

In press, *Psychological Science*

When are similar individuals a group?

Early reasoning about similarity and ingroup support

Lin Bian^{1*} and Renee Baillargeon²

¹Department of Psychology, University of Chicago

²Department of Psychology, University of Illinois at Urbana-Champaign

* Correspondence concerning this article should be addressed to Lin Bian, Department of Psychology, University of Chicago, 5848 S. University Avenue, Chicago, Illinois 60637, United States. E-mail: linbian@uchicago.edu

Abstract

Beginning in infancy, children expect individuals in a group to care for and be loyal to ingroup members. One prominent cue children use to infer that individuals belong to the same group is similarity. Does any salient similarity among individuals elicit an expectation of ingroup preference, or does contextual information modulate these expectations? In Experiments 1–2, 12-month-old infants expected ingroup preference between two individuals who wore the same novel outfit, but they dismissed this similarity if one of the outfits was used to fulfill an instrumental purpose. In Experiment 3, 26-month-old toddlers expected ingroup preference between two individuals who uttered the same novel labels, but they dismissed this similarity if the labels were used to convey incidental as opposed to categorical information about the individuals. Together, these results ($N = 96$) provide converging evidence that from early in life, children possess a context-sensitive mechanism for determining whether similarities mark groups.

Statement of Relevance

Beginning early in life, children expect individuals in a social group to prefer ingroup members over outgroup members. One prominent cue young children use to infer that individuals belong to the same group is similarity. The present research explored the nature of the relation between similarity and group membership in early moral cognition. In three violation-of-expectation experiments using minimum-group manipulations, 12- and 26-month-olds saw individuals with a salient novel similarity in their outfits or labels. When the experimental context allowed children to interpret this similarity as a signal to group membership, they expected the similar individuals to show ingroup preference. However, when the context undermined this interpretation, leading children to dismiss the similarity as a potential group marker, they no longer expected ingroup preference. From a young age, children thus refrain from interpreting a similarity as a group marker when contextual information suggests that this similarity carries little social significance.

When adults watch individuals in a scene, they typically represent who belongs to what group and use that information to predict how interactions will unfold within and between groups (Balliet et al., 2014; Roberts et al., 2017). One long-standing claim in the social sciences has been that these predictions are driven in part by a principle of *ingroup support*: Individuals in a group are expected to care for ingroup members and to show them loyalty (Brewer, 1999; Graham et al., 2013; Rai & Fiske, 2011; Shweder et al., 1997; Tajfel et al., 1971; Tooby et al., 2006). Recent findings indicate that this principle emerges early in life (Ting et al., 2020): Infants expect individuals to help ingroup members in need (Jin & Baillargeon, 2017), to refrain from helping wrongdoers who have harmed ingroup members (Ting et al., 2019), to comply with ingroup norms (Powell & Spelke, 2013), to prefer ingroup members over outgroup members (Spokes & Spelke, 2017), to reserve limited resources for ingroup members (Bian et al., 2018), and to side with ingroup members in conflicts with outgroup members (Pun et al., 2021). These results suggest that when infants represent individuals as members of the same group, the ingroup-support principle is then triggered, bringing forth rich expectations of ingroup care and loyalty.

What cues might suggest to infants that individuals belong to the same group? Much research has focused on infants' ability to use *similarity* to infer group membership. In first-party tasks, infants are more likely to support (e.g., prefer, endorse, help, and affiliate with) an individual who speaks their language (Begus et al., 2016; Buttelmann et al., 2013; Kinzler et al., 2007; Shutts et al., 2009), shares their food or toy preferences (Gerson et al., 2017; Mahajan & Wynn, 2012), or bounces to music in synchrony with them (Cirelli et al., 2014, 2016). Likewise, in third-party tasks, infants expect individuals to support others who are comforted by the same caretaker (Spokes & Spelke, 2015), have the same appearance (Bian et al., 2018; Rhodes et al., 2015), speak the same language (Lieberman et al., 2017), or share the same food preferences (Lieberman et al.,

2021). Additional findings indicate that infants can interpret as group markers not only similarities that provide meaningful information about the groups involved (e.g., about their language or preferences), but also shallow, minimal similarities (Dunham et al., 2011). Thus, infants expect individuals to support others who wear a similar novel outfit (Rhodes et al., 2015; Ting et al., 2019) or use a similar novel label to describe themselves (Jin & Baillargeon, 2017; Ting et al., 2019; for related findings with older children, see Jordan & Dunham, 2020; Sparks et al., 2017).

What mechanism might underlie infants' ability to interpret all of these different types of similarities as group markers? One possibility is that infants are equipped with a somewhat primitive and indiscriminate mechanism that responds to *any* salient similarity and simply views more similar individuals as more likely to belong to the same group. Another possibility is that sensitivity to similarity, even in infancy, depends on a sophisticated mechanism that takes into account contextual factors: Similarities are used to infer group membership when the context in which they are observed allows this inference; they are dismissed, however, when contextual information undermines this inference by suggesting that they carry little or no social significance.

