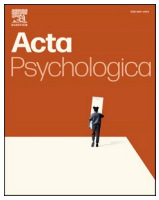
Acta Psychologica 234 (2023) 103864

Contents lists available at ScienceDirect

Acta Psychologica

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



Diversity vs. ingroup: How children generalize for the common good

Rongzhi Liu [a], Yi Zhang [b], Stella Christie [b,*]

[a] Department of Psychology, University of California, Berkeley, CA, United States
[b] Tsinghua Laboratory of Brain and Intelligence, Department of Psychology, Tsinghua University, China

ARTICLE INFO

Keywords:
Ingroup bias
Diversity-based reasoning
Inductive reasoning
Intergroup relations

ABSTRACT

How do children generalize for the common good? The present study investigated whether children are more likely to use the preference of their ingroup (ingroup rationale) or that of a diverse group (diversity rationale) as a basis for generalization about the broader community. In a series of studies, five-year-olds from two different cultures (US and China), yet living in environments with analogous ingroup majority-outgroup minority structure, were asked to generalize either the preference of a diverse sample or the preference of an ingroup sample to the majority. We found that children from both cultures have a default strategy to generalize from their ingroup (Study 1). However, Studies 2–4 show that this ingroup default is amenable to change, suggesting that children mostly use this strategy because ingroup members were salient and conveniently available. When ingroup was removed or reduced (Study 2), or when primed with photos of diverse populations (Studies 3 & 4), children changed their strategies and were more likely to use the diversity-rationale. In both cultures, the intergroup structure of children's living environment exerts similar pressures, resulting in analogous outcomes in generalizing for the common good.

1. Introduction

We form opinions about social others—prejudices, stereotypes—based on a small number of interactions. Such generalization from limited data is necessary; after all, we cannot expect to meet every single college student before forming an opinion about that demographic. But does the variability of the sample influence our confidence in making generalizations? Theories of inductive generalization predict that it should: one ought to be more confident to generalize from a diverse set of evidence than from a narrow one (Heit, 2000; Osherson et al., 1990). For example, we should be more likely to form a stereotype that college students are politically active after interacting with students from small liberal arts colleges and large state universities (diverse sample) than from interactions solely with students from small colleges (narrow sample).

Most work on comparison of evidence for generalization shows that adults and older children generally follow this *diversity rationale*: given a diverse vs. a narrow set of evidence, they are more likely to generalize from the diverse set (Gutheil & Gelman, 1997; Heit, 2000; Li et al., 2009; López et al., 1992; Osherson et al., 1990; Rhodes & Brickman, 2010; Rhodes, Brickman, et al., 2008; Rhodes, Gelman, et al., 2008). For example, 9-year-olds reasoned that if a cat and a buffalo (diverse sample) had property X, while a cow and a buffalo (narrow sample) had property Y, then it ought to be more likely to find X, but not Y, among all animals (López et al., 1992). But while prevalent in non-social reasoning, diversity rationale seems to be less prominent in the social world. If anything, there is even a documented tendency for people to adopt the opposite of the diversity rationale when they reason about social others. For example, like-minded individuals tend to associate with one another (Festinger, 1957; Kossinets & Watts, 2009), and sort themselves into like-minded communities that serve as echo chambers for pre-existing opinions (Bishop, 2009; Sunstein, 2009). One study analyzing millions of Facebook users' news-sharing behaviors found that individuals chose to be exposed to less diverse contents compared to Facebook's News Feed algorithmic ranking (Bakshy et al., 2015). We also see a social trend of favoring homogeneity rather than diversity, for example when communities resist immigration or integration. What are the root causes of this apparent lack of diversity rationale? How do they manifest in early childhood?

1.1. Ingroup rationale

One possible explanation to the lack of diversity rationale usage is that social reasoning is often biased by ingroup considerations. Rather

* Corresponding author at: Tsinghua University, Beijing 100084, China.
E-mail address: christie@tsinghua.edu.cn (S. Christie).

https://doi.org/10.1016/j.actpsy.2023.103864
Received 21 February 2022; Received in revised form 18 October 2022; Accepted 8 February 2023
Available online 21 February 2023

than privileging a diverse sample of evidence over a narrow one, people take the opinions of their (narrow) ingroup to be more generalizable to the whole population than that of the diverse others. There are at least two reasons for adopting the *ingroup rationale*: (i) Importance: thinking that one's own ingroup is more important than outgroups, giving more weights to the ingroup's opinions/preferences; (ii) Salience-Convenience: thinking that the ingroup is more representative of the whole population as they are more salient in one's mind and more convenient to access. Whichever the reason, when an individual uses the *ingroup* instead of the *diversity* rationale, this favors certain groups at a cost of potentially harming others. This cost is largest when the reasoning concerns the broader community, for example a distribution of resources. Partly motivated by a current concern regarding diversity and tolerance in many parts of the world, here we investigate the root mechanisms underlying social generalization. We asked whether young children are more likely to use the ingroup or diversity rationale when they make inferences about the general population's preferences, and whether their adherence to a particular rationale is robust.

1.1.1. Ingroup rationale is salient and convenient

To review, both reasons for using ingroup rationale are present in infants and young children. On the first reason—importance of ingroup over outgroup, studies with young children show that they do not treat ingroup and outgroup members equally. For example, 3-year-olds prefer to befriend children of the same gender (Shutts et al., 2013). Likewise, 5- to 6-years-old White American children report that they prefer White over Black individuals, and show an implicit pro-White attitude as measured by the Implicit Association Test (Baron and Banaji, 2006). Other than race and gender, children also show ingroup preference when language is the category marker, preferring to befriend and allocate more resources to children who speak the same language as themselves (Kinzler et al., 2007, 2011). Even when children are randomly assigned to a previously unfamiliar group based on arbitrary cues such as shirt color (i.e., the minimal group paradigm), they also show ingroup preferences (Dunham et al., 2011).

On the second reason—Salience-Convenience—there is evidence that infants perceive ingroup members to be more salient. For example, 3- to 4-month-olds prefer to look at faces of the same gender as their primary caregivers (Quinn et al., 2002). Five- to 6- month-olds prefer to look at an adult who previously spoke to them in their native language over an adult who previously spoke an unnatural language (natural speech played in reverse), a foreign language, or their native language with a foreign accent (Kinzler et al., 2007). Later on in life, demographic and geographic contexts may enhance ingroup saliency and convenience, as we often live closer to our ingroup members, thereby interacting more frequently with ingroup than with outgroup members. For instance, in the US people often find themselves living closer to people similar to them in terms of racial and ethnic identity, cultural background, and political ideologies (Bishop, 2009; Sunstein, 2009). Interestingly, 3-month-old infants prefer to look at faces of the same race as themselves, but only if they live in a monoracial environment (Bar-Haim et al., 2006).

1.2. Children's usage of diversity rationale

What about the diversity rationale? Are children also capable of using the diversity rationale in their social decisions? Prior studies in the non-social domain show that young children do not adeptly choose diverse over narrow evidence for generalization. For example, in López et al. (1992), 5-year-olds chose at random from diverse [cat-buffalo] or narrow [cow-buffalo] as evidence to generalize to all animals (see also Gutheil & Gelman, 1997; Li et al., 2009; Rhodes et al., 2008a, b; Rhodes & Brickman, 2010). A spontaneous, reliable preference for diverse evidence does not come until around nine years of age (Li et al., 2009; López et al., 1992).

However, two recent studies show that even five-year-olds understand the diversity rationale. In fact, the social domain enables an *easier* usage of it (Noyes & Christie, 2016; Shilo et al., 2019). In Noyes and Christie (2016), 5-year-olds failed to use the diversity rationale in the biology domain, choosing at random from diverse [lion-mouse] or narrow [lion-tiger] as evidence to generalize to all animals. However, the same group of 5-year-olds used diversity rationale when the question was framed as a social task. They reliably said they wanted to ask the diverse [a boy and a girl] more than the narrow [two boys] group to find out whether all children liked a novel toy X. The preference for diversity also held when the diversity was exemplified by race (e.g., diverse: White boy-Black boy, narrow: White boy-White boy). Using a similar task, Shilo et al. (2019) found that 5-year-olds, 8-year-olds, and adults in Israel and Germany used the diversity rationale when reasoning about psychological properties, regardless of whether they were reasoning about ingroups (Jews in Israel; Germans with German origins in Germany) or outgroups (Arabs in Israel; Germans with Turkish origins in Germany). However, when reasoning about biological properties, participants were less likely to use diversity rationale when reasoning about outgroups compared to reasoning about ingroups (i.e., the outgroup homogeneity effect).

