Overhearing Brief Negative Messages has Lasting Effects on Children's Attitudes Toward Novel Social Groups

Emily B. Conder & Jonathan D. Lane

Vanderbilt University, Peabody College, Department of Psychology & Human Development

Corresponding author:
Emily B. Conder, Vanderbilt University, Peabody College, Psychology and Human Development, 230 Appleton Place, PMB 552, Nashville, TN 37203.
Email: emily.b.conder@vanderbilt.edu

Acknowledgments
This work could not have been accomplished without the help of many others. We thank Steven Roberts and Josh Rottman for their thoughtful feedback on earlier drafts of this manuscript. We also thank the three anonymous reviewers for their feedback on the submitted manuscript. This research was conducted with the help of a team of hard-working research assistants and staff, especially Taylor Boothby, Marlotte de Jong, Carrie Sherry, Janelle Brown, Charlotte Buck, Rachel Calvosa, Leoncia Gillespie, Nicolette Granata, Mackenzie Jackson, Noora Jamal, Shreya Karak, Mary Tezak, Megan Wiebe, Alexandra Wolfe, and Anna Yarinsky. We are also grateful to the actresses who were willing to be a part of our study stimuli, Zoey Saylor, Mira D'Agostino, Allison Vreland, and Susannah Sutherland. We would like to thank the parents who volunteered their time to bring their children into our lab to participate in this study and the schools that allowed us to recruit children from their classrooms, Saint Ann School, Saint Bernard Academy, Explore Community School, and the Eighteenth Avenue Family Enrichment Center.

This is a pre-print version of a manuscript accepted for publication in *Child Development*.

Abstract

Societies are rife with out-group discrimination and mistreatment. One way that children might acquire social biases that lead to such outcomes is by overhearing derogatory or disparaging comments about social groups. Children (*n*=121) overheard a video call between a researcher and an adult or child caller who made negative claims (or no claims) about a novel social group. Immediately and following a two-week delay, older children (7-9 years) who overheard the message demonstrated stronger negative attitudes toward the group than children who heard no message. Younger children's (4-5-year-olds') attitudes were generally unaffected by these claims. Thus, overhearing brief, indirect messages from children or adults had robust and lasting effects on the social biases of children 7-years and older.

Keywords: Social Cognition, Intergroup Attitudes, Social Learning

Overhearing Brief Negative Messages has Lasting Effects on Children's Attitudes Toward Novel Social Groups

In recent years, many regions of the world have experienced substantial increases in the expression of intergroup biases. In the U.S., hate crimes involving race, ethnicity, religion or sexual orientation increased 17% from 2016 to 2017 (Federal Bureau of Investigation, 2017), and the number of hate groups rose 30% from 2017 to 2018 (Shanmugasundaram, 2018). The development of intergroup biases may be a common feature of human social cognition. Even young children typically demonstrate preferences towards in-groups relative to other social groups (Aboud, 2003; Nesdale, 2004; Raabe & Beelman, 2011). However, significant individual differences in the valence and magnitude of these biases suggest that socialization influences their development (Over & McCall, 2018). One way in which children's intergroup attitudes may be socialized is through information that they encounter in others' messages. Indeed, much of human conversation is focused on other people (Dunbar, 2004). Thus, children may have many opportunities to hear people make derogatory or disparaging comments about social groups, for example within others' conversations or from messages propagated on electronic media. We investigate how *overhearing* derogatory messages from an unfamiliar speaker—indirectly, from a caller on a nearby video chat—influences children's attitudes towards other social groups, and we explore potential developmental differences in the force of such messages between early and middle childhood.

During the preschool years, biases favoring one's in-group emerge as children learn about social categories that are prominent in their society (Nesdale, 2004). Preferences for one's own ethnic, racial, or national group typically strengthen throughout early childhood, peak between 5 and 7 years, and stabilize between 8 and 10 years (Raabe & Beelmann, 2011; Skinner

& Meltzoff, 2019). Experimental studies reveal that, during early and middle childhood (3-9 years), children prefer to be friends with, allocate more resources to, and like others who share their gender, race, ethnicity, or language, compared to those who do not (Kinzler, Shutts, DeJesus, & Spelke, 2009; Nesdale, Maass, Griffiths, & Durkin, 2003; Renno & Shutts, 2015). Similar biases are elicited when children are assigned to arbitrary 'minimal' groups—e.g., children are randomly assigned to either a "red" or "blue" group (Dunham, 2018). Immediately following assignment, 6-year-olds exhibit biases (Dunham & Emory, 2014), and after *repeated* exposures preschoolers (and older children, see Bigler, Jones & Lobliner, 1997) also demonstrate biases. In one study, preschoolers whose teachers emphasized differences between minimal groups over several weeks (e.g., by asking children to line-up based on their group's shirt color) more often wanted to play with other children wearing their group's color (compared to children whose teachers disregarded group membership; Patterson & Bigler, 2006). Collectively, this research suggests that intergroup biases can be hair-triggered and are perhaps typical of social-cognitive development. According to developmental intergroup theory (Bigler & Liben, 2007), children are prone to categorize other people into groups using perceptually-discriminable cues (e.g., gender, race, clothing). However, experience influences whether such cues become psychologically salient to children, and whether children associate these cues with specific attributes (e.g., traits, behaviors, roles) or affect (e.g., liking).

Children's intergroup attitudes may also be influenced by the opinions that others express about social groups. Experimental studies have revealed how messages provided directly to children can influence their intergroup attitudes. In one experiment, children ages 4-7 years who were told about a novel-group member who behaved anti-socially later predicted that a *different* member from that group would behave anti-socially (Baron, Dunham, Banaji, & Carey, 2014). In

another study, Jewish-Israeli 6-year-olds drew an Arab character further away from a Jewish character after they heard a story emphasizing ethnic essentialism, in comparison to children who did not hear the story (Diesendruck & Menahem, 2015). Hearing *positive* claims about a Black character decreased the pro-White bias (measured with the Child Implicit Associations Task) of Asian and White 8- to 12-year-olds, but those claims had no effect on the bias of 4- to 7-year-olds (Gonzalez, Steele, & Baron, 2017).

This emerging body of work reveals that direct messages about out-groups can influence children's attitudes toward those groups, and this influence seems to strengthen between early and middle childhood. However, not everything that children learn is directly taught to them. For example, messages that parents provide to their children about other social groups are generally neutral or positive (Cabrera, Kuhns, Malin, & Aldoney, 2016), yet correlational studies reveal associations between parents' and children's intergroup attitudes (Degner & Dalege, 2013). This suggests that parents may be unintentionally expressing negative biases around their children (or that both children and their parents are exposed to similar biased information). One way in which children (and adults) might be exposed to such biased information is by *overhearing* negative messages about social groups, e.g., from others' in-person conversations, video-chat conversations, or from pundits expressing their opinions on the radio or television. Indeed, after the 2016 Presidential election, elementary school children reported hearing and remembering Donald Trump's negative statements about immigrants and women (Patterson, et al., 2019). Studies have revealed that even toddlers can learn about novel words or tool use by overhearing conversations between adults who are nearby or on video (Akhtar, 2001, 2005; Boderé & Jaspaert, 2017; Floor & Akhtar, 2003; Gampe, Liebal, & Tomasello, 2012; O-Doherty et al.,

2011; Shneidman & Woodward, 2015). Preschoolers can also learn spatial information (the location of objects in a model) by overhearing adults' conversations (Foster & Hund, 2012).

Across early and middle childhood, children are increasingly capable of learning from overhearing others' conversations. For example, when 3- to 9-year-olds overheard video-recorded adults speak about the non-existence of a novel animal, there was an age-graded increase in children's tendency to believe the animal was indeed not real (Woolley, Ma, & Lopez-Mobilia, 2011). Likewise, 5-year-olds (but not 3-year-olds) were more likely to cheat on a guessing game (by peeking at a hidden card) after they overheard an adult describe the previous participant as "smart" (Zhao et al., 2019). Only one project (to our knowledge) has examined whether young children may *develop biases* about social groups by overhearing others' negative claims about those groups. In studies by Lane, Conder, and Rottman (2020), children ages 4- to 9-years played a game with a researcher who made derogatory claims about a novel group of people (e.g., "Gearoos are really bad people, they eat disgusting food and they wear such weird clothes, the Gearoos' language sounds so ugly"). In one condition, the researcher spoke directly to the child, making eye contact while uttering these statements. In another condition, the researcher received a cellphone call, turned to their side, and uttered these statements to the supposed caller (so that the child could *overhear* the researcher). In both conditions, children who heard the researcher's negative claims rated the novel group as being "less good", were less willing to try elements of the group's culture, less often wanted to be friends with a group member, and drew themselves further away from a member of the group, than children who heard no claims. Critical to the current study, these effects existed even for overheard claims (but to a lesser extent than direct claims).