One way to marshal evidence for the latter possibility is to demonstrate that infants *refrain* from interpreting a salient similarity as a group marker when contextual factors render this interpretation less likely. To date, only one manipulation has produced such a demonstration (Jin & Baillargeon, 2017). Seventeen-month-olds first heard two experimenters utter similar novel labels to convey information about themselves that was either categorical (“I am a bem!”, “I am a bem, too!”) or incidental (“I saw a bem!”, “I saw a bem, too!”). With the categorical information, infants inferred that the experimenters belonged to the same group, and they later expected them to help each other, in accordance with ingroup support. With the incidental information, however, infants drew no inference of group membership and held no expectation of ingroup support.

Similar results were obtained with 29-month-old toddlers (Ting et al., 2019). In the present research, we sought a stronger test of infants' ability to dismiss salient similarities as group markers: Experiments 1 and 2 presented 12-month-olds with a visual similarity that remained in view throughout the test events. Of interest was whether infants would refrain from interpreting this similarity as a group marker when contextual information undermined its social significance, but would interpret it as such otherwise, as in prior research.

In Experiment 1, a target experimenter (*TargetE*) wearing a novel outfit faced two experimenters, one with a similar outfit (*SimilarE*) and one with a different outfit (*DifferentE*). To manipulate whether infants construed the outfits as group markers, we built on findings that (a) children can be led by contextual information to adopt either a *ritual* or an *instrumental* stance when interpreting others' actions, and (b) they are more likely to imbue ritualistic actions with social significance, such as signaling group membership (Legare et al., 2015; Watson-Jones & Legare, 2016). For example, when an experimenter activated a light-box with her head, infants construed this action as a ritualistic, conventional action if her hands were free, but as an instrumental action if they were occupied (Gergely et al., 2002). Moreover, infants found it unexpected if two experimenters who used different conventional actions to activate the light-box later affiliated, but not if two experimenters who used different instrumental actions later affiliated (Lieberman et al., 2018). In line with these findings, we attempted to induce infants to adopt either a ritual or an instrumental stance toward the experimenters' outfits (for evidence that adults distinguish between ritualistic and instrumental outfits, see SOM).

Infants first saw TargetE put toys away either in a box (*two-group* condition) or in a pocket on her shirtfront, as though it was intended for that instrumental purpose (*no-group* condition). Infants later saw TargetE prefer either SimilarE or DifferentE. If our manipulation was successful,

responses should differ between conditions. In the two-group condition, infants should adopt a ritual stance toward the experimenters' outfits, view TargetE and SimilarE as ingroup members, and expect TargetE to prefer SimilarE over DifferentE, in accordance with ingroup support. In the no-group condition, in contrast, infants should adopt an instrumental stance toward the experimenters' outfits, dismiss them as group markers, and hold no expectation about whom TargetE would prefer.

Experiment 2 was identical to Experiment 1 except that a different manipulation was used to sway infants' interpretation of the experimenters' outfits: SimilarE either drank from a cup (two-group condition) or used a sponge attached to her outfit to clean messy lines on it (no-group condition). Finally, to provide further converging evidence for our conclusions, Experiment 3 examined whether 26-month-old toddlers would expect TargetE to prefer SimilarE when the two used similar novel labels to convey categorical (two-group condition) but not incidental (no-group condition) information about themselves.

Each of our three experiments thus tested whether children would expect TargetE to prefer SimilarE in the two-group but not the no-group condition. Finding this pattern across ages and manipulations would provide strong evidence that beginning early in life, the process of interpreting similarities among individuals is neither indiscriminate nor shallow: Even salient similarities may be dismissed if the context in which they are observed suggests that they carry little social significance.

Experiment 1

Infants in each condition received a familiarization trial, a pretest trial, and two test trials (Fig. 1). Only the familiarization trial differed across conditions; it served to manipulate whether infants adopted a ritual or an instrumental stance toward the experimenters' outfits. TargetE sat

alone at the back of a puppet-stage apparatus and, as explained above, stored toys either in a box (two-group condition) or in a pocket (no-group condition). In the pretest trial, TargetE was absent, and SimilarE and DifferentE sat at side windows and read identical picture books. In the test trials, TargetE approached either SimilarE (*approach-SimilarE* event) or DifferentE (*approach-DifferentE* event) to read with her. Evidence that infants in the two-group condition looked significantly longer at the approach-DifferentE than at the approach-SimilarE event, whereas infants in the no-group condition looked equally at the events, would indicate that although TargetE and SimilarE always wore the same outfits, infants refrained from interpreting this similarity as a group marker when shown that one of the outfits served an instrumental function.

Method

Design

Each trial had an initial phase and a final phase. During the initial phase, which was computer-controlled, infants saw the scripted events appropriate for the trial, ending with a paused scene. During the final phase, which was infant-controlled, infants watched this paused scene until the trial ended. Across trials, infants saw two outfits: a *pink* outfit (pink shirt, purple forehead band with a flower at the front) and a *blue* outfit (blue shirt, green fuzzy tiara, red scarf with multicolored dots). DifferentE wore one outfit (counterbalanced), and TargetE and SimilarE wore the other outfit; whichever outfit they wore also had a large black-and-white pocket on the shirtfront.

