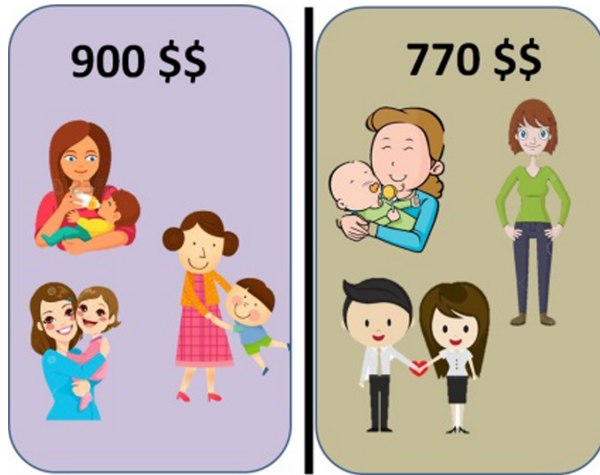


Figure 3. Example of item used in Study 2.
 Note. In this item, the homogenous sample (three in-group/out-group mothers) costs more than the heterogeneous one (an in-group/out-group mother, a single in-group/out-group woman, and an in-group/out-group woman in a relationship).

main effects or interactions involving only the between-subjects variables. There was a main effect of cost, $F(1, 88) = 85.76, p < .001, \eta_p^2 = .49$, with adults selecting more often the cheaper sample ($M = 7.00, SD = 1.26$) than the expensive one ($M = 5.30, SD = 1.62$). This indicates that the cost-factor did work as the intended counterweight against an exclusive choice of heterogeneous samples.

Crucially, the analyses revealed the same interaction between group membership and property type as the one found in Study 1, $F(1, 88) = 5.03, p = .03, \eta_p^2 = .05$. As in Study 1, when asked about psychological properties, there was no difference in sampling strategies for in-groups ($M = 6.30, SD = 1.12$) and out-groups ($M = 6.19, SD = 1.66$), $t(94) = 0.29, p = .77, d = 0.06$. In turn, when asked about biological properties, sampling strategies differed between in-groups and out-groups, with adults more often selecting heterogeneous samples when targets were in-groups ($M = 6.44, SD = 1.13$) than when they were out-groups ($M = 5.69, SD = 1.57$), $t(94) = 2.68, p = .009, d = 0.55$ (see Figure 4). Analyses against chance-responding revealed that both overall and after breaking down the analysis by the crucial variables (group membership, cost, property type) adults consistently chose heterogeneous samples

conducted using a repeated-measure ANOVA, with country, group membership, and gender as between-subjects variables, and property type and cost as within-subjects variables. There were no