In the current study, we focused on a period of development (4-9 years) when children's intergroup biases rapidly develop (e.g., Nesdale, 2004; Raabe & Beelmann, 2011). We expand upon the methods of Lane et al. (2020) and address additional questions about how derogatory, disparaging social messages influence children's intergroup attitudes. We exclusively and more thoroughly investigate the effects of *overhearing* claims on children's intergroup attitudes, and those claims are presented in a very different manner than in prior work. In contrast to Lane et al. (2020), who tested for the effects of hearing negative claims uttered by someone whom children had established rapport with and who was in the *same room* as the child, the current study explores the effects of negative messages provided by a complete *stranger* who is *not in the room* with the child and who is never introduced to the child. Specifically, children overhear a video-chat conversation in which a non-present, unfamiliar caller makes brief, negative claims about a novel social group. In control conditions, children overhear the same conversation, but claims about the group are omitted. In both conditions, children's attitudes toward the group are assessed using direct questions and indirect measures. Thus, the current methods mimic how children might overhear strangers (e.g., news anchors or politicians) make statements about social groups on electronic media, or how children might overhear comments made by parents' friends or co-workers during video calls. Another novel question that we begin to address is whether the effects of these messages differ depending upon the speaker's age (described in greater detail later). Finally, and critically, unlike past work, we evaluate the short-term *longitudinal effects* of these claims by measuring children's attitudes again, two-weeks after overhearing the video call.

The messages that children overhear in the current study are brief and uttered by absent, unknown persons; thus, they may simply constitute ambient noise without affecting the

intergroup attitudes of children at any age. We anticipated that claims would have little to no effect on the youngest children's (4-6 year-olds') attitudes, given prior findings that overheard messages (uttered by someone *in* the room *with* the child) had only modest and inconsistent effects on 4- to 6-year-olds' attitudes (Lane et al., 2020). As discussed earlier, children's capacity or tendency to learn by overhearing conversations increases during early childhood and into middle childhood (Dunham & Emory, 2014; Foushee & Xu, 2016; Lane et al., 2020; Woolley & Lopez-Mobilia, 2011; Zhao et al., 2019). Thus, we anticipated that older children who had overheard negative claims about a novel social group (compared to children who had not) would express stronger negative attitudes toward the group across all measures. We predicted that these effects would exist immediately after the video call and would maintain (to a lesser extent) following a two-week delay (for other work demonstrating that the effects of learning from others' claims may decrease over time, see Ronfard, Lane, Wang, & Harris, 2017; Rottman, Young, & Kelemen, 2017).

As noted earlier, as a supplementary question, we also explore whether the speaker's age moderates the force of negative messages on children's intergroup biases. Previous studies on this topic have employed only adult speakers (e.g., Lane et al., 2020). Yet, in their everyday lives, children may also hear other children discuss social groups. Thus, to test whether the influence of derogatory messages differs depending on the age of the speaker, we vary whether the video caller is a child or an adult. Even preschoolers appreciate that children and adults possess different types of expertise (Jaswal & Neely, 2006; Taylor, Cartwright, & Bowden, 1991), and as U.S. children transition into middle childhood, they spend more time with their peers (Eccles, 1999). Thus, older children *might* devote greater attention to or grant greater credence to other children's (vs. adults') claims about the social world. Consistent with this

reasoning, we predicted that messages uttered by a child (vs. an adult) caller would more powerfully influence children's intergroup attitudes, particularly among the oldest participants.

Method

Participants

Children (N = 134) ranging from 3.82 and 9.13 years (48% female, M_{age} = 6.40 years, SD = 1.23 years) were recruited from and tested in local charter schools in Nashville, TN (n = 118) or were recruited through Tennessee state birth records and tested in a campus laboratory (n = 16) between April 2017 and September 2018. Data were collected from all families that provided consent to participate, even after we met the recruitment goal determined by power analyses (described later). The study consisted of two sessions that were intended to take place approximately 2 weeks (14 +/- 5 days) apart. Six children did not participate in the second study session (ages 4.70, 6.34, 6.52, 6.68, 8.16, 8.20 years), one child was eliminated from analysis because of experimenter error at the second session (6.60 years), and one child was eliminated because their parent discussed the novel social group with them between the two study sessions (7.98 years). An additional 5 children were excluded because (due to scheduling logistics) the time between their two interviews was either fewer than 9 days (n = 3; 5.15, 5.16, 8.81 years) or more than 19 days (n = 2; 4.56, 6.78 years). (Ad-hoc analyses requested by a reviewer revealed that similar results are obtained whether or not these five participants are included in analyses. Partial correlations between delay length and children's evaluations of the novel group during the delayed session are reported in Supplementary Materials, Table S8). The final sample consisted of 121 children (46% female; M_{age} = 6.39 years, SD = 1.22 years).

We initially recruited a sample large enough to fulfill the requirements for an *a priori* power analysis in which age was a categorical, between-subjects variable (with two levels:

younger vs. older), consistent with much of the developmental literature on this topic (reviewed in the introduction). In previous work, the effects of social messages on children's intergroup attitudes have varied in magnitude, typically ranging from small-medium to medium-large (e.g., Diesendruck & Menahem, 2015; Gonzalez et al., 2017; Lane et al., 2020). Thus, we conducted a power analysis (using G*Power 3.1; Faul, Erdfelder, Lang, & Buchner, 2007) to determine the sample size necessary to detect medium sized main effects and interaction effects (e.g., fs $\geq .25$) for 2x2x2x2 repeated-measures ANOVAs (for the continuous dependent variables) and multi-level logistic regressions for the dichotomous dependent variable. In these analyses, desired statistical power was set to the recommended .80 level, $\alpha = .05$ (Cohen, 1992), the anticipated correlations among repeated measures was set to .5, and the non-sphericity correction was set to 1. This analysis identified that a sample of at least 98 participants was required. Consent forms were distributed to classrooms in mass, and the consent return rate was greater than anticipated. We interviewed all children whose parents provided consent and who were available to participate. Our final sample of 121 children exceeded the minimum sample size identified by our power analyses.

By the time data collection was completed, there was an increasing push in the developmental literature to analyze age as a continuous variable whenever possible. Thus, in response to this push, we report analyses in which age is a *continuous* (rather than a dichotomous) variable, using a series of hierarchical regressions for the continuous dependent variables and hierarchical logistic regressions for the dichotomous dependent variable.

Importantly, the results reported herein replicate those achieved with the initially planned ANOVAs and regressions.

Parents of children recruited from local charter schools had previously completed a consent form given to them by their child's teacher, and parents of children in the campus lab completed the form when they arrived with their child for the study. After the initial session, children who participated in their school were given a thank-you note to take home to their parents, which included a reminder to not discuss the purpose of the study with their child until the end of the second session. For children who participated in the lab, parents were given the same reminder on a note attached to the consent form.

All parents were asked to complete a voluntary questionnaire with demographic questions. Of the parents in our final sample (n = 121), 67% identified as White, 19% Black, 3% Asian, and 11% identified with multiple races or ethnicities. Self-reported education was fairly high: 16% of parents who completed the survey had a Doctorate, 17% a Master's degree, 37% a Bachelor's degree, 24% completed some college, 3% had a high school diploma, and 3% had not completed high school. Demographic data were not collected directly from or about child participants.

Procedure

Children participated individually with a female experimenter (E) either in a laboratory room on a local college campus or in a hall or room in the child's school. In the lab, sessions were video recorded using cameras hidden from view in decorative plants; in schools, sessions were video recorded with cameras hidden inside of a pencil box (22.9 cm x 6.2 cm) that had a small hole for the camera lens. Only children whose parents consented were video recorded. Children ages 6 years and younger provided verbal assent to participate, and children older than

6 years signed a written consent form. Children were told that they would play a picture-finding game (really, a filler task) and that later they would play another game on a nearby, open laptop computer. During the picture-finding game, E received a pre-recorded Skype video call alert (on the computer) from either an adult or child caller. During the call, the caller either did or did not utter a negative message about a novel social group (referred to as 'Flurps' or 'Gearoos'). Children in this age range may generalize negative messages about a group to individual members of that group (Roberts, Ho, & Gelman, 2019). Thus, to assess children's attitudes toward the group after the call, children completed tasks to evaluate their impressions of members of the novel group using both direct and indirect questions about the group or members of the group.

Approximately two weeks later (*M* = 13.64 days; *Range:* 9 – 19 days; average time between sessions did not significantly differ by children's age or condition), children were interviewed a second time. An average delay of two weeks was chosen because longer time periods between sessions would not have been feasible given the constraints of collecting data from and debriefing entire classrooms of children during the school year. To limit the possibility that children would simply perseverate in their answers across the two testing periods (to the extent that children remembered their earlier answers), a different experimenter (who was not present at all during the first testing session) worked with each child during the delayed session. During this session, children completed all dependent measures for a second time, in the same order as they had during the initial session. However, there was no filler task (picture-finding game), no video call, and no new exposure to information about the novel group.

Filler task (Picture-finding game). During the first study session, children were asked to find 9 objects in a book of elaborate, colorful photographs (Wick, 2011) prior to, during, and

following the video call (just as, in real life, children might be occupied with an activity when they overhear people comment about social groups). If children were unable to find a target object within 20 seconds, E said, "Let's move on to the next one." Six of the 9 target objects were the same for younger children (4–5 years) and older children (6–9 years), but younger children had an additional 3 'easy' objects to find and older children had an additional 3 'challenging' objects to find. In rare cases when a child found all of the objects before the end of the allotted time, E asked the child to find items from the other age-group's list. After starting and playing the game for 90 seconds, children were asked to find, "all of the people in the picture", and to count them. To prevent children from completing this task too quickly, they were told that there were, "a bunch of people, and some are hidden." Children who finished early were prompted to keep looking for more people in the picture.