In the (18-s) initial phase of the *familiarization* trial in the two-group condition, TargetE sat alone with a colorful open box on her left. Scattered on the apparatus floor were six colorful thin foam toys. TargetE grasped the back of the box with her left hand and then picked up each toy, one at a time, and placed it silently in the box. She then looked down at the apparatus floor and paused until the trial ended. The trial in the no-group condition was identical except that

TargetE opened the top of her pocket with her left hand and silently placed each toy in it (because the toys were thin and the pocket extended below the apparatus floor, TargetE's appearance was not changed by the addition of the toys to her pocket).

In the (18-s) initial phase of the *pretest* trial, TargetE was absent, and SimilarE and DifferentE sat at their side windows (counterbalanced) and looked at an identical picture book. They opened their books, flipped through the first three pages, each at her own rate, and then paused until the trial ended.

In the (7-s) initial phase of each *test* trial, all three experimenters sat at their windows; TargetE had no book, and SimilarE and DifferentE each had their book open to the first page. TargetE looked back and forth at them twice (side of first look was counterbalanced), then she approached either SimilarE or DifferentE to read with her. SimilarE and DifferentE looked at TargetE as she moved, and then all three experimenters paused, looking down at the books, until the trial ended. Test order was counterbalanced within each condition.

Power Analysis

In a previous violation-of-expectation report on early morality (Sloane et al., 2012), infants were shown a fair and an unfair event using a within-subject design. Infants in the experimental condition ($N = 16$ per condition) looked significantly longer at the unfair than at the fair event, whereas those in the inanimate-control condition looked equally at the events. The data yielded a significant Condition \times Event interaction with an effect size (η_p^2) of 0.142. A G*Power analysis (Faul et al., 2007) based on this value, with alpha set at .05 and power set at .80, suggested that the minimum number of participants per condition in our experiments—which focused on ingroup support rather than fairness but used a similar design—was 11. Nevertheless, we tested 16 participants per condition, in line with this and other previous reports on early morality (Choi &

Luo, 2015; Margoni et al., 2018). We ceased data collection for a given condition when we reached our target of $N = 16$.

Participants

Participants were 32 term 12-month-olds (15 male, $M = 12;3$, range = 11;10–12;24). No additional infant was tested but excluded. Sixteen infants were randomly assigned to each condition. Participants' names were obtained from a university-maintained database of parents interested in child-development research. Each participant's parent gave written informed consent, and the protocol was approved by the [xx] Institutional Review Board.

Apparatus

The apparatus consisted of a brightly lit white display booth (204 cm high \times 100 cm wide \times 74 cm deep) with a large opening (57 \times 93 cm) in its front wall; between trials, a supervisor lowered a curtain to hide this opening. TargetE sat at a window (72 \times 96 cm) in the back wall, and DifferentE and SimilarE sat at windows (57 \times 48 cm) in the side walls; each side window had a white curtain that could be drawn aside. Behind the experimenters, floor-to-ceiling white curtains hid the testing room from view. Stimuli included a colorful open box, six colorful thin foam toys, and two identical picture books.

During a testing session, the experimenters never made eye contact with the infant: As the events unfolded, they looked at each other or at the objects they acted on. Two cameras captured images of the infant and events; the two images were combined, projected onto a monitor located behind the apparatus, and checked by the supervisor to confirm that the trials followed the prescribed scripts. Recorded sessions were also checked off-line for experimenter and observer accuracy.

Procedure

Each infant sat on a parent's lap, and parents were instructed to remain silent and to close their eyes during the test trials. Two observers hidden on either side of the apparatus monitored each infant's looking behavior; the observers were blind to the infant's condition and test order. Looking times during the initial and final phases of each trial were computed separately, using the primary observer's responses. Infants were highly attentive during the initial phases of the trials and looked, on average, for 96% of each initial phase. The final phase of each trial ended when infants (a) looked away for 2 consecutive seconds after having looked for at least 6 cumulative seconds or (b) looked for a maximum of 60 cumulative seconds. The 6-s minimum value allowed infants to continue processing what they had seen before the trial could end. Across all test trials in Experiments 1–3, children took 6.71 s ($SE = 0.16$), on average, to complete the 6-s minimum look. Inter-observer agreement in the final phase of each test trial was calculated by determining the proportion of 100-ms intervals in which the two observers agreed. Across Experiments 1–3, agreement was calculated for 91/96 children and averaged 94% per trial.

Analyses of the familiarization and pretest data in Experiments 1–2 revealed no significant effect of condition (see SOM). Preliminary analyses of the test data in Experiments 1–3 revealed no significant interaction of condition and event with child's sex, TargetE's outfit/label, side of SimilarE's window, side of TargetE's first look in the test trials, and test order; the data were therefore collapsed across these latter five factors.

The data from Experiments 1–3 are available via the Open Science Framework at https://osf.io/k5fb2/?view_only=3f4acd20192d4aa9a41d77d56d888571.