1.3. The current study

To summarize, there are good reasons to think that young children—at least by five years of age—can employ either rationale—ingroup or diversity—in making generalizations in the social world. Prior studies, however, never contrasted these two rationales within the same social context. While ingroup rationale is heavily documented in children's social reasoning, diversity rationale is mostly investigated in the non-social domain. When the diversity rationale is observed in the social domain, it is mostly about making generalizations in the abstract. For example, the Noyes and Christie (2016) study only asked children about "which group we *should* ask to find out whether all children like a novel toy." However, children's ability to use the diversity rationale in the abstract (as observed in Noyes & Christie, 2016), does not automatically predict that they will *use* this rationale in real social situations.

Thus, in the current studies we wanted to investigate a social context akin to a real-world scenario where making a generalization matters and has consequences. This involves (at least) two things: i) where the group's preferences are already known (e.g., college students like lively restaurants, while senior citizens like elegant, refined ones), ii) decision based on one preference over the other impacts everyone (e.g., on this vacation the whole family—young and old—go to lively, not elegant restaurants). That is, generalization based on ingroup's preferences brings about different outcomes from generalization based on diverse group's preferences. To do this we set up a scenario where children are generalizing for the common good. We told participants about children from "the Jiffy school," and participants' task was to select a toy that made most Jiffy school children happy. Just like in the real world, subgroups of the Jiffy school children have different *known* preferences. The diverse group liked toy X, while the narrow-ingroup liked toy Y. Our question was whether children would use the diversity rationale and decide to buy toy X for all Jiffy children, or use the ingroup rationale and generalize a preference for Y to the Jiffy population. As participants could not see the toys (which were hidden), the only basis for judgment was the group itself—ingroup vs. diverse.

Because we are interested in social perception, we displayed group diversity based on either race or gender—the two most influential and salient social cues of diversity. To preface, in Study 1, five-year-olds showed a strong tendency to generalize the preference of the narrow-ingroup sample rather than the preference of the diverse sample. In Studies 2–4 we investigated the reasons and robustness of using the ingroup rationale. In Study 2 we removed the ingroup presence, to test whether children can use the diversity rationale per se when making generalizations for the broader population (as this has not been tested before). Studies 3 and 4 tested the Salience-Convenience

hypothesis—that children used the ingroup rationale because ingroup members were more salient or convenient to access. To do so, we first showed children a picture of the Jiffy school children showing a diverse population (for example, a photo of 50 % White and 50 % Black children) for a mere 1 minute. Following this prime, children were tested in the toy-generalization paradigm as in Study 1. If children's ingroup bias in Study 1 was due to Salience-Convenience, rather than because they deeply believed that ingroup's preference was inherently more important, we should expect this brief diversity prime to change children's ingroup bias.

We tested 5-year-olds in the 4 studies, since past research has shown that children at this age can use diversity rationale (Noyes & Christie, 2016; Shilo et al., 2019). Our studies were conducted at two sites: a Northeastern suburban town in U.S, and a major metropolis in China. Testing at these two sites from different cultures brings us closer to understanding the psychological phenomena of humans in general, rather than just the WEIRD (Western, educated, industrialized, rich, and democratic) population (Rad et al., 2018). Our goal is to understand the cognitive mechanism of how children make use of the ingroup vs. diversity rationale when they make generalizations about social groups, and whether the same cognitive mechanism is at play across the two cultures that we tested. Although the Chinese and US children come from different cultures, the ingroup majority-outgroup minority structures are analogous. Specifically, the racial demographics of both of our testing sites were relatively homogeneous: Our US testing site is a white-majority town, and the Chinese testing site is a Han Chinese-majority city. Thus, we predicted that there would not be any cross-cultural differences across testing sites in our studies.

2. Study 1: generalizing from diverse vs. narrow-ingroup samples

In Study 1, we asked whether 5-year-olds would generalize based on the preference of the diverse or the narrow-ingroup samples when reasoning for the general population. We told participants about a novel Jiffy school in a far-away place, and asked them to decide which item majority of the children in the novel school would prefer: the one liked by the diverse pair (e.g., a White child and a Black child) or the one liked by the narrow pair (e.g., two White children). To avoid potential conflations with motives of personal gain, the study was designed such that participants were not directly involved—the toy was for the Jiffy school children only.

European American (henceforth White) and Han Chinese (henceforth Chinese) children were tested with analogous (but not identical) stimuli on two types of trials: race and gender—where diversity is cued by these factors respectively. In the race trials, the diverse sample consisted of [a child of ingroup race and a child of outgroup race] while the narrow sample consisted of [two children of ingroup race]. We did not remove ingroup entirely from the diverse samples, as we want to mimic more of a real-life situation, where a diverse environment often contains one's ingroup. To make the stimuli even more comparable, we decided to use African American (henceforth Black) children as the outgroup race for both White and Chinese participants.

In the gender trials, the diverse evidence consisted of a boy and a girl of ingroup race, and the narrow evidence consisted of two boys or two girls of ingroup race.

2.1. Method

2.1.1. Participants

Thirty-five children (15 White, 20 Chinese; 21 males, 14 females, $M = 5.47$ years, $SD = 0.42$, range = 4.92–6.58 years) participated in Study 1. White children were recruited from local preschools in Swarthmore, US, where the majority of residents are White (84.5 % White, 5.4 % Asian, 4.9 % Black, 3.5 % Hispanic or Latino, 1.0 % American Indian and Alaska Native, 3.7 % two or more races). Chinese children were recruited using online announcements. All children are from Beijing, China, where the majority of residents are Chinese (99.5 % Chinese, 0.5 % foreigners). At both sites, parents received letters containing consent forms; only children with signed consent forms participated in the study. Children received a t-shirt or a book as a thank-you gift for participating. Our sample size was determined based on a previous study (Noyes & Christie, 2016), which tested 32 children per condition. The design of each condition in their study was similar to the design of our study.

2.1.2. Design

Participants had to decide whether a *known* preference of a diverse pair [e.g., girl-boy like toy X] versus a narrow-ingroup pair [e.g., boy-boy like toy Y] was more generalizable to the majority of children in the hypothetical "Jiffy" school. Group diversity was cued by race and gender.

In race trials, White children were shown diverse pairs consisting of one White child and one Black child of the same gender (either two boys or two girls), and narrow pairs consisting of two White children of the same gender (Fig. 1). Chinese children were shown diverse pairs consisting of one Chinese child and one Black child of the same gender, and narrow pairs consisting of two Chinese children of the same gender (Fig. 1).

In gender trials, White children were shown diverse pairs consisting of one White boy and one White girl, while the narrow pairs consisted of two White boys or two White girls (Fig. 1b). Chinese children were shown diverse pairs consisting of one Chinese boy and one Chinese girl, while the narrow pairs consisted of two Chinese boys or two Chinese girls (Fig. 1b). There were four race trials and four gender trials, for a total of eight trials. The order of trials was counterbalanced across participants.

2.1.3. Materials

We created a set of stimuli consisting of pictures of children. The pictures were headshot photographs of children's faces and upper torsos. In the US stimuli, we used thirty-two unique photographs (28 White children, 4 Black children; 16 females, 16 males); in the Chinese stimuli, we used another set of thirty-two unique photographs (28 Chinese children, 4 Black children; 16 females, 16 males). In all photographs, children were smiling, facing forward, and against a White background. We selected photographs of children who appeared similar in terms of attractiveness and happiness.

2.1.4. Procedure

Children were asked to help the experimenter to figure out which item would be preferred by a majority of children in the Jiffy School.

In each trial, children were told that there was an object of the same type (e.g., a toy) in each of the two bags/boxes. They were told that the hidden toy in one bag/box was liked by the diverse pair while the hidden toy in the other bag/box was liked by the narrow pair. For example, the experimenter said, "These Jiffy kids on the left like this toy. And these Jiffy kids on the right like this toy." Next, the child participant was asked, "Which toy do you think *more* of the other Jiffy kids like to play with? The toy these Jiffy kids like (pointing to one group of kids), or the toy these Jiffy kids like (pointing to the other group of kids)?" After the choice was made, the experimenter gave a generic and performance-independent praise (e.g., "great!") and the next trial was initiated immediately. There were eight questions for each participant; the order of the questions and the left/right placement of diverse groups were counterbalanced. All data and materials are available on the Open Science Framework: https://osf.io/w5yh4/?view_only=dbc9431fc6cd48758897cec6698c89a9.