Video call. As E introduced the child to the picture-finding game, she surreptitiously started a pre-recorded video on the laptop and moved the laptop screen so that it was facing away from the child (in order to limit distraction). The first 128 seconds of the video was a screen capture of an Apple computer desktop with a solid, blue background; thus, it *appeared* as if no video was playing on the screen. Approximately 20 seconds after the child began locating 'all of the people' in the picture (128 seconds into the video), a Skype video call embedded in the pre-recorded video began to ring on the laptop screen. E pretended to answer the call (saying to the child, "Hold on, someone's calling me on Skype. It's my big/little sister's friend, I'll answer it quickly") and turned the computer screen so it was facing themself and viewable to the child. In *adult-caller* conditions (n = 57), E referred to the caller as their "big sister's friend", for *child-caller* conditions (n = 64), E referred to the caller as their "little sister's friend." Callers were two White, female child actors (both 7-year-olds) and two White, female adult actors (M_{age} = 27

years). All actors spoke with U.S. English accents, had medium-length dark hair (pulled back away from their face), wore a black t-shirt, and did not wear glasses or other accessories. During the call, E maintained a neutral face and only looked toward the laptop screen, so as to not distract the participant or provide the participant with additional social cues.

The caller began the conversation by greeting E and saying, "Hi! I thought your sister would be using the computer, what are you doing?" E told the caller that they were playing games with the child and that they would later talk to the child about Gearoo or Flurp people (children were randomly assigned to hear about one of the two novel groups). In Negative Message conditions (n = 59) the caller responded to E, "those Flurps/Gearoos are really bad people. They eat disgusting food, and they wear such weird clothes. The Flurps'/Gearoos' language sounds so ugly." The caller ended the conversation by saying, "I'll let you get back to work! Tell your sister I called!" In No Message conditions (n = 62) participants heard the caller's greeting, the experimenter's reply, and the caller's ending of the conversation, but the caller did not say anything about the novel group. At the end of the conversation, E closed the laptop and asked the child to report how many people they found in the picture. The picture-finding game was played a second time, for 90 seconds, using a different page of the book. Each participant was randomly assigned to a Message condition (Negative Message vs. No Message) and a Caller condition (Adult vs. Child). Group Name (Gearoos vs. Flurps) was counterbalanced across participants.

Video chat stimuli. Video stimuli were constructed prior to the study. Each of the four actors called a researcher using Skype, and during their call, made negative statements using *both* novel group names within the same take (e.g., "The Flurps'…Gearoos'… language sounds so ugly"). The researcher screen-captured the entire conversation from the perspective of their

laptop screen. Using each of these original four recordings, eight new Negative Message videos (four about Gearoo people and four about Flurp people) were created. “Gearoo” and “Flurp” versions were created by editing-out one group name from the original recording. These procedures ensured that messages about the two novel groups were identical (in length, actors’ intonation, and facial expressions). To create videos for the No Message conditions, all four original recordings were trimmed to remove the callers’ claims about the group (approximately 10 seconds removed). Resulting Negative Message and No Message videos were identical, aside from differences in length and the presence of the caller’s message.

Measures

Social distance task. Immediately following the picture-finding game, children completed a social distance task based on work by Diesendruck and Menahem (2015). Children were presented 13 colored pencils inside a transparent box and asked to name their favorite and least favorite colors (as a warm-up to this task). E placed three stapled sheets of 21.7 x 35.7 cm black construction paper on top of a 21.7 X 35.7 sheet of white paper on the table. A rectangle was cut-out from the bottom, center of the first black sheet, exposing a 5.3 x 7.7 cm section of the white paper. Children were asked to use the colored pencils to draw themselves in this rectangle. The rectangle controlled the location in which children could draw themselves on the paper and the amount of space that could be occupied by that drawing. To limit the influence of E’s presence, she pretended to check her email (on the laptop) while the child was drawing.

When the participant finished their drawing, E asked, “Have you ever heard of Gearoo/Flurp people?” Children who answered “Yes” were asked, “Where did you hear about Gearoo/Flurp people?” Immediately following the caller’s message, 7% of participants said they had heard of Gearoo/Flurp people. Following the 2-week delay, with a different researcher, 56%

of children reported having heard of the group (excluding one participant who refused to answer this question and a second participant whose response was not collected due to experimenter error). Presumably, few children reported hearing about the group *immediately* after the call either because they interpreted this question as referring to their knowledge about the group *before* the study session, or because children felt uncomfortable admitting they had listened-in on the researcher's call. Next, E asked participants to, "Tell me about Gearoos/Flurps." We had no *a priori* hypotheses about how children might respond to this question but anticipated that responses might be informative. Two research assistants (blind to the study hypotheses) coded 20% of these data and achieved an inter-rater reliability of 95%; remaining data were coded by one of these research assistants. Analysis of open-ended responses revealed no robust patterns that speak to the current research questions, so they are not discussed further in this manuscript. The coding system and findings can be found in Supplementary Materials.

Following these questions, E placed the white paper (with the child's drawing of themself) beneath the second sheet of black construction paper, of which the entire bottom was removed, revealing 8 x 35.7 cm of the white paper. Thus, the child had equivalent blank space on each side of the drawing of themself (at center). E asked the child to draw a member of the novel group on that same white paper. If participants were confused about what to draw, they were instructed to, "Draw whatever you think a Gearoo/Flurp person looks like." The horizontal distance between children's drawings of themselves and the novel group member (i.e., the distance between the closest points on the two drawings) were measured in centimeters. When the child and the novel group member overlapped (n = 13) distance was recorded as 0 cm.

Direct questions. Next, participants were asked questions about the group or about members of the group. Prior to asking these questions, children were told that they were going to

talk more about Gearoo/Flurp people. E told the child, "there's no right or wrong answer to any of these questions. I just want to know how you really feel. So, you can tell me whatever you want. Okay?" For each question, if a child did not initially respond or if they could not decide between options (e.g., Yes or No), E gently prompted the child (e.g., "If you had to say something, what would you say?") and repeated the question. For exploratory purposes, after children's response to each question, E asked "Why?" We had no hypotheses about how children might respond to these questions but anticipated that responses might be informative. The same two (blind) research assistants, described earlier, coded 20% of these responses and achieved an average inter-rater reliability of 97% across all questions and categories, and the same individual coded the remaining data. As with the other open-ended data, analyses revealed no clear or informative patterns, so they are not discussed further in this manuscript. The coding system and findings can be found in Supplementary Materials (because we were interested in children's reasoning for negatively valanced responses, coding is included only for children's explanations for why they responded "no" to a given question).

***Affiliation decisions*.** E asked the child if they would want to be friends with a Gearoo/Flurp person. Participants could answer Yes (scored 1) or No (scored 0).

Goodness evaluations. Children were asked if they thought Gearoos/Flurps were good people (Yes or No) and if Gearoos/Flurps were very (not) good or just a little (not) good. Goodness evaluations were scored such that 'Very not good' = 0, 'A little not good' = .33, 'A little good' = .67, and 'Very good' = 1.00. For the interview session immediately following the claims, one participant responded that the group was 'good' but when asked if they were 'a little good' or 'very good', the participant could not decide, so that participant received a score of .67

(equivalent to ‘A little good’). For the delayed interview, one participant responded “I don’t know” to the question of whether or not the group was good, so they received a score of .5.

Cultural engagement. Children were asked whether they would engage with each of five elements of the novel group’s culture: (1) play a Gearoo/Flurp game, (2) attend a Gearoo/Flurp birthday party, (3) learn the Gearoo/Flurp alphabet, (4) try Gearoo/Flurp food, and (5) wear Gearoo/Flurp clothes. Children earned 1 point for each element that they agreed to try. Five participants provided an indecisive response (e.g., “I don’t know”) for one or two elements during the Immediate session, and two participants provided an indecisive response for one element during the Delayed session. Each indecisive response was scored .5. Each child’s *cultural engagement* score was the sum of the scores for all five elements (ranging from 0 to 5).

Resource allocation task. The final task was a version of Blake and Rand’s (2010) ‘Dictator Game’. Children selected their favorite set of stickers from a transparent box with 10 compartments; each compartment contained a different set of identical stickers. E removed the child’s chosen set and, while counting them aloud, placed the stickers in front of the participant. E covered the stickers with a blank folder and confirmed that the child understood there were 10 in total. E placed two white (10.4 cm x 24.0 cm) envelopes below the stickers on the table and told the child that they could keep all 10 stickers for themselves or they could give some to a Gearoo/Flurp girl/boy (matched to the participant’s gender). E wrote the child’s name on one envelope (to the child’s left) and instructed them to place stickers they wanted for themselves inside. Children were told to place stickers they wanted to *give to* the Gearoo/Flurp child in the blank envelope (to the child’s right). E confirmed that the child understood these instructions and assured them that they could decide for themself which envelopes to place the stickers inside. Next, E set up an occluder (composed of three manila envelopes stapled together; 23 cm x 29.7

cm) so that the child and stickers were on one side and E was on the other side. E turned around in her chair to face away from the child. The first 11 children in the study (evenly distributed across the age range and conditions) completed the task without an occluder. After the child made their decision, E gave the child their envelope to take home and placed the other envelope aside. After the session, E recorded the number of stickers inside this envelope.