Results

Looking times during the final phases of the test trials (Fig. 2) were analyzed using an ANOVA with condition (two-group, no-group) as a between-subject factor and event (approach-

SimilarE, approach-DifferentE) as a within-subject factor. The analysis yielded a marginally significant main effect of event, $F(1, 30) = 3.571, p = .068, \eta_p^2 = 0.106, 90\%-CI[0.000, 0.285]$, and a significant Condition \times Event interaction, $F(1, 30) = 7.793, p = .009, \eta_p^2 = 0.206, 90\%-CI[0.032, 0.389]$. Planned pairwise comparisons revealed that infants in the two-group condition looked significantly longer at the approach-DifferentE ($M = 29.58, SE = 4.20$) than at the approach-SimilarE ($M = 19.88, SE = 2.98$) event, $F(1, 30) = 10.957, p = .002, \eta_p^2 = 0.268, 90\%-CI[0.065, 0.446]$, whereas infants in the no-group condition looked equally at the approach-DifferentE ($M = 19.31, SE = 3.78$) and approach-SimilarE ($M = 21.18, SE = 2.78$) events, $F(1, 30) = 0.407, p = .528, \eta_p^2 = 0.013, 90\%-CI[0.000, 0.138]$. Non-parametric Wilcoxon signed-rank tests confirmed the results of the two-group ($Z = 2.637, p = 0.008$) and no-group ($Z = 0.724, p = 0.469$) conditions.

Discussion

Infants in the two-group condition adopted a ritual stance toward the experimenters' outfits, interpreted them as group markers, and brought to bear an expectation of ingroup support to predict whom TargetE would prefer. Conversely, infants in the no-group condition adopted an instrumental stance toward the experimenters' outfits, dismissed them as group markers, and had no alternative basis for predicting whom TargetE would prefer.

Experiment 2

Experiment 2 used a different instrumental-stance manipulation, with both outfits visible from the start. In two familiarization trials, SimilarE and DifferentE had identical picture books (TargetE was absent), and either SimilarE (*SimilarE* event) or DifferentE (*DifferentE* event) also had a cup. In the two-group condition, whoever had a cup drank from it before putting it away and reading her book. In the no-group condition, DifferentE again drank from her cup, but SimilarE used a sponge attached to her outfit to clean messy lines on her cup, before putting it away and

reading her book. If infants adopted an instrumental stance when SimilarE used her outfit to clean her cup, results should replicate those of Experiment 1.

Method

Design

Infants in each condition received two familiarization and two test trials (Fig. 3). In the two-group condition, the order of the SimilarE and DifferentE familiarization events was counterbalanced. At the start of the (30-s) initial phase in the SimilarE event, SimilarE and DifferentE sat facing their closed books. While DifferentE looked at her book, SimilarE picked up a red cup from the apparatus floor, silently drank from it twice, put it out of the apparatus (on a stool), and returned to her book. The trial then proceeded as in the pretest trial of Experiment 1: Each experimenter flipped through three pages and then paused until the trial ended. The DifferentE event was identical except that DifferentE drank from a green cup. The test trials were identical to those in Experiment 1.

The no-group condition was identical with four exceptions. First, two small light-blue sponges were attached to the forehead band of the pink outfit (on either side of the flower), and two small orange sponges were attached to the red scarf of the blue outfit. Second, each cup had a white band covering its midsection. Third, in the SimilarE event, which was always shown first, messy dark lines could be seen on the band of SimilarE's cup; she frowned at the lines, used one of her sponges to silently remove them, smiled at her clean cup, put it away, and readjusted her outfit (i.e., replaced her forehead band or smoothed down her scarf) as she returned to her book. Finally, the test trials were identical to those in the two-group condition except for the sponges on the experimenters' outfits.

Participants

Participants were 32 term 12-month-olds (15 male, $M = 12;8$, range = 11;11–12;23). Another 5 infants were excluded: 3 were distracted (e.g., by a pacifier), 1 looked the maximum allowed in both test trials, and 1 (in the no-group condition) had a difference between her two test trials that was over 3 standard deviations from the condition mean. Sixteen infants were randomly assigned to each condition.

Apparatus and Procedure

The apparatus and procedure were similar to those in Experiment 1 except for the changes noted above. Stimuli included the two picture books from Experiment 1 and two tall cups, one red and one green; both cups were empty, and the experimenters pretended to drink from them. Infants were highly attentive in the initial phases of the familiarization and test trials and looked, on average, for 95% of each initial phase.

Results

Looking times during the final phases of the test trials (Fig. 2) were analyzed as in Experiment 1. The analysis yielded a significant main effect of event, $F(1, 30) = 4.858, p = .035, \eta_p^2 = 0.139, 90\text{-CI}[0.005, 0.322]$, and a significant Condition \times Event interaction, $F(1, 30) = 5.184, p = .030, \eta_p^2 = 0.147, 90\text{-CI}[0.008, 0.330]$. Infants in the two-group condition looked significantly longer at the approach-DifferentE ($M = 25.69, SE = 3.62$) than at the approach-SimilarE ($M = 16.45, SE = 1.83$) event, $F(1, 30) = 10.040, p = .004, \eta_p^2 = 0.251, 90\text{-CI}[0.055, 0.431]$, whereas those in the no-group condition looked equally at the approach-DifferentE ($M = 17.08, SE = 2.33$) and approach-SimilarE ($M = 17.23, SE = 2.33$) events, $F(1, 30) = 0.003, p = .959, \eta_p^2 < 0.001, 90\text{-CI}[0.000, 0.000]$. Wilcoxon signed-rank tests confirmed the results of the two-group ($Z = 3.154, p = 0.002$) and no-group ($Z = 0.517, p = 0.605$) conditions.