2.2. Results and discussion

Children's selection of the diverse pairs (i.e., using diversity rationale) in race and gender trials are shown in Fig. 5 and Table 1. On

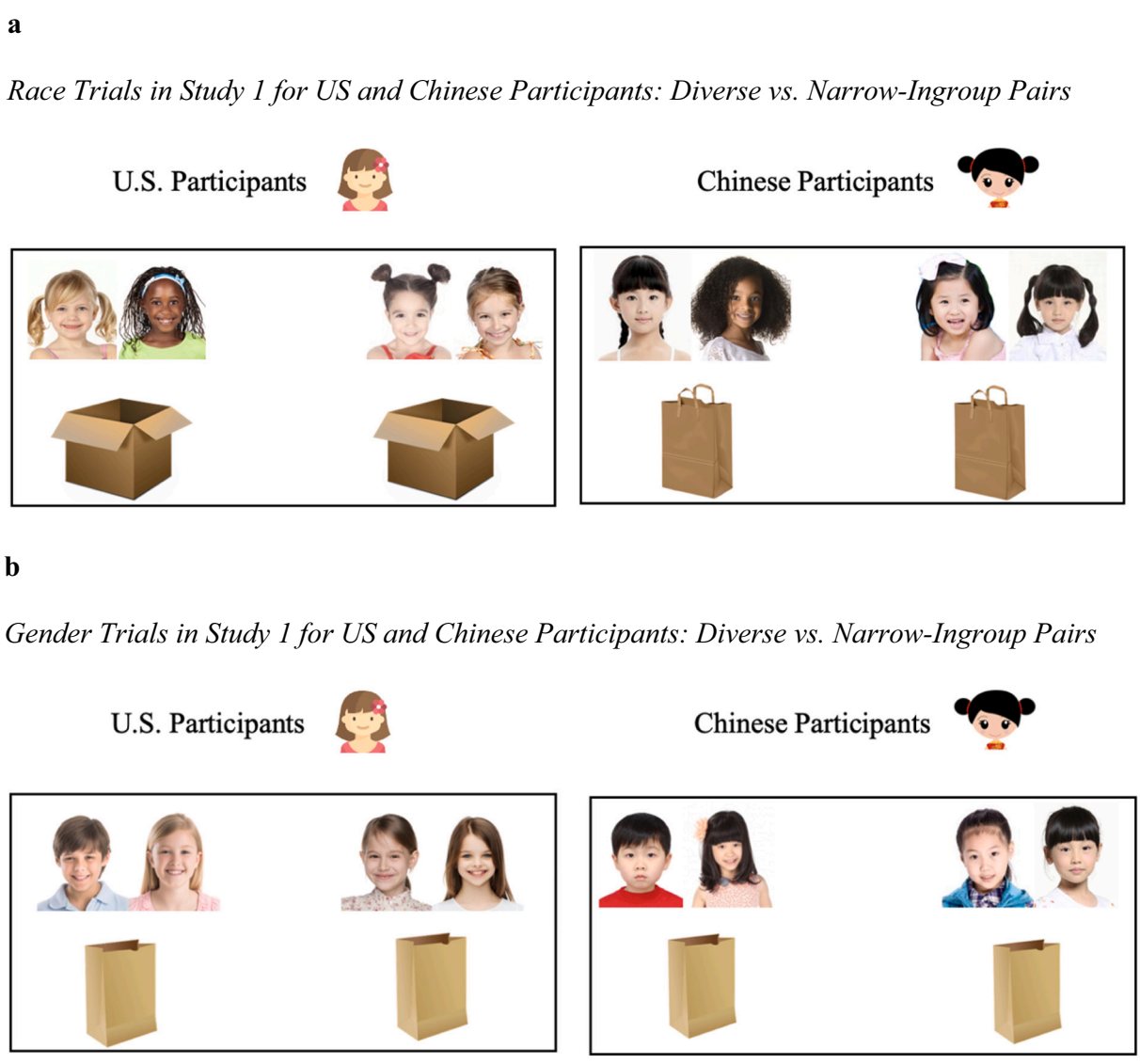


Fig. 1. a. Race Trials in Study 1 for US and Chinese Participants: Diverse vs. Narrow-Ingroup Pairs.
b. Gender Trials in Study 1 for US and Chinese Participants: Diverse vs. Narrow-Ingroup Pairs.
Note. Children were told that each group liked a different item (the diverse pair liked item X, while the narrow pair liked item Y); the items were always hidden in the bags. These stimuli were used in Studies 1, 3, and 4.

average, children selected the diverse pair in 37 % of the trials, which is below chance (Exact Binomial test: $P = .37$ [.31, .43]; $p < .001$). That is, children believed that the narrow-ingroup pair's preference was the one to be generalized. The usage of ingroup rationale is stronger in race trials than in gender trials. In race trials, children chose the diverse pair in 29 % of the trials, which is below chance ($P = .29$ [.22, .38]; $p < .001$), while in gender trials, children chose the diverse pair in 45 % of the trials, which is at chance ($P = .45$ [.36, .53], $p = .23$). It is possible that we found an attenuated usage of the ingroup rationale for gender trials because children had more opportunities to interact with children of other genders.

Next, we used mixed-effects logistic regression to predict children's choice (Diverse = 1, Non-Diverse = 0) from trial type (race trial or gender trial), children's age, children's gender, and their interactions, with random intercepts for country and participant. There was a significant effect of trial type. Children were more likely to select the diverse pair in gender trials than in race trials ($\beta = 0.72$, SE = 0.26, $p = .006$, Odds Ratio = 2.06). That is, children have a weaker tendency to use the ingroup rationale when diversity is cued by gender.

Overall, the results of Study 1 suggest that children from both US and China tend to use the ingroup over the diversity rationale when making generalizations about the broader community's preferences. In Study 2, we examined whether children could revert to diversity rationale when ingroup consideration is absent or reduced.

3. Study 2: generalizing from diverse vs. narrow-outgroup

In Study 2 we tested a new group of 5-year-olds from the same social environments as in Study 1, but this time removing or reducing the ingroup presence.

3.1. Method

3.1.1. Participants

Thirty-five children (14 White, 21 Chinese; 25 males, 10 females; $M = 5.38$ years, $SD = 0.33$, range = 5–6.17 years) participated in Study 2. Children were recruited from the two testing sites in the same manner as in Study 1.

Table 1
The proportion, 95 % CI, and exact binomial test of diverse selection by study and trial type.

	Trial type	*P*	95 % CI	*p*-value (Exact Binomial Test)
Study 1: Baseline	Both trials	.37	[.31, .43]	<.001
Study 1: Baseline	Race trial	.29	[.22, .38]	<.001
Study 1: Baseline	Gender trial	.45	[.36, .53]	.23
Study 2: Outgroup	Both trials	.57	[.51, .63]	.025
Study 2: Outgroup	Race trial	.56	[.47, .65]	.192
Study 2: Outgroup	Gender trial	.58	[.49, .66]	.076
Study 3: 50 %–50 % picture	Both trials	.48	[.42, .54]	.54
Study 3: 50 %–50 % picture	Race trial	.50	[.41, .59]	1
Study 3: 50 %–50 % picture	Gender trial	.46	[.37, .55]	.39
Study 4: 25 %–75 % picture	Both trials	.52	[.46, .58]	.60
Study 4: 25 %–75 % picture	Race trial	.42	[.33, .50]	.055
Study 4: 25 %–75 % picture	Gender trial	.62	[.53, .70]	.006

3.1.2. Design

Similar to Study 1, children were asked to decide whether the known preferences of the diverse versus the narrow pairs was more generalizable to the majority of children in the Jiffy school. But unlike Study 1, participants' ingroup race (White or Chinese) was not present in the stimuli of Study 2.

In race trials, White children were shown diverse pairs consisting of one Black child and one Asian child of the same gender, and narrow pairs consisting of two Black children of the same gender. Chinese children were shown the analogous diverse [a White child and a Black child], and narrow pairs [e.g. two Black children] (Fig. 2).

For gender trials, it is not possible to entirely remove an ingroup gender, as a child participant always belongs to one of the gender's group. As such, we attempted to reduce the ingroup factor by removing the *race* ingroup. That is, for example, while in Study 1 a Chinese girl participant saw the narrow-ingroup pair to consist of two Chinese girls, in Study 2 the narrow pair would be two White girls (Fig. 2b). The diverse pair stimuli consisted of a girl and a boy of the outgroup race (e.g., a White boy and a White girl).

All other design (e.g., the number of trials, the counterbalance orders, the presentation of the stimuli) was identical to Study 1.