Debriefing

Children who participated at school were debriefed, along with their classmates, when all participants in the school completed the study. Children who participated in the lab were debriefed individually after the second study session. A researcher told children that Gearoos and Flurps "were not real groups of people, but if they were real, they would probably be very nice people". The researcher answered any questions children had about the groups or the study.

Results

Statistical Approach and Model Reduction Plan

Our planned analyses consisted of a series of repeated-measures, multilevel regression models to test each continuous dependent variable (goodness evaluations, cultural engagement, drawing distance, and resource allocation) for between-subjects effects of Message (Negative Message vs. No Message), Caller (Adult vs. Child), and Child Age (as a continuous variable), for within-subjects effects of Time (Immediate vs. Two-Week Delay, nested within participant), and all combinations of interactions among these variables. For the dichotomous dependent variable (affiliation), a series of multilevel logistic regressions were performed. Repeated-measures models are special types of multilevel models (see Maindonald & Braun, 2013); henceforth, we simply describe these models as multilevel. All Analyses were conducted using Stata 15 (Stata-

Corp, College Station, TX). Supplementary Materials contain model output (in Tables S1-S7), and STATA syntax for each model.

The following model reduction plan was applied, in the following order, to identify the best fitting and most parsimonious final model for each dependent variable.

1. Full model: Multilevel model, including Message, Caller, Child Age (as a continuous variable), Time (nested within participant), and all interaction effects.

2A. If the full model does *not* reveal a significant interaction (or interactions) with Time, this suggests that the model fit does not differ for the immediate and delayed sessions. Remove non-significant 4-way interaction term and run model.

2Ai. Remove non-significant 3-way interaction terms and run model.

2Aii. (a) If there are no significant 3-way interaction terms in the previous step, remove non-significant 2-way interaction terms, and run model. (b) If there are significant 3-way interaction terms in the previous step, retain all 2-way interaction terms.

2B. If the full model reveals a *significant* interaction (or interactions) with Time, this suggests that the model fit *does* differ for the immediate and delayed sessions. Thus, run ordinary least squares (OLS) regression models, separately for the immediate session and for the delayed session. There are multiple post-hoc options for exploring significant interactions with Time. We chose to run OLS regressions separately for the Immediate session and the Delayed session because it allowed us to more easily interpret and visualize trends that differed between the two time points.

2Bi. For each OLS model, if 3-way interaction term is non-significant, remove term and run model.

2Bii. (a) For each OLS model, if 3-way interaction term is non-significant in previous step, remove non-significant 2-way interaction terms and run model. (b) If 3-way interaction term is significant in the previous step, retain all 2-way interaction terms.

To help interpret a significant interaction involving Child Age (a continuous variable) in any of the final models, exploratory pairwise comparisons will be conducted (using the postestimation -test- command in Stata 15) for fitted values at ages 5, 6, 7 and 8 years.

Preliminary analyses revealed that scores for all dependent measures were equivalent whether the group was called “Gearoos” or “Flurps”. As well, results did not differ between the two child or two adult actresses who were featured in the video calls. Thus, the name of the group and specific caller are not considered in further analyses.

Goodness Evaluations

Children’s evaluations of the novel group’s *goodness* were scored such that 0 = ‘Very not good’, .33 = ‘A little not good’, .67 = ‘A little good’, or 1 = ‘Very good’. Children’s average evaluation of the group’s goodness was .65 (95% CI [.59, .71]) immediately after the call and .72 (95% CI [.67, .77]) two weeks later.

A series of multilevel regression analyses were conducted on children’s evaluations of the novel group’s ‘goodness’ using the model-reduction plan outlined earlier (see Supplementary Materials, Table S1). Initial regression models revealed no significant 4-way interaction and no significant interaction involving Time (Model 1), no significant 3-way interaction effects (Model 2), and only one significant 2-way interaction of Message X Child Age (Model 3). The final model (Model 4) retained this significant 2-way interaction and all main effects. This model revealed significant effects of Time (β = .07, *SE* = .03, *z* = 2.47, *p* = .013, 95% CI [.01, .12]) and

Caller (β = .14, *SE* = .04, *z* = 3.50, *p* < .001, 95% CI [.06, .22]) — goodness evaluations generally increased between the immediate session and the two-week delay and participants rated the group as less good when the caller was an adult vs. a child (adult caller: *M* = .57, 95% CI [.49, .65]; child caller: *M* = .72, 95% CI [.64, .79]). More critically, this analysis continued to identify a significant interaction of Message x Child Age (β = -.08, *SE* = .03, *z* = -2.47, *p* = .014, 95% CI [-.15, -.02]), indicating that age-related trends (i.e., age slopes) differed significantly depending on whether or not participants heard the negative message. As depicted in Figure 1, there was an age-related *increase* in evaluations of the groups' goodness when children did *not* hear a message, but there was an age-related *decrease* in evaluations of the groups' goodness when children *did* hear the negative message.

To further interpret this interaction of Message x Child Age, post-hoc comparisons were conducted at ages 5, 6, 7 and 8 years, comparing fitted values for participants who did vs. did not overhear the negative message, using Stata's -test- command. At 5 years, children's evaluation of the group's goodness was similar whether or not they had overheard the negative message, $\chi^2(1)=.03$, *p* = .855. In contrast, from 6 years onward, children evaluated the group as less 'good' if they had heard the negative message (vs. not), (6 years: $\chi^2(1) = 5.02$, *p* = .025; 7 years: $\chi^2(1)=14.71$, *p* < .001; 8 years: $\chi^2(1)=14.19$, *p* < .001). No other significant main effects or significant interactions were found (see Supplementary Materials, Table S1).

Cultural Engagement

Participants were asked about their willingness to experience two elements of the novel group's culture that had been mentioned by the caller in Negative Message conditions (food and clothing) and three elements *not* mentioned in the caller's message (alphabet, game, and party). Participants earned a *cultural engagement* score ranging from 0 (not willing to engage with any

elements) to 5 (willing to engage with all elements). Immediately after the video call, children wanted to engage with 3.21 elements on average (95% CI [2.92, 3.51]), and two weeks later wanted to engage with 3.14 elements on average (95% CI [2.83, 3.45]).

A series of multilevel regression analyses were conducted on children's cultural engagement scores, using the model-reduction plan outlined earlier (see Supplementary Materials, Table S2). Initial regression models revealed no significant 4-way interaction and no significant interaction with Time (Model 1), no significant 3-way interaction (Model 2), and just one significant 2-way interaction of Message X Child Age (Model 3). The final model (Model 4) retained this significant 2-way interaction and all main effects. This model revealed a significant effect of Caller (β = .63, *SE* = .26, *z* = 2.39, *p* = .017, 95% CI [.11, 1.15]) — children were less willing to engage with elements of the novel group's culture when the caller was an adult vs. a child (adult caller: *M* = 2.92 elements, 95% CI [2.47, 3.37]; child caller: *M* = 3.48 elements, 95% CI [3.09, 3.86]). More critical to the goals of this study, the analysis continued to identify a significant interaction of Message x Child Age (β = -.47, *SE* = .22, *z* = -2.12, *p* = .034, 95% CI [-.90, -.04]), indicating that age-related trends (i.e., age slopes) differed significantly depending on whether or not participants heard the message. As depicted in Figure 2, there was an age-related *increase* in children's willingness to try elements of the group's culture when children did *not* overhear a message, but there was an age-related *decrease* in willingness to try elements of the group's culture when children *did* overhear the message.

To further interpret the interaction of Message x Child Age, post-hoc comparisons were conducted at ages 5, 6, 7 and 8 years, comparing fitted values for participants who did vs. did not overhear the message. At 5- and 6-years of age, there was no difference in children's willingness to engage with elements of the group's culture, whether or not they had overheard the negative

message (5 years: $\chi^2(1)=.88$, $p = .348$; 6 years: $\chi^2(1)=.12$, $p = .725$). At 7-years, children were (marginally) less willing to engage with elements of the culture if they had overheard the message (vs. not) ($\chi^2(1)=3.49$, $p = .062$), and by 8-years this difference was statistically significant, $\chi^2(1)=5.22$, $p = .022$. No other significant main effects or significant interactions were found (see Supplementary materials, Table S2).

Affiliation Decisions

On average, 52% of children decided to be friends with the novel group member immediately after the video call, and 64% decided to be friends two weeks later. A series of multilevel logistic regressions were conducted on children's affiliation decisions, using the model-reduction plan outlined earlier (see Supplementary Materials, Table S3). Initial regression models revealed no significant 4-way interaction and no interaction involving Time (Model 1), no significant 3-way interaction (Model 2), and only one significant 2-way interaction of Message X Child's Age (Model 3). This interaction effect and all main effects were retained in the final model (Model 4), which revealed a significant main effect of Time ($\beta = .82$, $SE = .36$, $z = 2.27$, $p = .023$, 95% CI [.11, 1.53]) —children's willingness to be friends with someone from the new group increased between the immediate session and the 2-week delay. More critically, this analysis revealed a significant effect of Message ($\beta = 5.84$, $SE = 2.97$, $z = 1.97$, $p = .049$, 95% CI [.02, 11.66]), and a significant interaction of Message x Child Age ($\beta = -.47$, $SE = .22$, $z = -2.12$, $p = .034$, 95% CI [-.90, -.04]), which indicated that age-related trends (i.e., age slopes) differed significantly between conditions. These effects are illustrated in Figure 3 (created using Stata's -margins- and -marginsplot- functions). For children who had *not* overheard the message, there was a significant age-related increase in children's decisions to be friends with a group

member. In contrast, for children who *did* overhear the message there was an age-related *decrease* in their decisions to be friends with a group member.