Discussion

As in Experiment 1, infants expected TargetE to prefer SimilarE when the experimental context allowed a ritual construal of the experimenters' outfits, but they held no such expectation when the context supported an instrumental construal instead.

Experiment 3

To provide converging evidence for our conclusion that early sensitivity to similarity as a group marker is modulated by contextual information, Experiment 3 tested 26-month-old toddlers using novel labels ("lutak", "scobbie"), instead of novel outfits. In two familiarization trials, TargetE and SimilarE uttered one label (counterbalanced), and DifferentE uttered the other label; labels were inserted in "I am an X!" (two-group condition) or "I saw an X!" (no-group condition) phrases (Jin & Baillargeon, 2017). In the test trials, TargetE asked SimilarE or DifferentE whether she could read with her. If toddlers construed the experimenters' labels as group markers in the two-group but not the no-group condition, results should be the same as before.

Method

Design

Toddlers in each condition received two identical familiarization trials (to help them remember what each experimenter said), one pretest trial, and two test trials (Fig. 4). In contrast to Experiments 1–2, the familiarization and pretest trials were computer-controlled; we were concerned that seeing multiple paused scenes in which the experimenters did not interact might affect toddlers' test responses.

Only the *familiarization* trials differed between conditions. In each (20-s) trial, the three experimenters spoke in turn, in two rounds from right to left, using the phrases appropriate for each condition (e.g., two-group: SimilarE: "I'm a lutak!", TargetE: "I'm a lutak, too!", DifferentE: "I'm a scobbie!"; no-group: SimilarE: "I saw a lutak!", TargetE: "I saw a lutak, too!", DifferentE:

“I saw a scobbie!”). The experimenters looked at each other naturally as they spoke.

The (25-s) *pretest* trial was similar to that in Experiment 1: TargetE was absent, and SimilarE and DifferentE read identical picture books until the trial ended. Finally, as in Experiments 1–2, the *test* trials had an initial phase and final phase. In the 13-s initial phase of the trial, TargetE looked at SimilarE and DifferentE (counterbalanced) and their closed books twice, and then she asked either SimilarE (approach-SimilarE event) or DifferentE (approach-DifferentE event) if she could read with her (“Can I read with you?”, stated three times). SimilarE and DifferentE looked at her as she spoke, and then all three experimenters looked down and paused until the trial ended.

Participants

Participants were 32 26-month-olds (16 male, $M = 26;10$, range = 25;0-27,8). Another 3 toddlers were excluded because they were fussy (1), distracted (1), or spoke and caused the parent and experimenters to laugh (1). Sixteen toddlers were randomly assigned to each condition.

Apparatus and Procedure

The apparatus and procedure were similar to those in Experiment 1 except for the changes noted above. In addition, the experimenters wore different plain shirts (green, blue, and red) that provided no group markers, and the primary observer left the room during the familiarization trials to remain naïve about the experimenters’ utterances. Toddlers were highly attentive during the (computer-controlled) familiarization and pretest trials and during the initial phases of the test trials; they looked, on average, for 98% of each trial.

Results

Looking times during the final phases of the test trials (Fig. 2) were analyzed as in Experiment 1. The analysis yielded a significant main effect of event, $F(1, 30) = 4.676$, $p = .039$,

$\eta_p^2 = 0.135$, 90%-CI[0.004, 0.317], and a significant Condition \times Event interaction, $F(1, 30) = 6.043$, $p = .020$, $\eta_p^2 = 0.168$, 90%-CI[0.015, 0.351]. Toddlers in the two-group condition looked significantly longer at the approach-DifferentE ($M = 37.68$, $SE = 5.18$) than at the approach-SimilarE ($M = 21.18$, $SE = 2.91$) event, $F(1, 30) = 10.675$, $p = .003$, $\eta_p^2 = 0.262$, 90%-CI[0.062, 0.441], whereas those in the no-group condition looked equally at the approach-DifferentE ($M = 31.35$, $SE = 3.81$) and approach-SimilarE ($M = 32.41$, $SE = 4.08$) events, $F(1, 30) = 0.044$, $p = .836$, $\eta_p^2 = 0.001$, 90%-CI[0.000, 0.070]. Wilcoxon signed-rank tests confirmed the results of the two-group ($Z = 2.379$, $p = 0.017$) and no-group ($Z = 0.000$, $p = 1.000$) conditions.

Discussion

As expected, toddlers expected TargetE to prefer SimilarE when they construed the experimenters' labels as group markers, but not otherwise.