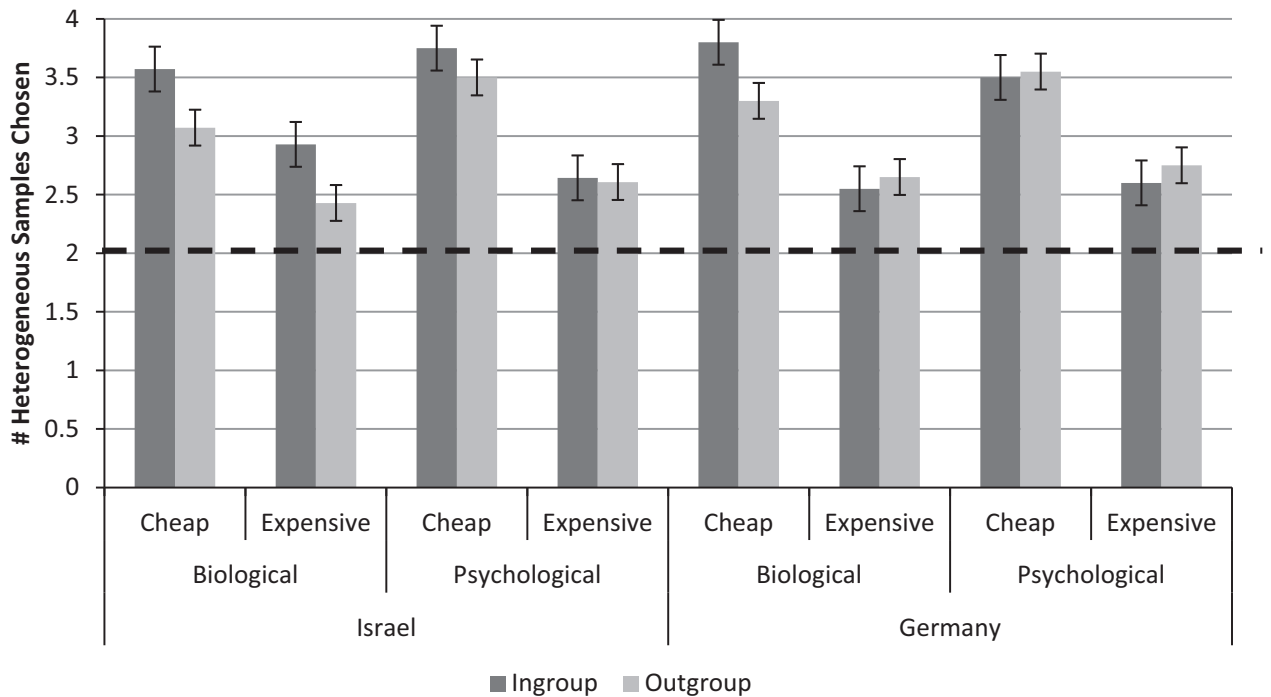


Figure 4. Mean numbers of heterogeneous samples chosen by adults (range = 0–4), in Study 2.
 Note. Sample choices depicted as a function of the between-subjects factors country (Israel/Germany) and group membership (in-group/out-group) as well as the within-subjects factors property type (biological/psychological) and costs (heterogeneous sample being cheap/expensive). Error bars indicate standard errors. The dashed line marks chance level.

more often than would be expected by chance: $M = 12.29$, $SD = 2.30$, $t(95) = 18.30$, $p < .001$, $d = 1.88$.

In sum, even though adults evinced the general tendency to prefer diverse samples for making generalizations across instances (Heit, 2000; Osherson et al., 1990), they nonetheless manifested a similar sensitivity to that revealed by children regarding group membership and property type. Namely, adults too were more likely to select heterogeneous samples for drawing inferences about in-groups than about out-groups regarding biological properties, but not regarding psychological properties, highlighting the developmental stability of this bias.

General Discussion

The present studies investigated out-group homogenization across development, a cognitive tendency thought to underlie intergroup biases. We tested whether children (Study 1) and young adults (Study 2) in two different national contexts (Israel and Germany) conceived of out-group members as more similar to one another than in-group members. To this end, we compared participants' sampling strategies when inductively reasoning about in- and out-group members' biological and psychological properties.

The main findings were the following: First, subjects tended to sample less diversely when reasoning about out- compared to in-group members, thus revealing a basic out-group homogeneity bias. Interestingly, however, this sampling pattern varied as a function of property domain. Namely, the pattern differed when reasoning about in- and out-group members' biological, but not psychological, properties. Second, this pattern was found in much the same ways both in kindergarten children, children of elementary school age, and in young adults. Finally, we did not find substantial differences between the Israeli and the German samples in any of the age groups.

This pattern of findings is interesting in several respects. With regard to children's inductive reasoning competencies, the findings suggest that even before school entrance, children are capable of diversity-sensitive sampling. Much prior developmental research had failed to find evidence for such sampling competence at this young age (e.g., Rhodes, Brickman, & Gelman, 2008; Rhodes, Gelman, et al., 2008). This study, in contrast, converges with another recent study (Noyes & Christie, 2016), in suggesting that under suitable circumstances,

even 5-year-olds can take sample diversity into account when confronted with inductive problems. More systematic future research is needed to investigate under which conditions children do or do not translate this general competence into successful performance.