3.1.3. Materials and procedure

Two different sets of stimuli were created in the same manner as in Experiment 1. In the US stimuli, we used thirty-two unique photographs (28 Black children, 4 Asian children; 16 females, 16 males); in the Chinese stimuli, we used thirty-two unique photographs (24 White children, 8 Black children; 16 females, 16 males). Children were tested in exactly the same way as in Study 1.

3.2. Results and discussion

On average, children selected the diverse pair in 57 % of the trials, which is above chance (Exact Binomial test: $P = .57$ [.51, .63]; $p = .025$). That is, when ingroup race was removed from the samples, children preferred to use the diversity rationale when generalizing for the broader population.

As noted above, while ingroup was removed entirely from the race trials, it was merely reduced in the gender trials. As such, one could expect that in gender trials, children still chose to generalize the narrow (ingroup)'s preference. However, there was no difference between the two types of trials, children chose the diverse pair 56 % of the time in race trials ($P_{race} = .56$ [.47, .65]; $p = .192$), and 58 % of the time in gender trials ($P_{gender} = .58$ [.49, .66], $p = .076$). Furthermore, mixed-effects logistic regression did not reveal any effect of trial type, age, gender, or their interactions.

We next compared the data from Studies 1 and 2 to analyze whether removing ingroup presence increased the usage of diversity rationale. We used mixed-effects logistic regression to predict children's choice (Diverse = 1, Non-Diverse = 0) from study (1 or 2), trial type, age, gender, and their interactions, with random intercepts for country and participant. Indeed, we found a main effect of study: overall, children were more likely to choose the diverse pairs in Study 2 than in Study 1 ($\beta = 0.85$, SE = 0.20, $p < .001$, OR = 2.35). We also found a main effect of trial type: children were more likely to choose the diverse pairs in gender trials than in race trials ($\beta = 0.37$, SE = 0.18, $p = .036$, OR = 1.45). There was no interaction between study and trial type – the difference between Studies 1 and 2 was observed in race trials ($\beta = 1.37$, SE = 0.37, $p < .001$, OR = 3.93) as well as in gender trials ($\beta = 0.59$, SE = 0.30, $p = .049$, OR = 1.81). That is, children were more likely to use the diversity rationale both when ingroup was completely absent in the race trials, and when ingroup was relatively absent in the gender trials. One possible interpretation of the gender trials results is that in Study 1, the narrow-ingroup pair consisted of two children who are of the same race and gender as the participants, thus participants have a strong ingroup preference for the narrow-ingroup pair. In contrast, in Study 2, the narrow-ingroup pair consisted of two children who are of the same gender as the participants, but not of the same race. Thus, participants' ingroup preference for the narrow-ingroup pair is weaker than in Study 1, and they were more likely to use the diversity rationale instead.

4. Study 3: mixed race and gender population

Study 2 showed that 5-year-olds were more likely to use the diversity rationale—generalizing from diverse over narrow samples of evidence—when ingroup consideration was absent or reduced. This result raises the possibility that children in Study 1 generalized the preference of their ingroup simply because it was more convenient to do so (the Salience-Convenience hypothesis), rather than because they believed deeply in the importance of their ingroup (Importance hypothesis). Since our children samples come from a majority-ingroup environment (in both US and China), they may have interacted more frequently or even exclusively with ingroup members, making ingroup members salient and/or available. Thus, in the absence of explicit information about the Jiffy school population, children may simply assume that most Jiffy school children are just like themselves.

If this *Salience-Convenience* hypothesis is correct, then giving some information about the population should change children's strategy. At the very least, they should be *less* likely to use the ingroup rationale. Alternatively, already by 5 years of age, children living in a majority-ingroup environment think that their ingroup truly matters; its preference ought to be generalized regardless. If this *Importance* hypothesis was the reason behind Study 1, then we should expect the ingroup rationale to be robust. To test this, in Study 3 we first showed participants a Jiffy school photo consisting of 50 % ingroup race and 50 % outgroup race, half girls and half boys. After children saw this one photo very briefly, they were tested with the identical stimuli as in Study 1, where ingroup rationale was pitted against diversity rationale. Our question was whether this minimal information about the population impacted children's inferences about the broader community.

4.1. Method

4.1.1. Participants

Thirty-five children (15 White, 20 Chinese; 22 males, 13 females; *M*

a

Race Trials in Study 2 for US and Chinese Participants: Diverse vs. Narrow Pairs

b

Gender Trials in Study 2 for US and Chinese Participants: Diverse vs. Narrow Pairs

Fig. 2. a. Race Trials in Study 2 for US and Chinese Participants: Diverse vs. Narrow Pairs.
b. Gender Trials in Study 2 for US and Chinese Participants: Diverse vs. Narrow Pairs.
Note. Children were told that each group liked a different item (the diverse pair liked X, while the narrow pair liked Y); the items were hidden in a bag.

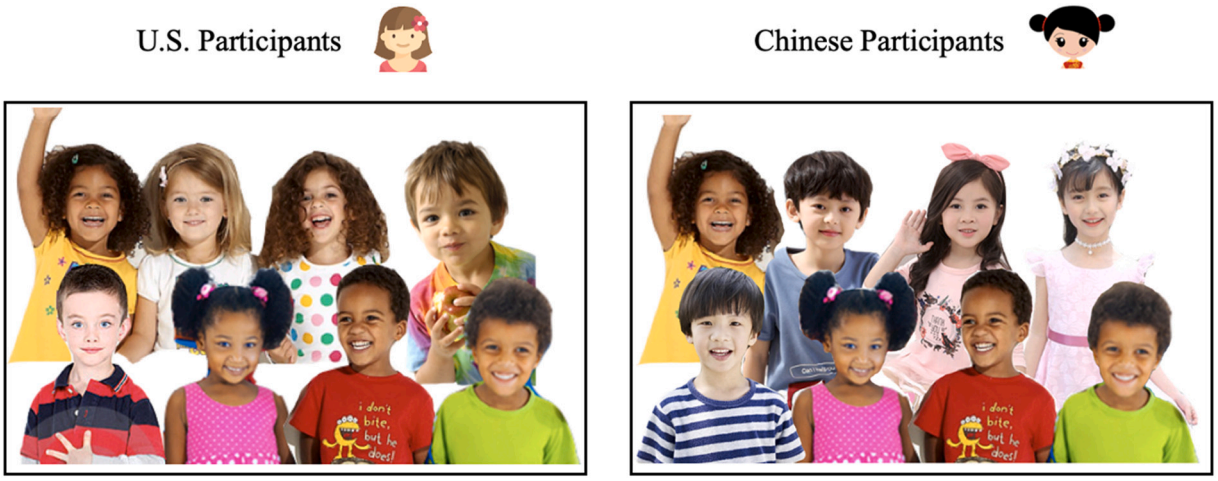


Fig. 3. Jiffy School Photo (50 % Ingroup Race-50 % Outgroup Race, Half Girls and Boys) Shown at the Beginning of Study 3.

= 5.29 years, *SD* = 0.42, range = 4.17–6.25 years) participated in Study 3. One additional Chinese child was tested, but excluded from the final sample because he answered the memory check question (see below) incorrectly. Children were recruited from the two testing sites in the same manner as in Study 1.

4.1.2. Design, materials, and procedure

The design was the same as in Study 1, except that children first saw a photo of Jiffy school children showing the racial and gender makeup of the school. Children were told, "Look! Here is a photo of the Jiffy kids!" The photo consisted of 50 % ingroup race and 50 % outgroup race, and the gender makeup was 50 % girls and 50 % boys. White US children saw a photo consisting of 2 White girls, 2 Black girls, 2 White boys, and 2 Black boys (Fig. 3). Chinese children saw a photo consisting of 2 Chinese girls, 2 Black girls, 2 Chinese boys, and 2 Black boys (Fig. 3). The mixed race-gender photo was only shown once at the beginning of the study for approximately 1 minute. Following this, children proceeded to the eight generalization trials identical to Study 1.

After the generalization trials, we asked a manipulation check question to examine whether the diverse school photo affected children's assumptions about the racial and/or gender makeup of the children in the Jiffy school. We showed children 4 pictures of groups of children, none of which was the same as the school picture shown at the beginning of the study. The 4 pictures consisted of photos of children who were 1) of mixed race and gender, 2) all ingroup race, mixed gender, 3) all ingroup race, all boys, 4) all ingroup race, all girls. Children were told, "One of these 4 pictures is a picture of the Jiffy kids," and were asked, "Which one do you think is the picture of the Jiffy kids?"