Overall Analyses

To assess the robustness of our findings, we conducted two additional sets of analyses. In the first set, we combined the test data of Experiments 1–3 and performed two random-effects meta-analyses, one for each condition (Cumming, 2014). Children in the two-group condition ($N = 48$) looked significantly longer at the approach-DifferentE than at the approach-SimilarE event, mean effect size Cohen's $d^+ = 0.79$, 95%-CI[0.38, 1.21], $Z = 3.73$, $p < .001$, whereas children in the no-group condition ($N = 48$) looked equally at the events, mean effect size Cohen's $d^+ = -0.08$, 95%-CI[-0.48, 0.32], $Z = 0.38$, $p = 0.708$. Fixed-effects meta-analyses yielded identical results. Finally, heterogeneity Q -tests yielded non-significant results, $ps \geq .954$, suggesting that in each condition, Experiments 1–3 estimated the same effect size.

The second set of analyses focused on the alternative possibility (mentioned in the Introduction) that children might initially expect similar individuals to belong to the same group

and support each other irrespective of the context in which they observed these similar individuals. This possibility predicted that across Experiments 1–3, children in the no-group condition would look differentially (experimental hypothesis) as opposed to equally (null hypothesis) at the two test events. To assess whether the negative result obtained in that condition merely failed to reject the null hypothesis or actually provided evidence for it, we conducted a Bayes factor (BF) analysis using the Jeffreys-Zellner-Siow (JZS) prior (Rouder et al., 2009). According to conventional cut-offs, a BF above 3 indicates at least moderate support for a hypothesis (Jarosz & Wiley, 2014). We obtained a Scaled JZS BF of 5.66 in favor of the null hypothesis, indicating that the data of the no-group condition were over five times more likely to occur under the null than the experimental hypothesis. (For the two-group condition, we obtained a Scaled JZS BF of 4357.91 in favor of the experimental hypothesis.) Whether similar individuals are expected to belong to the same group and support each other is thus modulated by contextual information beginning early in life.

General Discussion

Infants and toddlers expected TargetE to prefer SimilarE over DifferentE when the experimental context allowed them to interpret the similarity between TargetE and SimilarE as a signal that the two belonged to the same group. However, they held no expectation about whom TargetE would prefer when the context undermined such an interpretation. Thus, in Experiments 1–2, infants dismissed the similar outfits worn by TargetE and SimilarE as a group marker when these outfits were used to fulfill an instrumental purpose (e.g., store or clean objects). Likewise, in Experiment 3, toddlers dismissed the similar labels spoken by TargetE and SimilarE as a group marker when these labels appeared to convey only incidental information about the two experimenters (e.g., “I saw a lutak!”, “I saw a lutak, too!”). Together, these results provide

converging evidence across two ages (12 and 26 months), two types of similarities (outfits and labels), and two types of manipulations (instrumental and incidental) that young children refrain from interpreting a salient similarity as a signal to group membership when contextual information suggests that this similarity carries little social significance.

Our results provide further evidence that an expectation of ingroup support emerges early in life, and that minimal similarities are sufficient to trigger this expectation. Our results also break new ground by making clear that the basic structure of human moral cognition includes not only an abstract principle of ingroup support but also an abstract concept of group. Beginning early in life, children do not use any or all salient similarities among individuals to form expectations about their interactions; only similarities that are interpreted as signaling group membership are used for this purpose, and subtle computations encompassing a wide range of factors contribute to these determinations. Future research can further explore these computations. For example, in what contexts might infants refrain from using meaningful (as opposed to minimal) similarities as group markers? Are there similarities that infants might be likely (or unlikely) to view as group markers irrespective of context? Could infants understand that some similarities are simulated rather than real and represent affiliative bids or shows of loyalty (e.g., white lies)? Such research should help us better understand the nature of the relation between similarity and group membership in early moral cognition.

Acknowledgment

This research was supported by John Templeton Foundation grants 52034 and 60847 to xx.

References

- Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin, 140*, 1556-1581.
- Begus, K., Gliga, T., & Southgate, V. (2016). Infants' preferences for native speakers are associated with an expectation of information. *Proceedings of the National Academy of Sciences, 113*, 12397-12402.
- Bian, L., Sloane, S., & Baillargeon, R. (2018). Infants expect ingroup support to override fairness when resources are limited. *Proceedings of the National Academy of Sciences, 115*, 2705-2710.
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*, 429-444.
- Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective imitation of in-group over out-group members in 14-month-olds infants. *Child Development, 84*, 422-428.
- Choi, Y. J., & Luo, Y. (2015). 13-month-olds' understanding of social interactions. *Psychological Science, 26*, 274-283.
- Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science, 17*, 1003-1011.
- Cirelli, L. K., Wan, S. J., & Trainor, L. J. (2016). Social effects of movement synchrony: Increased infant helpfulness only transfers to affiliates of synchronously moving partners. *Infancy, 21*, 807-821.
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*, 7-29.
- Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development, 82*, 793-811.
- 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.
- Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, 415, 755.
- Gerson, S. A., Bekkering, H., & Hunnius, S. (2017). Do you do as I do?: Young toddlers prefer and copy toy choices of similarly acting others. *Infancy*, 22, 5-22.
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 55-130.
- Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, 7, 2-9.
- Jin, K., & Baillargeon, R. (2017). Infants possess an abstract expectation of ingroup support. *Proceedings of the National Academy of Sciences*, 114, 8199-8204.
- Jordan, A., & Dunham, Y. (2020). Are category labels primary? Children use similarities to reason about social groups. *Developmental Science*, e13013.
- Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, 104, 12577-12580.
- Legare, C. H., Wen, N. J., Herrmann, P. A., & Whitehouse, H. (2015). Imitative flexibility and the development of cultural learning. *Cognition*, 142, 351-361.
- Lieberman, Z., Kinzler, K. D., & Woodward, A. L. (2018). The early social significance of shared ritual actions. *Cognition*, 171, 42-51.
- Lieberman, Z., Kinzler, K. D., & Woodward, A. L. (2021). Origins of homophily: Infants expect people with shared preferences to affiliate. *Cognition*, 212, 104695.