The differences in subjects' sampling strategies between psychological and biological properties may be a window into underlying processes of social categorization. In particular, the finding that group membership did not affect sample choices in any of the age groups, in both national contexts, when participants reasoned about psychological properties, extends Noyes and Christie's (2016) findings regarding toy preferences. This result is consistent with the finding that children expect people's preferences to be malleable (Kalish, 2002) and that even infants are not surprised when preferences are not shared across individuals (Henderson & Woodward, 2012). In general, these findings suggest that from a young age on, children seem ready to entertain the possibility that people differ in their mental states and behaviors.

When asked about biological properties, in contrast, children and adults in both countries chose homogeneous samples significantly more often for out- as compared to in-groups. For instance, children were less likely to choose a diverse sample of out-group members than of in-group members to determine whether all group members have 12 bones in their hands. This finding suggests that children expect out-group members to be more similar in their biological constitution and, thus, more broadly, that they are more prone to construe such social groups as natural kinds (Atran, 1996; Rhodes, 2013). This is an important datum informing the debate as to whether children's social essentialism derives from their biological essentialism, from a modular naïve sociology, or from more general cognitive processes (cf. Atran, 1996; Cimpian & Salomon, 2014; Diesendruck & Eldror, 2011; Gelman, 2003; Gil-White, 2001; Hirschfeld, 1996).

Last, the finding that similar biased sampling patterns were found across age groups and cultures has potentially interesting implications concerning the role of cultural socialization in the development of intergroup biases. In both Israel and Germany, participants belonged to a majority characterized by Western secular values, and we assessed their reasoning about members of a Muslim minority. Although both countries clearly do differ in the degree of intergroup tensions, and thus also in terms of the everyday salience of the given out-groups, we found the same pattern of results in our

Israeli and German samples. In the Israeli sample, this pattern is in line with prior developmental findings indicating an essentialist construal of ethnic groups (Birnbaum et al., 2010; Deeb et al., 2011). The fact that the same pattern was found in Germany for an out-group based on nationality may suggest a similar construal of national groups—a type of social category that, so far, has rarely received attention in research on social essentialism (Haslam & Whelan, 2008). Thus, although the specific social categories that are homogenized vary between the cultures, the readiness to homogenize at least *some* out-groups on *some* property dimensions, seems to be structurally similar across these cultures. This is consistent with more general findings on the prevalence of social essentialism across cultures (Diesendruck, 2013), highlighting the importance of identifying the factors in children's cultural environments that prompt the construal of particular social categories in these ways.

In this regard, one important factor that has been suggested as contributing to children's group concepts and attitudes is language. In particular, Rhodes, Leslie, and Tworek (2012) have highlighted the importance of generic language (e.g., in our case: "Arabs are X") in the cultural transmission of social essentialism. In a series of studies, they found that hearing generic language about a novel social category led both preschool-age children and adults to develop essentialist beliefs about the category and that in children this effect persisted over time. Furthermore, they found that experimentally inducing parents to hold essentialist beliefs about a novel social category led them to produce more generic language when discussing the category with their children. In a related study, Segall, Birnbaum, Deeb, and Diesendruck (2015) examined the extent to which parents' attitudes toward an ethnic out-group, and the language that they used to tell a novel story about in- and out-group members to their child (with parameters that included: valence, stereotypes, contrasts, labeling, generics), contributed to the child's essentialist beliefs about ethnicity. They found that parents' speech—particularly the frequency of labeling and generics—made a significant contribution to children's ethnic essentialism, even more so than parents' own attitudes and beliefs, or their endorsement of ethnic stereotypes. Evidently much more work on this matter is welcome.