Lastly, we asked a memory check question to examine whether children remembered the photo that they saw at the beginning of the study. We showed children two pictures of groups of children: a novel photo (ingroup race, mixed gender) and the identical initial photo. Children were asked, "Do you remember which picture is the Jiffy kids' picture that I showed you at the beginning?" Children who answered the memory check question incorrectly were excluded from the final sample (only one child did so).

The memory check simply examined whether children remembered the school photo shown initially, whereas the manipulation check examined whether children generalized the diversity of the group of children in the initial photo to a new group of Jiffy children.

4.2. Results and discussion

On average, children selected the diverse pair in 48 % of the trials, which is at chance (Exact Binomial test: $P = .48$ [.42, .54]; $p = .54$). Children's selections in both race and gender trials were at chance ($P_{race} = .50$ [.41, .59], $p = 1$; $P_{gender} = .46$ [.37, .55], $p = .39$). Mixed-effects logistic regression did not reveal any effect of trial type, age, gender, or their interaction.

Because Study 3 is essentially identical to Study 1 except for the diverse school photo shown at the beginning, we next compared data from Studies 1 and 3. We used mixed-effects logistic regression to predict children's choice (Diverse = 1, Non-Diverse = 0) from study (1 or 3), trial type, age, gender, and their interactions, with random intercepts for country and participant. There was an interaction between study and trial type: in the race trials, children were more likely to select the diverse pair in Study 3 than in Study 1 ($\beta = 0.93$, SE = 0.28, $p = .001$, OR = 2.53), but in gender trials, children's choices did not differ between Study 1 and 3 ($\beta = 0.06$, SE = 0.27, $p = .84$). Thus, we found that the minimal manipulation had an effect in race trials. This was quite impressive given that children in Study 3 only saw the diverse school photo for less than a minute.

In the manipulation check question, children were asked to pick one of four pictures that showed children from the Jiffy school. If children believed that the population to be generalized (Jiffy children) were truly diverse, they should choose the photo of mixed gender-race children. Eighteen out of 35 children (51 %) chose that picture, 6 chose the ingroup race-diverse gender picture, 3 chose ingroup race-ingroup gender picture, and 8 chose the ingroup race-outgroup gender photo. The overall distribution of children choosing the 4 options differed from what would be expected from chance (Chi-Square Goodness of Fit Test: X^2 (3) = 14.49, $p = .002$). Post-hoc comparisons revealed that the proportion of children who chose the mixed gender-race picture was significantly higher than expected by chance ($p < .001$).

Altogether, results of Study 3 show that a minimal manipulation (showing children a picture of a diverse population just once) affects children's notion of the population and subsequently, their reasoning for generalization. Although children did not use the diversity rationale above chance, in race trials they were more likely to do so compared to Study 1. This suggests that the ingroup rationale we saw in Study 1 was more due to *Salience-Convenience*, rather than due to *Importance*, since children were not fixated to generalizing from the ingroup when faced with new information about the population.

5. Study 4: mixed race population, 25 % Outgroup-75 % Ingroup

To further probe whether an ingroup bias we saw in Study 1 was due to mere availability vs. thinking that ingroup is very important, we replicated Study 3, but this time with 25 % outgroup race-75 % ingroup race composition, half girls and half boys. Such a composition is probably more reminiscent to real life situations compares to the 50–50 ingroup-outgroup race composition shown in Study 3. The question is whether a lower presence of outgroup race (25 %) would be sufficient to change children's ingroup rationale.

5.1. Method

5.1.1. Participants

Thirty-six children (15 White, 21 Chinese; 17 male, 19 female; *M* = 5.40 years, *SD* = 0.37, range = 4.83–6.17 years) participated in Study 4. Children were recruited from the two testing sites in the same manner as in Study 1.

5.1.2. Design, materials and procedure

The design was the same as in Study 3, except that the composition of the Jiffy School photo that we showed children at the beginning was changed to 25 % outgroup race and 75 % ingroup race, half girls and half boys. Note that only the racial composition changed from Study 3 (50 % outgroup race and 50 % ingroup race); the gender composition remained the same. US White children saw a photo consisted of 3 White girls, 1 Black girl, 3 White boys, and 1 Black boy (Fig. 4). Chinese children saw a photo consisted of 3 Chinese girls, 1 Black girl, 3 Chinese boy, and 1 Black boy (Fig. 4). The rest of the procedure was the same as in Study 3.

5.2. Results and discussion

On average, children selected the diverse pair in 52 % of the trials, which is at chance (Exact Binomial test: $P = .52$ [.46, .58]; $p = .60$). In race trials, children chose the diverse pair in 42 % of the trials, which is marginally below chance ($P_{race} = .42$ [.33, .50]; $p = .055$). But surprisingly, in gender trials, children chose the diverse pair in 62 % of the trials, which is above chance ($P_{gender} = .62$ [.53, .70], $p = .006$).

We used mixed-effects logistic regression to predict children's choice (Diverse = 1, Non-Diverse = 0) from trial type (race trial or gender trial), age, gender, and their interactions, with random intercepts for country and participant. There was a significant effect of trial type. Children were more likely to select the diverse pairs in gender trials than in race trials ($\beta = 0.88$, SE = 0.25, $p < .001$, OR = 2.41). Next, we compared children's choices from Study 1 and 4; recall that these two studies are identical except for the diverse Jiffy school photo shown in Study 4. We used mixed-effects logistic regression to predict children's choice

Fig. 4. Jiffy School Photo (25 % Outgroup Race-75 % Ingroup Race, Half Girls and Boys) Shown at the Beginning of Study 4.

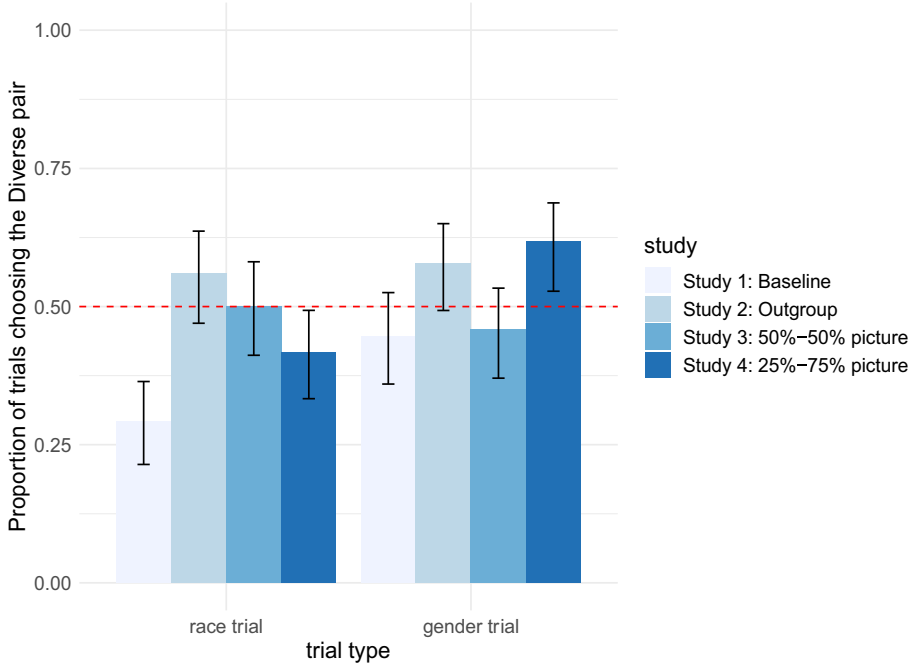


Fig. 5. Children's Selection of the Diverse Pairs by Study and Trial Type
Note. The dashed line indicates chance selection (0.5), and the error bars indicate bootstrapped 95 % CIs.

(Diverse = 1, Non-Diverse = 0) from study (1 or 4), trial type, age, gender, and their interactions, with random intercepts for country and participant. There was a main effect of study: overall, children were more likely to select the diverse pairs in Study 4 than in Study 1 (β = 0.68, SE = 0.23, p = .003, OR = 1.97). The difference between Studies 1 and 4 was statistically significant in gender trials (β = 0.80, SE = 0.32, p = .012, OR = 2.22), and marginally significant in race trials (β = 0.72, SE = 0.42, p = .087, OR = 2.06). There was also a main effect of trial type: children were more likely to select the diverse pairs in gender trials than in race trials (β = 0.80, SE = 0.18, p < .001, OR = 2.24).

In the manipulation check question, children were asked to pick one of four pictures that showed children from the Jiffy school. Nineteen out of 36 children (47 %) chose the mixed gender-race picture, 5 chose the ingroup race diverse gender picture, 5 chose ingroup race ingroup gender and 6 chose the ingroup race outgroup gender photo, one did not make a choice. The overall distribution differed from chance (Chi-Square Goodness of Fit Test: X^2 (3) = 16.09, p = .001). Post-hoc comparisons revealed that the proportion of children who chose the mixed gender-race picture is significantly higher than expected by chance (p < .001).