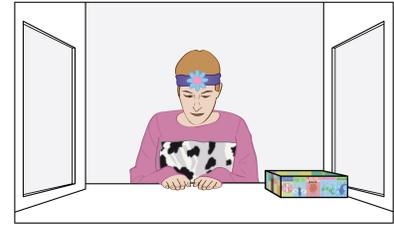
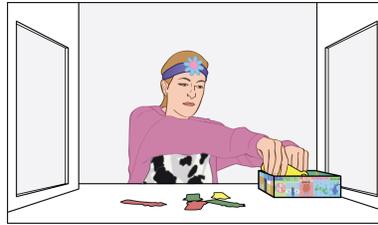
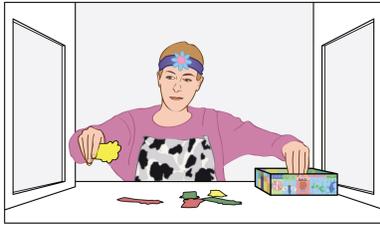
- Lieberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). Preverbal infants infer third-party social relationships based on language. *Cognitive Science*, *41*, 622-634.
- Mahajan, N., & Wynn, K. (2012). Origins of “Us” versus “Them”: Prelinguistic infants prefer similar others. *Cognition*, *124*, 227-233.
- Margoni, F., Baillargeon, R., & Surian, L. (2018). Infants distinguish between leaders and bullies. *Proceedings of the National Academy of Sciences*, *115*, E8835-E8843.
- Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, *110*, 3965-3972.
- Pun, A., Birch, S. J., Baron, A. S. (2021). The power of allies: Infants’ expectations of social obligations during intergroup conflict. *Cognition*, *211*, 104630.
- Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*, 57-75.
- Rhodes, M., Hetherington, C., Brink, K., & Wellman, H. M. (2015). Infants' use of social partnerships to predict behavior. *Developmental Science*, *18*, 909-916.
- Roberts, S. O., Williams, A. D., & Gelman, S. A. (2017). Children’s and adults’ predictions of Black, White, and Multiracial children’s friendship patterns. *Journal of Cognition and Development*, *18*, 189-208.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225-237.
- 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*, 1-17.
- Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, *23*, 196-204.

- Shweder, R. A., Much, N. C., Mahapatra, M. & Park, L. (1997). The “big three” of morality (autonomy, community and divinity) and the “big three” explanations of suffering. In A. M. Brandt & P. Rozin (Eds.), *Morality and health*, (pp. 119–169). New York: Routledge.
- Sparks, E., Schinkel, M. G., & Moore, C. (2017). Affiliation affects generosity in young children: The roles of minimal group membership and shared interests. *Journal of Experimental Child Psychology*, *159*, 242-262.
- Spokes, A. C., & Spelke, E. S. (2017). The cradle of social knowledge: Infants’ reasoning about caregiving and affiliation. *Cognition*, *159*, 102-116.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*, 149–178.
- Ting, F., Buyukozer Dawkins, M., Stavans, M., & Baillargeon, R. (2020). Principles and concepts in early moral cognition. In J. Decety (Ed.), *The social brain* (pp 41–65). Cambridge, MA: MIT Press.
- Ting, F., He, Z., & Baillargeon, R. (2019). Toddlers and infants expect individuals to refrain from helping an ingroup victim’s aggressor. *Proceedings of the National Academy of Sciences*, *116*, 6025-6034.
- Tooby, J., Cosmides, L., & Price, M. E. (2006). Cognitive adaptations for n-person exchange: The evolutionary roots of organizational behavior. *Managerial and Decision Economics*, *27*, 103–129.
- Watson-Jones, R. E., & Legare, C. H. (2016). The social functions of group rituals. *Current Directions in Psychological Science*, *25*(1), 42-46.