In general, future research needs to address whether the found intergroup bias in the tendency to homogenize groups in terms of their biological properties is driven by a general contrast between

in- and out-groups or rather by more local contrasts applied to the specific groups investigated here. The latter interpretation cannot be ruled out but seems not too plausible in light of the lack of cross-cultural differences. The selected groups in Israel and Germany differ in important ways (e.g., degree of conflict between groups, and relative salience), both of which give reason to believe that the present findings reflect a more general pattern. However, future studies could investigate these questions in more stringent ways, independent from the contingencies of particular real groups, for instance by implementing artificial and controlled minimal group paradigms.

A further related question to be addressed in future studies is whether the present findings indeed reflect an *out-group* homogeneity bias rather than a *minority* homogeneity bias. The participants tested in Israel and Germany both came from local majority groups and thus the out-groups were identified with minority groups. Thus, the bias we found could reflect either one. Prior research on the perception of group variability as a function of group status has indeed documented a minority homogenization bias. It has been found, for example, that both majority (e.g., White) and minority (e.g., Black) children perceived faces of minority children in a more homogenizing way than faces of in-group members (Guinote et al., 2007). Similarly, Birnbaum et al. (2010) found that if minority members (in that case, Arab children) had been the target group in an induction task, both majority and minority children were more likely to draw inferences based on the ethnicity of the characters, thus implying similarity between minority but not between majority members. These findings align with research on adults' perception of group variability, in which the moderating role of group status, power, and group size has been addressed (Lorenzi-Cioldi, 1993; Simon & Brown, 1987). Future research will therefore need to systematically compare majority and minority children in their perception of the respective in- and out-groups.

A further important question for future studies concerns the moderating role of property domains. In two sets of items, developed separately for children and adults, we found that perceptions of out-group variability differed between biological and psychological properties. As mentioned earlier, this finding could be taken to reflect a naturalistic construal of the given out-groups. However, it is also worth noting that the biological items mostly referred to properties invariant between and within

individuals (e.g., the existence of a brain region called Amygdala), whereas the psychological items described more fluctuating properties (e.g., one's favorite food). It is plausible that this difference in variability is not merely an artifact of the present items but a fair representation of prototypical exemplars of the two domains. Nonetheless, it could be valuable to more systematically investigate this issue. Moreover, a comparison of several property domains might be of interest to better understand which kinds of properties are subject to homogeneity biases between groups. The above qualifications notwithstanding, our studies yielded a stable pattern, of children and adults, in both countries, settling for less diverse samples *only* when drawing inferences about *out-groups'* biological properties.

A final open question pertains to the direction of the bias in sampling in- versus out-group members. It remains unclear whether the bias in terms of biological group variability reflects a homogenization of out-group members or rather a heterogenization of in-group members. Future studies will need to address this issue more comprehensively by systematically comparing inductive sampling patterns for a larger set of different social categories, including those for which questions of participants' own group membership do not even arise and thus notions of in- versus out-group do not apply.

In summary, the present studies found that children and young adults in Israel and in Germany perceived out-group members as biologically less diverse than in-group members. Studies in adults have shown that such out-group homogenization may lead to various pernicious consequences, such as stereotyping, prejudice, and dehumanization. The finding that this construal emerges so early on in development alerts to the need to intervene even before children get into schools.