To further interpret the interaction of Message x Child Age, post-hoc comparisons were conducted at ages 5, 6, 7 and 8 years, comparing fitted probabilities for participants who did vs. did not overhear the message (using Stata's -test- command). At 5 and 6 years, children were equally willing to be friends with a member of the novel group regardless of whether they had overheard the negative message (5 years: $\chi^2(1)=.70$, $p = .40$; 6 years: $\chi^2(1)=.53$, $p = .47$). In contrast, from 7 years of age onward, children were *less willing* to be friends with a group member if they had overheard the message (vs. not), (7 years: $\chi^2(1)=5.89$, $p = .015$; 8 years: $\chi^2(1)=8.51$, $p = .004$). No other significant main effects or significant interactions were found (see Supplementary materials, Table S3).

Social Distance

An indirect measure of each child's sentiment toward the novel group was the distance (in cm) that they included between their drawings of themself and the novel group member. We excluded one participant's drawing from the immediate session because they did not comply with drawing instructions. One additional participant was excluded from analysis due to experimenter error during the delayed session. On average, children drew themselves and the novel group member 3.63 cm apart (95% CI [3.08 cm, 4.17 cm]) immediately after the call, and 2.94 cm apart (95% CI [2.46 cm, 3.43 cm]) two weeks later.

An initial multilevel logistic regression (Supplementary Materials, Table S4) predicting distance between drawings identified multiple interaction effects involving Time, including a significant interaction of Time x Message (β = -9.89, SE = 4.63, z = -2.14, p = .032, 95% CI [-18.96, -.83]), Time x Child Age (β = -1.13, SE = .54, z = -2.12, p = .034, 95% CI [-2.18, -

.09]), and Time x Message x Child Age (β = 1.65, *SE* = .73, *z* = 2.27, *p* = .023, 95% CI [.22, 3.07]). These interaction effects suggest that associations between the predictor variables and drawing distance differ significantly between the Immediate and Delayed testing sessions. Thus, in accordance with our model reduction plan, we conducted separate ordinary least squares (OLS) regressions for the Immediate session and for the Delayed session.

For the Immediate session, a series of OLS regressions predicted drawing distance based on whether or not children overheard the Message, Caller, and Child Age (see Supplementary Materials, Table S5, top). These regressions revealed no significant 3-way interaction (Model 1), and no significant 2-way interaction (Model 2). The final model (Model 3) included only main effects, revealing a significant age-related *increase* in children's drawing distance (β = .59, *SE* = .23, *t* = 2.60, *p* = .010, 95% CI [.14, 1.04]), illustrated on the left side of Figure 4. The lack of an interaction between Message and Child Age indicates that age-related trends were similar for the two conditions. No other significant main effects or interaction effects were found.

For the Delayed session, a similar series of OLS regressions were conducted (see Supplementary Materials, Table S5, bottom). These regression analyses revealed no significant 3-way interaction (Model 1) and only one significant 2-way interaction of Message X Child Age (Model 2). This interaction effect and all main effects were retained in the final model (Model 3), in which the interaction effect remained significant (β = .81, *SE* = .41, *t* = 2.00, *p* = .048, 95% CI [.01, 1.62]). As depicted on the right side of Figure 4, there was an age-related *decrease* in distance between drawings among children who had *not* overheard the message, but an age-related *increase* in distance between drawings among children who *had* overheard the negative message. To further interpret the interaction of Message x Child Age, post-hoc comparisons were conducted at ages 5, 6, 7 and 8 years, comparing fitted values for participants who did vs.

did not overhear the message. At 5 and 6 years, there was no difference in the distance between children's drawings whether or not they had overheard the message (5 years: $F(1, 114) = .90$, $p = .345$; 6 years: $F(1, 114) = .06$, $p = .799$). By 7-years, children drew the novel group member (marginally) further away from themselves when they had (vs. had not) overheard the message ($F(1, 114) = 2.87$, $p = .093$), and by 8-years this difference was statistically significant, $F(1, 114) = 4.42$, $p = .038$. No other significant main effects or interactions were found.

Resource Allocation

Recall that children could donate up to 10 stickers to a member of the novel group. On average, children donated 4.08 stickers (95% CI [3.61, 4.56]) immediately after the call, and 3.47 stickers (95% CI [3.03, 3.90]) two weeks later. An initial multilevel logistic regression (Supplementary Materials, Table S6) predicting donation quantity identified a significant interaction of Time x Message x Caller ($\beta = 9.38$, $SE = 4.58$, $z = 2.05$, $p = .044$, 95% CI [.41, 18.35]). These interaction effects suggest that associations between the predictor variables and donation amount differ significantly between the Immediate and Delayed testing sessions. Thus, according to our model reduction plan, we conducted separate ordinary least squares (OLS) regressions for the Immediate session and for the Delayed session.

For the Immediate session, a series of OLS regressions predicted donation amount based on whether or not children overheard the Message, Caller, and Child Age (see Supplementary Materials, Table S7, top). These analyses revealed no significant interaction effects or main effects (Models 1-3) predicting children's donations. For the Delayed session, a similar series of OLS regressions were conducted (see Supplementary Materials, Table S7, bottom), again revealing no significant associations between the predictor variables and donations (Models 1-3).

Thus, at neither testing session did any variable significantly predict children's resource allocation.

Discussion

One way to inform work on the development of intergroup bias is to examine how verbal messages might influence children's intergroup attitude formation. Children have many opportunities to hear both positive and negative information about groups of people in their daily lives. Indeed, in the current media-saturated world, children may often encounter these messages in passing; for example, as they overhear the musings of radio hosts, television news anchors, or others' video chat conversations. In this study, we investigated whether overhearing brief (approximately 10-second) messages about a novel social group, from an absent caller on video chat, influenced 4- to 9-year-olds' attitudes toward that group. We also explored whether such effects varied by children's age, differed by the age of the caller, and whether they persisted following a two-week delay.

Conceivably, we might have found that exposure to this sort of information would *not* influence children's intergroup attitudes — the messages were brief, were presented only once, the speaker was a stranger who was not in the room with the child, children were occupied with another activity (finding pictures in a book) and were never directed to attend to the video call. Yet, across multiple measures of children's intergroup attitudes, there were robust effects of hearing these messages. Children who heard the caller's message demonstrated stronger, negative attitudes toward the group than children who heard no message. Consistent with our predictions and prior research on children's developing intergroup attitudes, the influence of such messages increased with children's age; effects were stronger among the oldest participants (Gonzalez, Steele, & Baron, 2017; Jordan & Hernandez-Reif, 2009; Lane et al., 2020), and these

effects persisted following a two-week delay. Contrary to our predictions, children's attitudes were equally influenced by negative messages whether the caller was an adult or a child.

Immediately following the video call, and again two weeks later, children 6-years and older evaluated the group as being less 'good' if they had overheard the negative message (relative to children who had not overheard the message). Immediately following the call, and two weeks later, children 7-years and older were additionally less willing to affiliate with a member of the group, and were less willing to engage with elements of the group's culture if they had overheard the message. As well, two weeks after overhearing the negative message, children 7-years and older drew themselves further away from a member of the novel group (relative to children who had not overheard the message), suggesting that children felt more socially distant from (or wanted to be further from) members of that group (Bar-Tal & Teichman, 2009; Diesendruck & Menahem, 2015). Although we had no hypotheses about the specific ages at which the effects of these messages would emerge for each dependent measure, it is interesting that children's general evaluation of the group (which required minimal extrapolation from the caller's claim that the group was "bad") was influenced earlier in development than children's anticipated willingness to affiliate with the group and their sentiment toward the group (which arguably required more extrapolation from the caller's claims). Lane et al. (2020) also found that the effects of overhearing derogatory claims about novel social groups manifested earlier in children's general evaluations of the group. Perhaps when younger children hear disparaging claims about a social group, they initially treat that information as a non-personal factoid, and that information does not yet influence their *personal* feelings toward the group or desire to interact with the group. Slightly later in development, those disparaging remarks have farer-reaching implications in how children characterize, feel

toward, and act toward the other social group. In sum, by middle childhood, these messages had broad and robust effects on children's attitudes toward the maligned group.