Lastly, we analyzed the combined data from Study 1, 2, 3, and 4. We used mixed-effects logistic regression to predict children's choice (Diverse = 1, Non-Diverse = 0) from study (1, 2, 3, or 4), trial type, age, gender, and their interactions, with random intercepts for country and participant. There was a significant interaction between study and trial type. In race trials, children were more likely to select the diverse pair in Study 2, 3, and 4 than in Study 1 (Study 2: β = 1.17, SE = 0.28, p < .001, OR = 3.25; Study 3: β = 0.92, SE = 0.28, p < .001, OR = 2.51; Study 4: β = 0.57, SE = 0.27, p = .039, OR = 1.76). In gender trials, children were more likely to select the diverse pair in Study 2 and 4 than in Study 1 (Study 2: β = 0.56, SE = 0.27, p = .04, OR = 1.75; Study 4: β = 0.73, SE = 0.27, p = .006, OR = 2.08), but their choices did not differ between

Study 3 and Study 1 ($\beta = 0.06$, SE = 0.27, $p = .83$). Overall, children were more likely to choose the diverse pair in gender trials than in race trials ($\beta = 0.69$, SE = 0.25, $p = .007$, OR = 2.00).

Similar to Study 3, the results of Study 4 suggest that showing a picture of a diverse population primed children to use more diversity rationale compared to Study 1. Interestingly, while the manipulation photo in Study 3 (50 % ingroup race, 50 % outgroup race) only changed children's reasoning in race trials, Study 4's manipulation photo (25 % outgroup race and 75 % ingroup race) only affected children's reasoning in gender trials. The photo in Study 4 had a weaker effect on children's reasoning in race trials compared to Study 3, potentially because the racial composition of the photo in Study 4 was less diverse than the photo in Study 3. However, the result in gender trials is puzzling since the gender composition of the initial photos were the same in Studies 3 and 4 (both had 50 % boys and 50 % girls). A possible explanation, which accounts for the difference between Studies 3 and 4, is that racial diversity obfuscates gender diversity, the latter coming to the foreground only in racially less diverse samples.

6. General discussion

Different social groups are often bound by common norms, laws, and policies. The process whereby these laws and policies are formulated—inferences and generalizations about the broader population—critically affects intergroup relations. At times, it can lead to conflicts, for example when there is a perceived disparity between considerations of ingroups vs. diverse groups. Here we investigated the origins of the ability to make inferences and generalizations about the broader population by asking 5-year-olds from two different cultures (US and China) to choose a toy that makes most (hypothetical) Jiffy school children happy. We distinguished two rationales for the decision: following the preference of their ingroup (*ingroup rationale*) vs. following the preference of a diverse group (*diversity rationale*). The results showed that 5-year-olds from both cultures chose to generalize the preference of their ingroup to the general population (Study 1). However, when the ingroup presence was omitted or reduced (Study 2), children were more likely to use the diversity rationale when making a generalization. Furthermore, in Studies 3 and 4 we found that the ingroup default is malleable. Simply by showing a photo with a somewhat diverse target population (Jiffy school children), children's strong bias for privileging ingroup over diversity could be reduced. To our knowledge, this is the first study that directly pits diversity-based reasoning against ingroup preference in the context of generalizing for the common good. Our results suggest that although 5-year-olds across two different cultures have a strong ingroup preference, a brief exposure to diversity can shift them away from the ingroup preference toward using diversity-based reasoning.

6.1. Ingroup rationale over diverse rationale

Why do children use the ingroup rationale as a default strategy? A priori, both rationales—ingroup and diversity—could be expected to shape children's generalization. First, a prior study shows that 5-year-olds understand diversity-based reasoning: when generalizing in the abstract (whom to ask to find out whether all children like toy X), they were more likely to use the diverse (e.g., a White child and a Black child) over the narrow evidence (e.g., two White children) (Noyes & Christie, 2016). Second, while many studies suggest that children starting at around four years already have an ingroup bias, the focus of these studies has been on tasks involving decisions affecting the *self*: maintaining one's own reputation (Engelmann et al., 2013), selecting who to be friends with (Kinzler et al., 2007; Shutts et al., 2013), or who to learn from (Master & Walton, 2013). The current research, however, explicitly asked children to make a decision for another population (the Jiffy school children), which would ostensibly not affect the child participant her/himself. Since children themselves derived neither benefit nor loss, there was no utilitarian reason to favor one's ingroup. This design was intentional, as we wanted to circumvent the complexity of personal loss-gain in social reasoning.

6.1.1. Ingroup is more salient and available

Nevertheless, in the current studies, children across two cultures with starkly contrasting characteristics—for example, along the independent (US) vs. interdependent (China) axis (Markus & Kitayama, 1991)—showed a strong tendency to make a generalized decision based on the ingroup's preference. Why is this the case? For one, the social set-up in the current study differs from that of Noyes and Christie (2016), where children were reasoning in the abstract (which group *should* one ask to make inferences about all children?). In contrast, here children were told that each group already had a *known* preference (the diverse group likes X while the narrow-ingroup likes Y), which is arguably closer to everyday social situations. As such, these two sets of results suggest that while children may understand the diversity rationale, they do not always *use* it. This discrepancy between knowledge vs. usage parallels the "knowledge-action gap" in the morality literature (Blasi, 1980; Walker, 2004), where people's moral knowledge (what they should do) is often disconnected from their moral actions (what they do in real moral situations).

Another reason why 5-year-olds in the current study did not use the diversity rationale may have to do with the *scope of generalization*—who makes up the general population. In making generalizations about the broader population, children may have used convenience or salience in deciding the scope of generalization. A priori, there is no unequivocally correct scope of generalization and the same sample can be taken to represent many different populations. Schematically, one may opt for (a) a truly general population (all people), (b) one's ingroup (people similar to myself), (c) one's social environment, or other hybrid or more restrictive options. Taking an example of a Black child living in a majority-White neighborhood, she could take the general population to comprise Black and White people (truly general), majority Black (ingroup), or majority White (social environment but not ingroup).

Both our US and Chinese participants live in environments which conflate options (b) and (c) above: the ingroup race members form the overwhelming majority. As such, it is possible that when children were reasoning about what toy to buy for the Jiffy school children, their ingroup members were more salient or more convenient to access. This *Salience-Convenience* hypothesis is akin to availability heuristics (Tversky & Kahneman, 1974) in that ingroup members are more available and obvious. In the introduction, we distinguished Salience-Convenience from ingroup importance as two possible motors for preferring one's ingroup. The results of Studies 2–4 are more consistent with the *Salience-Convenience* hypothesis because they show a malleability of children's preference. When we removed the ingroup entirely from the race trials in Study 2, and when we showed children that the population contained diverse groups in Studies 3 and 4—thereby making the diverse group more available—children's calculation of the scope of generalization changed. Across Studies 2–4, children were significantly more likely to generalize the preference of the diverse group compared to their tendency to do so in Study 1.

Our findings parallel another set of findings showing that statistical and social information can change children's scope of generalization for individuals' preferences. Past studies have shown that children are reluctant to generalize an individual's preference to other individuals (e.g., Kalish, 2012; Repacholi & Gopnik, 1997), and when they do, they only generalize preferences within social categories (e.g., Birnbaum et al., 2010; Heyman & Gelman, 2000; Shutts et al., 2009, 2011). However, when strong statistical and social cues were provided, children generalized preferences even across social categories (Diesendruck et al., 2015).

6.1.2. Gender distribution

Children in our study grow up in mixed gender environments or at

least attend mixed gender preschools. Hence, we should expect them to think that the general population contains gender-diverse groups. Indeed, while there was a strong bias for using ingroup rationale in race trials of Study 1, this tendency was weaker in gender trials. Children selected between the ingroup pairs and the diverse pairs at chance in gender trials. That is, children were not using diversity rationale either. One possible explanation for this is that children's micro-environment—other children they play with, those they consider friends—is more single gender (Albert & Porter, 1983; La Freniere et al., 1984; Leaper, 1994; Maccoby & Jacklin, 1987; Martin et al., 1999; Martin & Fabes, 2001; Yee & Brown, 1994). Cultural norms can further shape this tendency; in both Chinese and US cultures, children are encouraged to value gender roles and subscribe to binary gender norms (Borowski et al., 2021;Fouts et al., 2021 ; Shu, 2005). For example, when choosing toys for children, some parents and teachers think that it is only proper for girls to gravitate toward "feminine" items and boys to gravitate toward "masculine" items (Fouts et al., 2021; Shu, 2005). There is evidence that as young as 3 years, children have absorbed cultural messages about gender in developing their gender identity (Hoffmann & Powlishta, 2001). By age 5—the children in our studies—it is possible that cultural norms about gender roles have reinforced gender segregation, affecting outgroup generalization.