References

- Aboud, F. E. (2003). The formation of in-group favoritism and out-group prejudice in young children: Are they distinct attitudes? *Developmental Psychology, 39*(1), 48–60. <https://doi.org/10.1037/0012-1649.39.1.48>
- Atran, S. (1996). *Cognitive foundations of natural history: Towards an anthropology of science*. Cambridge, UK: Cambridge University Press.
- Benozio, A., & Diesendruck, G. (2015). Parochialism in preschool boys' resource allocation. *Evolution and Human Behavior, 36*, 256–264. <https://doi.org/10.1016/j.evolhumbehav.2014.12.002>
- Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science, 16*, 162–166. <https://doi.org/10.1111/j.1467-8721.2007.00496.x>
- Birnbaum, D., Deeb, I., Segall, G., Ben-Eliyahu, A., & Diesendruck, G. (2010). The development of social essentialism: The case of Israeli children's inferences about Jews and Arabs. *Child Development, 81*, 757–777. <https://doi.org/10.1111/j.1467-8624.2010.01432.x>
- Boldry, J. G., & Gaertner, L. (2006). Separating status from power as an antecedent of intergroup perception. *Group Processes & Intergroup Relations, 9*, 377–400. <https://doi.org/10.1177/1368430206064640>
- Boldry, J. G., Gaertner, L., & Quinn, J. (2007). Measuring the measures: A meta-analytic investigation of the measures of outgroup homogeneity. *Group Processes & Intergroup Relations, 10*, 157–178. <https://doi.org/10.1177/1368430207075153>
- Brauer, M., & Er-rafiy, A. (2011). Increasing perceived variability reduces prejudice and discrimination. *Journal of Experimental Social Psychology, 47*, 871–881. <https://doi.org/10.1016/j.jesp.2011.03.003>
- Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.
- Buttelmann, D., & Böhm, R. (2014). The ontogeny of the motivation that underlies in-group bias. *Psychological Science, 25*, 921–927. <https://doi.org/10.1177/0956797613516802>
- Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences, 37*, 461–480. <https://doi.org/10.1017/S0140525X13002197>
- Deeb, I., Segall, G., Birnbaum, D., Ben-Eliyahu, A., & Diesendruck, G. (2011). Seeing isn't believing: The effect of intergroup exposure on children's essentialist beliefs about ethnic categories. *Journal of Personality and Social Psychology, 101*, 1139–1156. <https://doi.org/10.1037/a0026107>
- Diesendruck, G. (2013). Essentialism: The development of a simple, but potentially dangerous, idea. In M. R. Banaji & S. Gelman (Eds.), *Navigating the social world: The early years* (pp. 263–268). New York, NY: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199890712.003.0048>
- Diesendruck, G., & Eldror, E. (2011). What children infer from social categories. *Cognitive Development, 26*, 118–126. <https://doi.org/10.1016/j.cogdev.2010.11.001>
- Diesendruck, G., & Menahem, R. (2015). Essentialism promotes children's inter-ethnic bias. *Frontiers in Psychology, 6*. <https://doi.org/10.3389/fpsyg.2015.01180>
- Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development, 82*, 793–811. <https://doi.org/10.1111/j.1467-8624.2011.01577.x>
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individualizing processes: Influences of information and

- motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York, NY: Elsevier. [https://doi.org/10.1016/s0065-2601\(08\)60317-2](https://doi.org/10.1016/s0065-2601(08)60317-2)
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195154061.001.0001>
- Gil-White, F. J. (2001). Are ethnic groups biological “species” to the human brain? Essentialism in our cognition of some social categories. *Current Anthropology*, *42*, 515–553. <https://doi.org/10.1086/321802>
- Guinote, A., Mouro, C., Pereira, M. H., & Monteiro, M. B. (2007). Children’s perceptions of group variability as a function of status. *International Journal of Behavioral Development*, *31*, 97–104. <https://doi.org/10.1177/0165025407073930>
- Gutheil, G., & Gelman, S. A. (1997). Children’s use of sample size and diversity information within basic-level categories. *Journal of Experimental Child Psychology*, *64*, 159–174. <https://doi.org/10.1006/jecp.1996.2344>
- Hamlin, J. K., Mahajan, N., Liberman, Z., & Wynn, K. (2013). Not like me = bad: Infants prefer those who harm dissimilar others. *Psychological Science*, *24*, 589–594. <https://doi.org/10.1177/0956797612457785>
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*, 252–264. https://doi.org/10.1207/s15327957pspr1003_4
- Haslam, N., & Whelan, J. (2008). Human natures: Psychological essentialism in thinking about differences between people. *Social and Personality Psychology Compass*, *2*, 1297–1312. <https://doi.org/10.1111/j.1751-9004.2008.00112.x>
- Heit, E. (2000). Properties of inductive reasoning. *Psychonomic Bulletin & Review*, *7*, 569–592. <https://doi.org/10.3758/BF03212996>
- Henderson, A. M. E., & Woodward, A. L. (2012). Nine-month-old infants generalize object labels, but not object preferences across individuals. *Developmental Science*, *15*, 641–652. <https://doi.org/10.1111/j.1467-7687.2012.01157.x>
- Hewstone, M., & Hamberger, J. (2000). Perceived variability and stereotype change. *Journal of Experimental Social Psychology*, *36*, 103–124. <https://doi.org/10.1006/jesp.1999.1398>
- Hirschfeld, L. A. (1996). *Race in the making: Cognition, culture, and the child’s construction of human kinds*. Cambridge, MA: MIT Press. <https://doi.org/10.2307/2654625>. ISSN 0094-3061
- 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>
- Israeli Central Bureau of Statistics. (2015). *Local government report*. Retrieved from http://www.cbs.gov.il/reader/cw_usr_view_SHTML?ID=357
- Judd, C. M., & Park, B. (1988). Out-group homogeneity: Judgments of variability at the individual and group levels. *Journal of Personality and Social Psychology*, *54*, 778–788. <https://doi.org/10.1037/0022-3514.54.5.778>
- Kalish, C. W. (2002). Children’s predictions of consistency in people’s actions. *Cognition*, *84*, 237–265. [https://doi.org/10.1016/S0010-0277\(02\)00052-5](https://doi.org/10.1016/S0010-0277(02)00052-5)
- Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, *88*, 686–702. <https://doi.org/10.1037/0022-3514.88.4.686>
- Kinzler, K. D., & Dautel, J. B. (2012). Children’s essentialist reasoning about language and race: Language and race. *Developmental Science*, *15*, 131–138. <https://doi.org/10.1111/j.1467-7687.2011.01101.x>
- Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 12577–12580. <https://doi.org/10.1073/pnas.0705345104>
- Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, *21*, 1134–1140. <https://doi.org/10.1177/0956797610376652>
- Li, F., Cao, B., Li, Y., Li, H., & Deák, G. (2009). The law of large numbers in children’s diversity-based reasoning. *Thinking & Reasoning*, *15*, 388–404. <https://doi.org/10.1080/13546780903343227>
- Linville, P. W., Fischer, G. W., & Yoon, C. (1996). Perceived covariation among the features of ingroup and outgroup members: The outgroup covariation effect. *Journal of Personality and Social Psychology*, *70*, 421–436. <https://doi.org/10.1037/0022-3514.70.3.421>
- Lopez, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development*, *63*, 1070–1090. <https://doi.org/10.1111/j.1467-8624.1992.tb01681.x>
- Lorenzi-Cioldi, F. (1993). They all look alike, but so do we . . . sometimes: Perceptions of in-group and out-group homogeneity as a function of sex and context. *British Journal of Social Psychology*, *32*, 111–124. <https://doi.org/10.1111/j.2044-8309.1993.tb00990.x>
- Noyes, A., & Christie, S. (2016). Children prefer diverse samples for inductive reasoning in the social domain. *Child Development*, *87*, 1090–1098. <https://doi.org/10.1111/cdev.12522>
- Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*, 185–200. <https://doi.org/10.1037/0033-295X.97.2.185>
- Ostrom, T. M., Carpenter, S. L., Sedikides, C., & Li, F. (1993). Differential processing of in-group and out-group information. *Journal of Personality and Social Psychology*, *64*(1), 21–34. <https://doi.org/10.1037/0022-3514.64.1.21>
- Park, B., & Hastie, R. (1987). Perception of variability in category development: Instance- versus abstraction-based stereotypes. *Journal of Personality and Social*