However, across all measures, there were no effects of these messages among the youngest participants (4- to 5-year-olds), consistent with prior work in which overhearing denigrating claims about social groups (from an adult standing near the child) had little influence on preschoolers' intergroup attitudes (e.g., Lane et al., 2020). There are several viable interpretations of this finding. Although previous research has identified that toddlers and preschoolers are capable of learning about new words and object functions by overhearing others' conversations (Akhtar, 2001, 2005; Boderé & Jaspaert, 2017; Floor & Akhtar, 2003; Gampe, Liebal, & Tomasello, 2012; O-Doherty et al., 2011), these studies differ from ours in important ways. For example, in past work, overheard information was usually *repeated* several times and what was being labeled or described was typically *visible* to the child. In the current study, perhaps the youngest children had difficulty processing information from the overheard message, which was communicated just once, during a video chat conversation, by an absent speaker, and which referred to people who were never depicted. Of note, we purposely never depicted the group because we did not want to evoke or reinforce any existing intergroup biases (e.g., based on race or attire). It is also possible that this social information was simply less intriguing and relevant to the youngest children. Future studies might examine if overhearing these sorts of messages influences preschoolers' intergroup attitudes when messages are repeated or in circumstances where the information is more personally relevant (e.g., if children expect that they will actually meet someone from the group).

We found no differences in children's donations to the group, whether or not children heard the caller's negative message. This was inconsistent with our predictions and with several

studies demonstrating that children in this age range donate fewer resources to out-group members (e.g., Buttelmann & Böhm, 2014; de França & Monteiro, 2013; Jordan, McAuliffe, & Warneken, 2014; Renno & Shutts, 2014; Rizzo & Killen, 2018). However, several noteworthy differences exist between the resource allocation task used in this study and those used in past work. We believe that the most important of these differences is that, in the current study, there was a *benefit* in *not* donating resources (children could keep the stickers that they did not give away); whereas in past studies there was no personal benefit — children had to forfeit resources they did not donate (e.g., Buttelmann & Böhm, 2014; de França & Monteiro, 2013; Renno & Shutts, 2014; Rhodes et al., 2017; Rizzo & Killen, 2018). Thus, in the current study children might have been inclined to keep stickers for themselves regardless of their sentiment toward the novel group. If children did not have the opportunity to keep the stickers, perhaps we would have found that they donated fewer stickers after hearing negative claims about the novel group. Future research can test this possibility.

Finally, contrary to our hypotheses, the age of the caller (whether the caller was an adult or child) did not significantly moderate effects of the caller's message on older children's attitudes. Indeed, no interactions of Message x Caller or Message x Caller x Child Age were statistically significant. It is possible that the children in this sample were too young for us to detect effects of caller age (the oldest participants were 9 years). In a study with older participants, 10- to 18-year-olds who initially reported that it was wrong to exclude another child were more willing to agree with a peer's claim (vs. a parent's claim) that social exclusion was acceptable (Killen et al., 2002). Another factor that might influence the force of these messages is the age of the child caller. Although child actors in this study were 7 years old (close to the average age of our participants) an older child caller might be considered a more authoritative

source of social information, as children might value older children's intergroup claims more (French, 1984). Future work should evaluate the effects of these messages on the intergroup attitudes of older children and adolescents, and should examine whether these messages are more forceful when they are provided by older children. There were two significant *main effects* of caller— if the caller had been an adult (vs. child), children evaluated the novel group as less 'good' and were less willing to engage with the group's culture. However, we had no hypotheses about main effects of Caller (only hypotheses for interactions of Message x Caller) and there is no obvious post-hoc explanation for why there would be significant main effects of Caller, so we hesitate to interpret those two effects.

The current findings inspire additional avenues for research. In this study, messages about the novel group were purposely brief but detailed; they included multiple pieces of information about the group's culture (food, language, and clothing) that children and adults commonly use to infer group membership (Dunham, Baron, & Banaji, 2008; Liberman, Woodward, & Kinzler, 2017; Mahajan & Wynn, 2012). The group was also described as generally "bad." Future studies might investigate how *specific aspects* of these messages influence children's attitudes and behavior toward social groups. For example, would merely describing the group as "bad" be enough to induce the biases found in the current study? The informants in the current study were all unfamiliar to our participants; future work is needed to explore whether children's familiarity with speakers moderates the impact of their claims. As well, would children's attitudes be affected to the same extent if the video caller were speaking *directly* to them? In a study on the effects of derogatory messages uttered by someone in the *same room* as children, messages provided directly to children (vs. overheard messages) more forcefully influenced children's intergroup attitudes (Lane et al., 2020).

Whether direct messages about social groups provided via other means (e.g., phone calls or video calls) similarly influence children's intergroup attitudes remains an open question. The current sample was heterogenous, and the distribution of participants across racial groups was roughly similar to distributions at the national level (US Census Bureau, 2019). An intriguing line of research would involve evaluating how defamatory messages about novel social groups influence the attitudes of children who are themselves members of marginalized groups, compared to children from non-marginalized groups. To adequately address this question, future studies will need to over-sample participants from marginalized populations.

In sum, the current work provides evidence that, during the early elementary-school years, children's social attitudes can be profoundly influenced by denigrating messages that they may *overhear* about other groups. Children 7-years and older who overheard one brief, negative message from a person who was *not physically present* demonstrated stronger, negative attitudes toward a novel social group than children who did not hear the message. The effects of these messages were robust — most effects were maintained longitudinally (following a two-week delay) and existed whether the message was uttered by an adult or a child. These results highlight that children's *indirect* exposure to others' media use can influence their developing concepts of the social world, a topic that warrants increased attention. Further work on the effects of overheard claims about social groups can help to identify ways in which exposure to other types of message (e.g., positive claims) may reduce or counter the development of intergroup biases.

References

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, *72*, 416–430. doi: 10.1111/1467-8624.00287

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, *8*, 199–209. doi: 10.1111/j.1467-7687.2005.00406.x

Baron, A. S., Dunham, Y., Banaji, M., & Carey, S. (2014). Constraints on the acquisition of social category concepts. *Journal of Cognition and Development*, *15*, 238–268. doi: 10.1080/15248372.2012.742902

Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development*, *68*, 530–543. doi: 10.1111/j.1467-8624.1997.tb01956.x

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. doi: 10.1111/j.1467-8721.2007.00496.x

Blake, P. R., & Rand, D. G. (2010). Currency value moderates equity preference among young children. *Evolution and Human Behavior*, *31*, 210–218. doi: 10.1016/j.evolhumbehav.2009.06.012

Boderé, A., & Jaspaert, K. (2017). Six-year-olds' learning of novel words through addressed and overheard speech. *Journal of Child Language*, *44*, 1163–1191. doi: 10.1017/S0305000916000465

Cabrera, N., Kuhns, C., Malin, J. L., & Aldoney, D. (2016). Helping children navigate a diverse world: Parents' contributions. In S. S. Horn, M. D. Ruck, & L. S. Liben (Eds.), *Advances in Child Development and Behavior* (81–102). Elsevier Inc. doi:

10.1016/bs.acdb.2016.05.002

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159. doi: 10.1037/0033-2909.112.1.155

de França, D. X., & Monteiro, M. B. (2013). Social norms and the expression of prejudice: The development of aversive racism in childhood: Social norms and the expression of prejudice. *European Journal of Social Psychology, 43,* 263–271. doi: 10.1002/ejsp.1965

Degner, J., & Dalege, J. (2013). The apple does not fall far from the tree, or does it? A meta-analysis of parent–child similarity in intergroup attitudes. *Psychological Bulletin, 139,* 1270–1304. doi: 10.1037/a0031436

Diesendruck, G., & Menahem, R. (2015). Essentialism promotes children's inter-ethnic bias. *Frontiers in Psychology, 6.* doi: 10.3389/fpsyg.2015.01180

Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology, 8,* 100–110. doi: 10.1037/1089-2680.8.2.100

Dunham, Y. (2018). Mere membership. *Trends in Cognitive Sciences, 22,* 780–793. doi: 10.1016/j.tics.2018.06.004

Dunham, Y., & Emory, J. (2014). Of affect and ambiguity: The emergence of preference for arbitrary ingroups. *Journal of Social Issues, 70,* 81–98. doi: 10.1111/josi.12048

Eccles, J. S. (1999). The development of children ages 6 to 14. *The Future of Children, 9,* 30–44. doi: 10.2307/1602703

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191. doi: 10.3758/BF03193146

Federal Bureau of Investigation. (2017). *Hate Crime Statistics, 2017.* Retrieved from

ucr.fbi.gov/hate-crime/2017

Floor, P., & Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy, 9*, 327–339. doi: 10.1207/s15327078in0903_4

Foster, E. K., & Hund, A. M. (2012). The impact of scaffolding and overhearing on young children's use of the spatial terms between and middle. *Journal of Child Language, 39*, 338–364. doi: 10.1017/S0305000911000158

Foushee, R., & Xu, F. (2016). Active overhearing: Development in preschoolers' skill at 'listening in' to naturalistic overheard speech. In Papafragou, A., Grodner, D., Mirman, D., & Trueswell, J.C. (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society: Recognizing and Representing Events* (pp. 2178 – 2182). Philadelphia, PA: Cognitive Science Society, Inc.