6.2. Ingroup default is malleable

Despite defaulting to ingroup bias when making generalizations about the broader community, children from both cultures were able to change their rationale when faced with new evidence. In Studies 3 and 4, after being shown *one* photo of mixed gender-race Jiffy school children for a mere one minute, children were more likely to generalize the preference of the diverse group compared to those who did not see such a photo (children in Study 1). This finding aligns with the theory that increased intergroup contacts benefit intergroup relations (Allport, 1954; Pettigrew & Tropp, 2005). It adds to the body of evidence, which shows that daily intergroup contact among children (attending heterogenous or integrated schools) can improve their attitudes toward cross-race friendships (McGlothlin & Killen, 2010) or decrease children's essentialist beliefs about ethnic categories (Deeb et al., 2011). While our studies did not directly probe the effect of children attending heterogenous vs. homogenous schools, it is encouraging to see that even for children growing up in an environment where their race ingroup is the majority, children's decisions for the broader community do not dogmatically adhere to the ingroup bias. This finding is especially interesting in light of the Chinese children's social environment, where they interact with the children of the same race practically 100 % of the time. In addition to the homogenous racial composition during normal times, our data collection took place during the pandemic (end of 2021), during which China has closed its borders to nearly all foreigners beginning in March 2020.

Our finding that children's ingroup bias can be changed just by a simple diversity priming raises the important issue of real-life intervention. In some parts of the world, attending an integrated or heterogenous school is not an option due to policy, economic, or social demographic reasons. China's population is a case in point, with 95 % Han Chinese population and a non-Chinese population of around 0.06 % (National Bureau of Statistics, 2020). Under this circumstance, the only scalable channel for increasing the acceptance of diversity and intergroup relations is by exposure rather than prolonged contact, as prescribed by Allport (1954). Our findings hint that even a sporadic exposure may be helpful for increasing the salience of diversity.

6.2.1. Scope of ingroup vs. scope of population

The explanation we offer for why even brief exposure affects children's reasoning for the broader community is that it changes the scope of generalization—the coverage of the target population. One might also ask whether the exposure can redraw the lines of the ingroup. Recent works with young children have shown that social inclusion, understood as granting other children a perceived membership in one's ingroup, is relatively difficult to come by among young children (Mulvey et al., 2018; Toppe et al., 2019, 2020, 2021). One study even found that while 3–7-year olds understood that someone would feel hurt when socially excluded, they were not more likely to include that socially excluded person in their ingroup activities (Stengelin et al., 2022). Our results suggest that changing the scope of the population under consideration is easier to effect in young children than is social inclusion, perhaps because it is a change less emotional and more statistical in nature, one which children are certainly capable of doing (Denison & Xu, 2019; Kushnir et al., 2010; Xu & Garcia, 2008). We find that if changing the scope of reasoning for the broader community is a goal then a simple intervention might suffice—a finding that is encouraging, for example, for the recent series of works in adults, which advocate a more diverse news distribution as a means to break the "filter bubble" problem (e.g., Fletcher et al., 2020; Helberger, 2019; Loecherbach et al., 2020; Nielsen, 2021). More research is needed given that the sample size of our study is relatively small. But it is encouraging that children from two very different cultures (US vs. China), yet sharing analogous social make-ups (their ingroup is the majority population), reason about diversity similarly.

6.3. Conclusion

It is important to stress that the task we gave to our child participants has no one correct answer. Balancing the benefit of the majority against consideration of minorities is as old as history itself, and is a central theme in both political theory and philosophy. Indeed, the ingroup rationale (as opposed to the diversity rationale) is not without a logical basis: most people would agree that providing an essential service to the majority (say, health care) at the cost of a minor inconvenience to a narrow minority (higher taxes for the extremely rich) is justified. In adults, one might expect this type of argument to form a declarative foundation for preferring one's ingroup in reasoning about the common good.

Our study was designed to circumvent these types of complications. First, we studied responses of children, whom we expect to have fewer entrenched biases either way. Second, we left the preference of the ingroup and the diverse group deliberately open to some interpretation. For example, we did not say by how much the ingroup preferred their best toy over the toy favored by the diverse group; children were free to surmise that the differential preference may be marginal or absolute. In this way, we were asking children to project their own idea of the general population, rather than perform a utilitarian computation. Third, the child participants did not stand to gain or lose from either decision; they were choosing toy for others and not for themselves. This disentangles the results from a potential conflating factor of personal gain/loss. Finally, we tested children at two locales, which are different in nearly every aspect (culture, language, urban/suburban, etc.) except one: both are very homogeneous. The results from both testing sites showed no major differences, attesting to the robustness of our results. In the end, we believe that our findings faithfully reflect children's own ideas about the scope of the population.

Overall, the results suggest that when ingroup and diversity rationale are directly pitted as bases for generalization for the common good, 5-year-olds' default is to privilege the ingroup rationale. This default, however, does not seem to stem from a deep conviction about ingroup's importance. Rather, thinking about ingroup is more convenient and/or salient. Consequently, simple perceptual exposure about the diversity of the population is sufficient to allow children to move away from the ingroup rationale. Future studies should investigate how to leverage this finding in everyday situations, to increase usage of diversity rationale in intergroup interactions.

Declaration of competing interest

All authors declare that they have no conflicts of interest.

Data availability

We have shared the link to our data in our paper.

Acknowledgements

We thank Alexander Noyes and Cynthia Siego for conducting early pilot work of this study. We also thank Abigail Robinson, Kiera Parece, and other members of Swarthmore College Cognition and Development Laboratory for help in recruiting participants; the children, families, schools, and organizations who participated in this study. Funding was provided through the Guoqiang Foundation to SC (2019GQB0001) and Swarthmore College Summer Research Fellowships to RL.

References

Albert, A. A., & Porter, J. R. (1983). Age patterns in the development of children's gender-role stereotypes. *Sex Roles, 9*, 59–67. https://doi.org/10.1080/152483707018366261O.1007/BF00303110

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

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science, 348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science, 17*(2), 159–163. https://doi.org/10.1111/j.1467-9280.2006.01679.x

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science, 17*(1), 53–58. https://doi.org/10.1111/j.1467-9280.2005.01664.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: Development of social essentialism. *Child Development, 81*(3), 757–777. https://doi.org/10.1111/j.1467-8624.2010.01432.x

Bishop, B. (2009). *The big sort: Why the clustering of like-minded America is tearing us apart.* Houghton Mifflin Harcourt.

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin, 88*(1), 1–45. https://doi.org/10.1037/0033-2909.88.1.1

Borowski, S. K., Spiekerman, A. M., & Rose, A. J. (2021). The role of gender in the friendships of children and adolescents. In *Reference module in biomedical sciences.* Elsevier. https://doi.org/10.1016/B978-0-12-818872-9.00044-3.

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*(6), 1139–1156. https://doi.org/10.1037/a0026107

Denison, S., & Xu, F. (2019). Infant statisticians: The origins of reasoning under uncertainty. *Perspectives on Psychological Science, 14*(4), 499–509. https://doi.org/10.1177/1745691619847201

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*(2), 370–380. https://doi.org/10.1080/15248372.2013.848870

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "Minimal" group affiliations in children: Minimal group affiliations in children. *Child Development, 82*(3), 793–811. https://doi.org/10.1111/j.1467-8624.2011.01577.x

Engelmann, J. M., Over, H., Herrmann, E., & Tomasello, M. (2013). Young children care more about their reputation with ingroup members and potential reciprocators. *Developmental Science*, 952–958. https://doi.org/10.1111/desc.12086

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford University Press.