- Psychology*, 53, 621–635. <https://doi.org/10.1037/0022-3514.53.4.621>
- Park, B., & Judd, C. M. (1990). Measures and models of perceived group variability. *Journal of Personality and Social Psychology*, 59, 173–191. <https://doi.org/10.1037/0022-3514.59.2.173>
- Park, B., Ryan, C. S., & Judd, C. M. (1992). The role of meaningful subgroups in explaining differences in perceived variability for in-groups and out-groups. *Journal of Personality and Social Psychology*, 63, 553–567. <https://doi.org/10.1037/0022-3514.63.4.553>
- Pauker, K., Ambady, N., & Apfelbaum, E. P. (2010). Race salience and essentialist thinking in racial stereotype development. *Child Development*, 81, 1799–1813. <https://doi.org/10.1111/j.1467-8624.2010.01511.x>
- Pickett, C. L., & Brewer, M. B. (2001). Assimilation and differentiation needs as motivational determinants of perceived in-group and out-group homogeneity. *Journal of Experimental Social Psychology*, 37, 341–348. <https://doi.org/10.1006/jesp.2000.1469>
- Prentice, D. A., & Miller, D. T. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science*, 16, 202–206. <https://doi.org/10.1111/j.1467-8721.2007.00504.x>
- Raabe, T., & Beelmann, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences. *Child Development*, 82, 1715–1737. <https://doi.org/10.1111/j.1467-8624.2011.01668.x>
- Rhodes, M. (2013). How two intuitive theories shape the development of social categorization. *Child Development Perspectives*, 7(1), 12–16. <https://doi.org/10.1111/cdep.12007>
- Rhodes, M., & Brickman, D. (2010). The role of within-category variability in category-based induction: A developmental study. *Cognitive Science*, 34, 1561–1573. <https://doi.org/10.1111/j.1551-6709.2010.01137.x>
- Rhodes, M., Brickman, D., & Gelman, S. A. (2008a). Sample diversity and premise typicality in inductive reasoning: Evidence for developmental change. *Cognition*, 108, 543–556. <https://doi.org/10.1016/j.cognition.2008.03.002>
- Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, 59, 244–274. <https://doi.org/10.1016/j.cogpsych.2009.05.001>
- Rhodes, M., Gelman, S. A., & Brickman, D. (2008b). Developmental changes in the consideration of sample diversity in inductive reasoning. *Journal of Cognition and Development*, 9, 112–143. <https://doi.org/10.1080/15248370701836626>
- Rhodes, M., Leslie, S.-J., Saunders, K., Dunham, Y., & Cimpian, A. (2017). How does social essentialism affect the development of inter-group relations? *Developmental Science*. <https://doi.org/10.1111/desc.12509>
- Rhodes, M., Leslie, S.-J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 13526–13531. <https://doi.org/10.1073/pnas.1208951109>
- Rhodes, M., & Liebenson, P. (2015). Continuity and change in the development of category-based induction: The test case of diversity-based reasoning. *Cognitive Psychology*, 82, 74–95. <https://doi.org/10.1016/j.cogpsych.2015.07.003>
- Ryan, C. S., Judd, C. M., & Park, B. (1996). Effects of racial stereotypes on judgments of individuals: The moderating role of perceived group variability. *Journal of Experimental Social Psychology*, 32, 71–103. <https://doi.org/10.1006/jesp.1996.0004>
- Segall, G., Birnbaum, D., Deeb, I., & Diesendruck, G. (2015). The intergenerational transmission of ethnic essentialism: How parents talk counts the most. *Developmental Science*, 18, 543–555. <https://doi.org/10.1111/desc.12235>
- Simon, B., & Brown, R. (1987). Perceived intragroup homogeneity in minority-majority contexts. *Journal of Personality and Social Psychology*, 53, 703–711. <https://doi.org/10.1037/0022-3514.53.4.703>
- Stewart, T. L., & Vassar, P. M. (2000). The effect of occupational status cues on memory for male and female targets. *Psychology of Women Quarterly*, 24, 161–169. <https://doi.org/10.1111/j.1471-6402.2000.tb00197.x>
- Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology*, 94, 1033–1047. <https://doi.org/10.1037/0022-3514.94.6.1033>

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Table S1. Samples Used for Induction in Study 1

Table S2. Samples Used for Induction in Study 2