French, D. C. (1984). Children's knowledge of the social functions of younger, older, and same-age peers. *Child Development, 55*, 1429-1433. doi: 10.2307/1130012

Gampe, A., Liebal, K., & Tomasello, M. (2012). Eighteen-month-olds learn novel words through overhearing. *First Language, 32*, 385–397. doi: 10.1177/0142723711433584

Gonzalez, A. M., Steele, J. R., & Baron, A. S. (2017). Reducing children's implicit racial bias through exposure to positive out-group exemplars. *Child Development, 88*, 123–130. doi: 10.1111/cdev.12582

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science, 17*, 757–758. doi: 10.1111/j.1467-9280.2006.01778.x

Jordan, J. J., McAuliffe, K., & Warneken, F. (2014). Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of*

Sciences, 111, 12710–12715. doi: 10.1073/pnas.1402280111

Killen, M., Lee-Kim, J., McGlothlin, H., Stangor, C., & Helwig, C. (2002). How children and adolescents evaluate gender and racial exclusion. *Monographs of the Society for Research in Child Development, 67,* i – 129.

Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition, 27,* 623–634. doi: 10.1521/soco.2009.27.4.623

Lane, J. D., Conder, E. B., & Rottman, J. (2020). The influence of direct and overheard messages on children's attitudes toward novel social groups. *Child Development, 91,* 829-845. doi: 10.1111/cdev.13238

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). Preverbal infants infer third-party social relationships based on language. *Cognitive Science, 41,* 622–634. doi: 10.1111/cogs.12403

Mahajan, N., & Wynn, K. (2012). Origins of "us" versus "them": Prelinguistic infants prefer similar others. *Cognition, 124,* 227–233. doi: 10.1016/j.cognition.2012.05.003

Maindonald, J., & Braun, W. J. (2013). *Data Analysis and Graphics Using R.* Cambridge University Press. doi: 10.1017/CBO9781139194648

Nesdale, D. (2004). Social identity processes and children's ethnic prejudice. In M. Bennett & F. Sani (Eds.), *The development of the social self,* (pp. 219-245). New York, NY: Psychology Press.

Nesdale, D., Maass, A., Griffiths, J., & Durkin, K. (2003). Effects of in-group and out-group ethnicity on children's attitudes towards members of the in-group and out-group. *The British Journal of Developmental Psychology, 21,* 177-192.

doi: 10.1348/026151003765264039

O'Doherty, K., Troseth, G. L., Shimpi, P. M., Goldenberg, E., Akhtar, N., & Saylor, M. M. (2011). Third-party social interaction and word learning from video. *Child Development, 82*, 902–915. doi: 10.1111/j.1467-8624.2011.01579.x

Over, H., & McCall, C. (2018). Becoming us and them: Social learning and intergroup bias. *Social and Personality Psychology Compass, 12*. doi: 10.1111/spc3.12384

Patterson, M. M., & Bigler, R. S. (2006). Preschool children's attention to environmental messages about groups: Social categorization and the origins of intergroup bias. *Child Development, 77*, 847–860. doi: 10.1111/j.1467-8624.2006.00906.x

Patterson, M. M., Bigler, R. S., Pahlke, E., Brown, C. S., Hayes, A. R., Ramirez, M. C., Nelson, A. (2019). Toward a developmental science of politics. *Monographs of the Society for Research in Child Development, 84*, 1 – 185. doi: 10.1111/mono.12410

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. doi: 10.1111/j.1467-8624.2011.01668.x

Renno, M. P., & Shutts, K. (2015). Children's social category-based giving and its correlates: Expectations and preferences. *Developmental Psychology, 51*, 533–543. doi: 10.1037/a0038819

Rizzo, M. T., & Killen, M. (2018). Theory of mind is related to children's resource allocations in gender stereotypic contexts. *Developmental Psychology, 54*, 510–520. doi: 10.1037/dev0000439

Roberts, S. O., Ho, A. K., & Gelman, S. A. (2019). The role of group norms in evaluating uncommon and negative behaviors. *Journal of Experimental Psychology: General, 148*,

374–387. doi: 10.1037/xge0000534

Ronfard, S., Lane, J. D., Wang, M., & Harris, P. L. (2017). The impact of counter-perceptual testimony on children's categorization after a delay. *Journal of Experimental Child Psychology, 163,* 151–158. doi: 10.1016/j.jecp.2017.06.006

Rottman, J., Young, L., & Kelemen, D. (2017). The impact of testimony on children's moralization of novel actions. *Emotion, 17,* 811–827. doi: 10.1037/emo0000276

Shanmugasundaram, S. (2018). *Hate Crimes, Explained.* Retrieved from https://www.splcenter.org/20180415/hate-crimes-explained

Shneidman, L., & Woodward, A. L. (2016). Are child-directed interactions the cradle of social learning? *Psychological Bulletin, 142,* 1–17. doi: 10.1037/bul0000023

Skinner, A. L., & Meltzoff, A. N. (2019). Childhood experiences and intergroup biases among children. *Social Issues and Policy Review, 13,* 211-240. doi: 10.1111/sipr.12054

Taylor, M., Cartwright, B. S., & Bowden, T. (1991). Perspective taking and theory of mind: Do children predict interpretive diversity as a function of differences in observers' knowledge? *Child Development, 62,* 1334–1351. doi: 10.2307/1130810

US Census Bureau (2019). https://www.census.gov/quickfacts/fact/table/US/PST045219

Wick, W. (2011). *Can You See What I See? Toyland Express: Picture Puzzles to Search and Solve.* New York, New York: Scholastic Inc.

Woolley, J. D., Ma, L., & Lopez-Mobilia, G. (2011). Development of the use of conversational cues to assess reality status. *Journal of Cognition and Development, 12,* 537–555. doi: 10.1080/15248372.2011.554929

Zhao, L., Chen, L., Sun, W., Compton, B. J., Lee, K., & Heyman, G. D. (2019). Young children are more likely to cheat after overhearing that a classmate is smart. *Developmental*

Science, e12930. doi: 10.1111/desc.12930

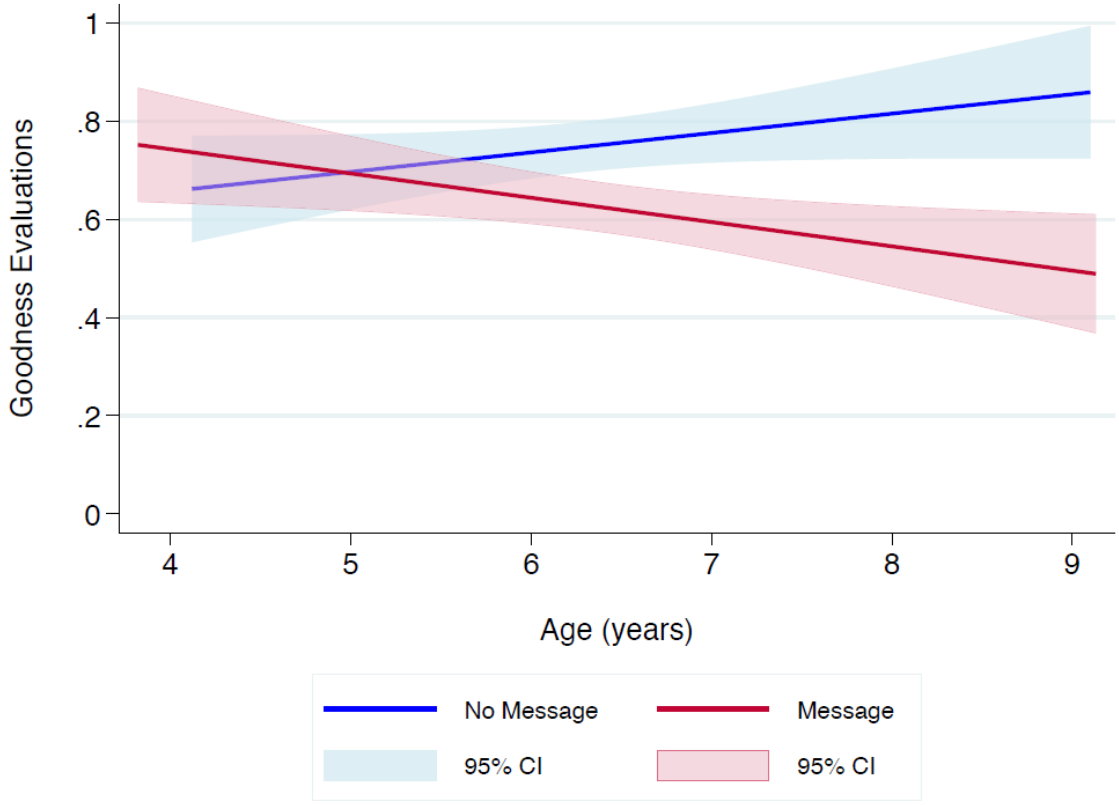

Figure 1. Evaluation of a novel group's 'goodness', by participant's age. Goodness evaluations ranged from 0 (very not good) to 1.00 (very good). Children overheard a video call in which the caller either made negative claims about a novel social group (red line) or did not (blue line). Results are averaged across choices made immediately following the video call and two weeks later. Shaded regions represent 95% confidence intervals.