Fletcher, R., Cornia, A., & Nielsen, R. K. (2020). How polarized are online and offline news audiences? A comparative analysis of twelve countries. *The International Journal of Press/Politics, 25*(2), 169–195. https://doi.org/10.1177/1940161219892768

Fouts, H. N., Bader, L. R., Neitzel, C. L., & Salinas, D. A. (2021). Ethnicity as a predictor of gender segregation among young children in an informal urban settlement in Kenya. *Social Development, 30*(3), 713–729. https://doi.org/10.1111/sode.12503

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*(2), 159–174. https://doi.org/10.1080/152483707018366261O.1006/jecp.1996.2344

Heit, E. (2000). Properties of inductive reasoning. *Psychonomic Bulletin & Review, 7*, 569–592. https://doi.org/10.1080/152483707018366261O.3758/BF03212996

Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism, 7*(8), 993–1012. https://doi.org/10.1080/21670811.2019.1623700

Heyman, G. D., & Gelman, S. A. (2000). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology, 77*(1), 1–19. https://doi.org/10.1006/jecp.1999.2555

Hoffmann, M. L., & Powlishta, K. K. (2001). Gender segregation in childhood: A test of the interaction style theory. *The Journal of Genetic Psychology, 162*(3), 298–313. https://doi.org/10.1080/00221320109597485

Kalish, C. W. (2012). Generalizing norms and preferences within social categories and individuals. *Developmental Psychology, 48*(4), 1133–1143. https://doi.org/10.1037/a0026344

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers: Selective trust in native-accented speakers. *Developmental Science, 14*(1), 106–111. https://doi.org/10.1111/j.1467-7687.2010.00965.x

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences, 104*(30), 12577–12580. https://doi.org/10.1073/pnas.0705345104

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology, 115*(2), 405–450. https://doi.org/10.1086/599247

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science, 21*(8), 1134–1140. https://doi.org/10.1177/0956797610376652

La Freniere, P., Strayer, F. F., & Gauthier, R. (1984). The emergence of same-sex affiliative preferences among preschool peers: A developmental/ethological perspective. *Child Development, 55*(5), 1958–1965. https://doi.org/10.1080/152483707018366261O.2307/1129942

Leaper, C. (1994). Exploring the consequences of gender segregation on social relationships. In C. Leaper (Ed.), *Childhood gender segregation: Causes and consequences* (pp. 67–86). San Francisco: Jossey-Bass. https://doi.org/10.1080/152483707018366261O.1002/cd.23219946507.

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*(4), 388–404, 10.1080/152483707018366261O.1080/13546780903343227.

Loecherbach, F., Moeller, J., Trilling, D., & van Atteveldt, W. (2020). The unified framework of media diversity: A systematic literature review. *Digital Journalism, 8*(5), 605–642. https://doi.org/10.1080/21670811.2020.1764374

López, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development, 63*(5), 1070–1090. https://doi.org/10.2307/1131519

Maccoby, E. E., & Jacklin, C. N. (1987). Gender segregation in childhood. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 20, pp. 239–287). Academic Press. https://doi.org/10.1016/S0065-2407(08)60404-8.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Martin, C. L., & Fabes, R. A. (2001). The stability and consequences of young children's same-sex peer interactions. *Developmental Psychology, 37*(3), 431. https://doi.org/10.1080/152483707018366261O.1037/0012-1649.37.3.431

Martin, C. L., Fabes, R. A., Evans, S. M., & Wyman, H. (1999). Social cognition on the playground: Children's beliefs about playing with girls versus boys and their relations to sex segregated play. *Journal of Social and Personal Relationships, 16*(6), 751–771. https://doi.org/10.1080/152483707018366261O.1177/0265407599166005

Master, A., & Walton, G. M. (2013). Minimal groups increase young children's motivation and learning on group-relevant tasks. *Child Development, 84*(2), 737–751. https://doi.org/10.1111/j.1467-8624.2012.01867.x

McGlothlin, H., & Killen, M. (2010). How social experience is related to children's intergroup attitudes. *European Journal of Social Psychology, 40*(4), 625–634. https://doi.org/10.1002/ejsp.733

Mulvey, K. L., Boswell, C., & Niehaus, K. (2018). You don't need to talk to throw a ball! Children's inclusion of language-outgroup members in behavioral and hypothetical scenarios. *Developmental Psychology, 54*(7), 1372–1380. https://doi.org/10.1037/dev0000531

National Bureau of Statistics. (2020). 2020 population census of China. Retrieved from http://www.stats.gov.cn/tjsj/pcsj/.

Nielsen. (2021, September 9). *Being seen on screen 2021: The importance of quantity and quality representation on TV*. Retrieved from. The Nielsen Company (US), LLC https://www.nielsen.com/us/en/insights/report/2021/being-seen-on-screen/.

Noyes, A., & Christie, S. (2016). Children prefer diverse samples for inductive reasoning in the social domain. *Child Development, 87*(4), 1090–1098. https://doi.org/10.1080/152483707018366261O.1111/cdev.12522

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97*, 185. https://doi.org/10.1080/152483707018366261O.1037/0033-295X.97.2.185

Pettigrew, T. F., & Tropp, L. R. (2005). Allport's intergroup contact hypothesis: Its history and influence. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 262–277). Blackwell Publishing. https://doi.org/10.1002/9780470773963.ch16.

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception, 31*(9), 1109–1121. https://doi.org/10.1068/p3331

Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of Homo sapiens: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences, 115*(45), 11401–11405. https://doi.org/10.1073/pnas.1721165115

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*(1), 12–21. https://doi.org/10.1037/0012-1649.33.1.12

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/10.1080/15248370701836626 10.1111/j.1551-6709.2010.01137.x

Rhodes, M., Brickman, D., & Gelman, S. A. (2008). Sample diversity and premise typicality in inductive reasoning: Evidence for developmental change. *Cognition, 108*, 543–556. https://doi.org/10.1080/15248370701836626 10.1016/j.cognition.2008.03.002

Rhodes, M., Gelman, S. A., & Brickman, D. (2008). Developmental changes in the consideration of sample diversity in inductive reasoning. *Journal of Cognition & Development, 9*, 112–143. https://doi.org/10.1080/15248370701836626

Shilo, R., Weinsdörfer, A., Rakoczy, H., & Diesendruck, G. (2019). The out-group homogeneity effect across development: A cross-cultural investigation. *Child Development, 90*(6), 2104–2117. https://doi.org/10.1111/cdev.13082

Shu, X. (2005). Market transition and gender segregation in urban China. *Social Science Quarterly, 86*(s1), 1299–1323. https://doi.org/10.1111/j.0038-4941.2005.00347.x

Shutts, K., Banaji, M. R., & Spelke, E. S. (2011). Social categories guide young children's preferences for novel objects: Social categories guide preferences. *Developmental Science, 13*(4), 599–610. https://doi.org/10.1111/j.1467-7687.2009.00913.x

Shutts, K., Kinzler, K. D., McKee, C. B., & Spelke, E. S. (2009). Social information guides infants' selection of foods. *Journal of Cognition and Development, 10*(1–2), 1–17. https://doi.org/10.1080/15248370902966636

Shutts, K., Roben, C. K. P., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development, 14*(1), 35–62. https://doi.org/10.1080/15248372.2011.638686

Stengelin, R., Toppe, T., Kansal, S., Tietz, L., Sürer, G., Henderson, A. M. E., & Haun, D. B. M. (2022). Priming third-party social exclusion does not elicit children's inclusion of out-group members. *Royal Society Open Science, 9*(1), Article 211281. https://doi.org/10.1098/rsos.211281

Sunstein, C. R. (2009). *Going to extremes: How like minds unite and divide*. Oxford University Press.

Toppe, T., Hardecker, S., & Haun, D. B. M. (2019). Playing a cooperative game promotes preschoolers' sharing with third-parties, but not social inclusion. *PLOS ONE, 14*(8), Article e0221092. https://doi.org/10.1371/journal.pone.0221092

Toppe, T., Hardecker, S., & Haun, D. B. M. (2020). Social inclusion increases over early childhood and is influenced by others' group membership. *Developmental Psychology, 56*(2), 324–335. https://doi.org/10.1037/dev0000873

Toppe, T., Hardecker, S., Zerres, F., & Haun, D. B. M. (2021). The influence of cooperation and competition on preschoolers' prosociality toward in-group and out-group members. *Royal Society Open Science, 8*(5), Article 202171. https://doi.org/10.1098/rsos.202171

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131. http://www.jstor.org/stable/1738360.

Walker, L. J. (2004). Gus in the gap: Bridging the judgment-action gap in moral functioning. In D. K. Lapsley, & D. Narvaez (Eds.), *Moral development, self, and identity* (pp. 1–20). Mahwah, NJ: Erlbaum.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences, 105*(13), 5012–5015. https://doi.org/10.1073/pnas.0704450105

Yee, M., & Brown, R. (1994). The development of gender differentiation in young children. *British Journal of Social Psychology, 33*(2), 183–196. https://doi.org/10.1111/j.2044-8309.1994.tb01017.x