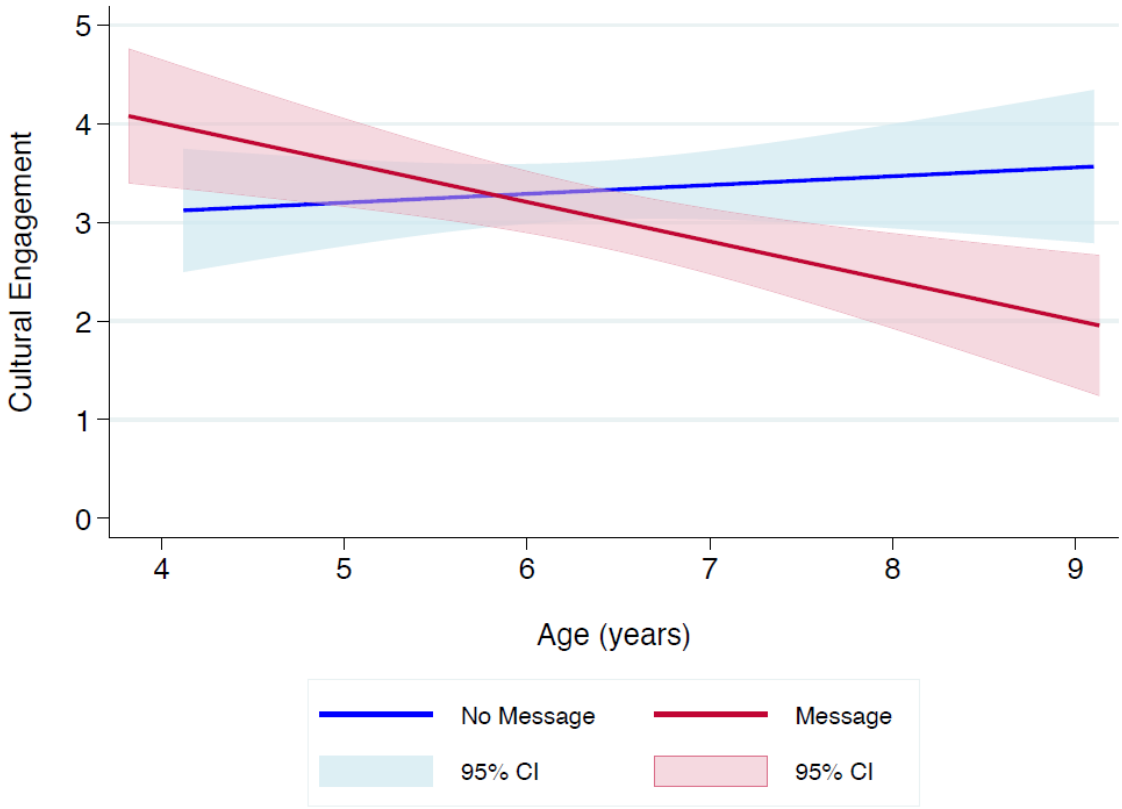

Figure 2. Average number of items from the novel group's culture that children chose to engage with (out of five), by participant's age. Children overheard a video call in which the caller either made negative claims about a novel social group (red line) or did not (blue line). Results are averaged across choices made immediately following the video call and two weeks later. Shaded regions represent 95% confidence intervals.

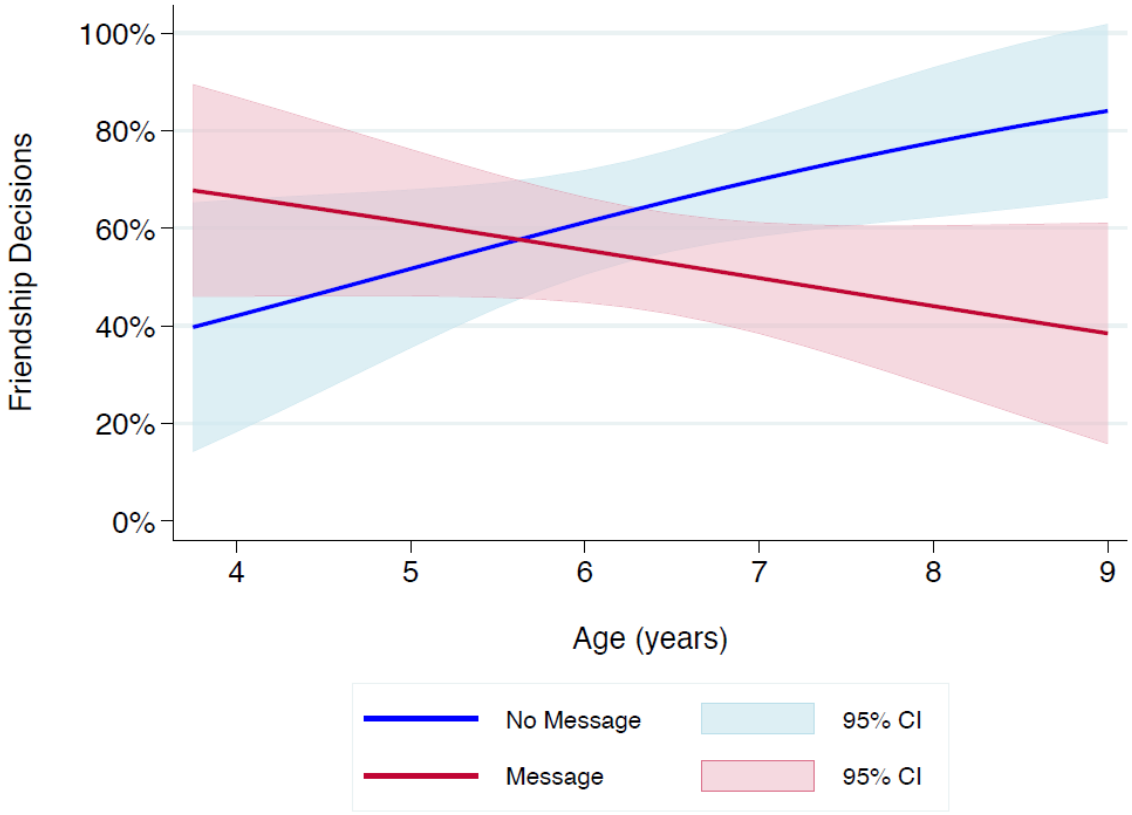

Figure 3. Probability of children choosing to be friends with the novel group member, by children's age. Children overheard a video call in which the caller either made negative claims about a novel social group (red line) or did not (blue line). Results are averaged across choices made immediately following the video call and two weeks later. Shaded regions represent 95% confidence intervals.

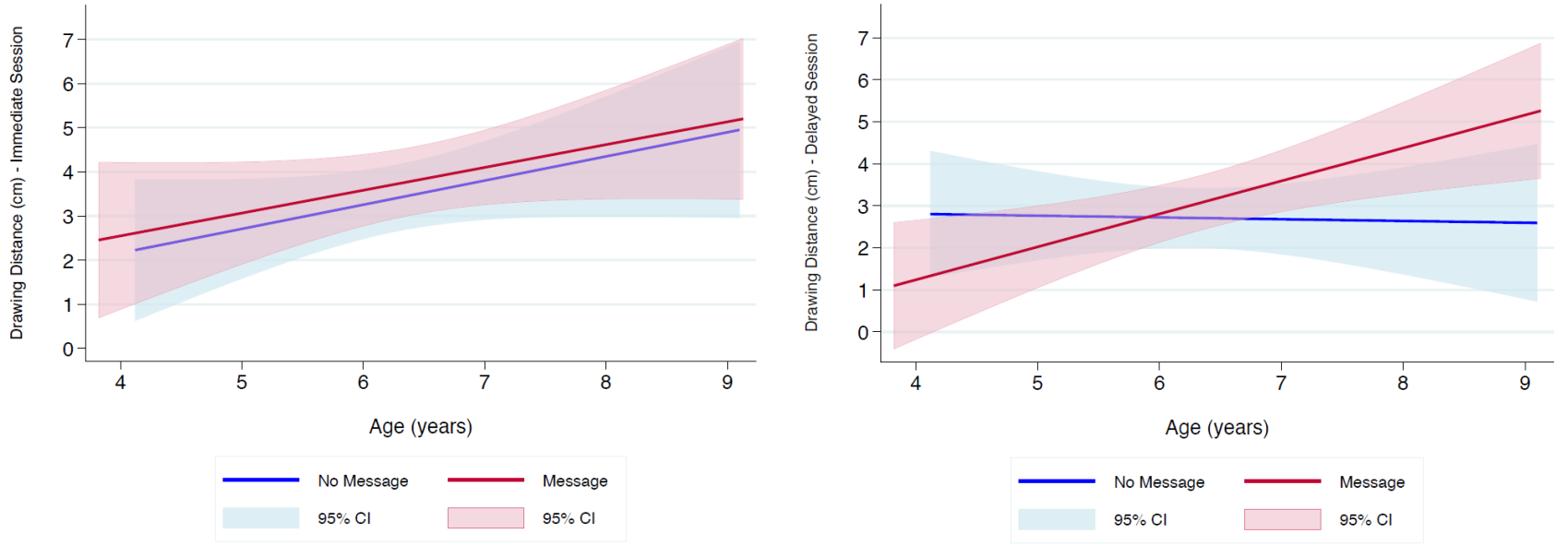

Figure 4. Average distance (cm) between children's drawings of themselves and a novel group member. Children overheard a video call in which the caller either made negative claims about a novel social group (red lines) or did not (blue lines). Children created these drawings immediately following the video call (left) and two weeks later (right). Shaded regions represent 95% confidence intervals.