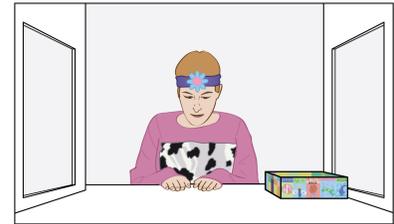
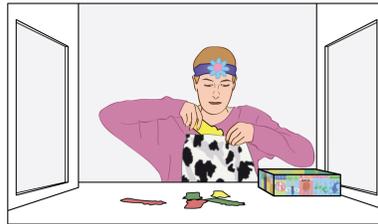
Experiment 1

Familiarization Trial

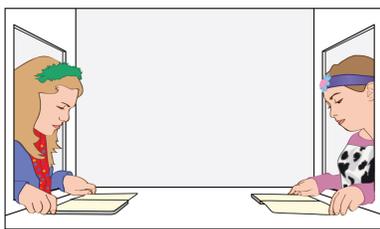
Two-Group Condition



No-Group Condition



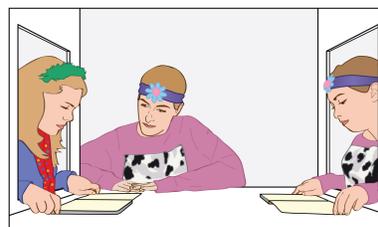
Pretest Trial



Test Trials



Approach-SimilarE Event



Approach-DifferentE Event

Fig. 1: Experiment 1 familiarization trial in the two-group and no-group conditions, followed by the pretest and test trials in both conditions.

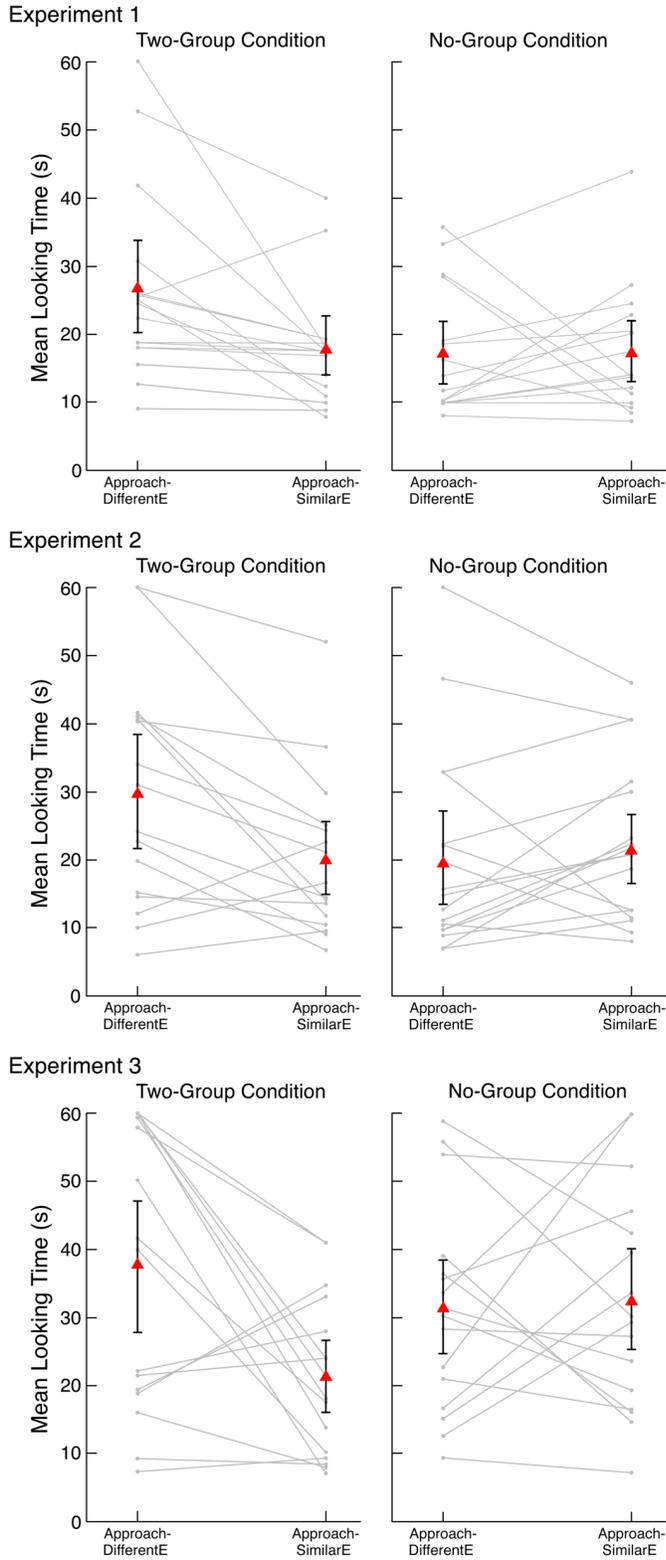
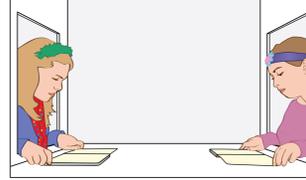
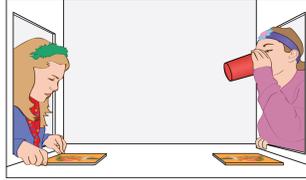
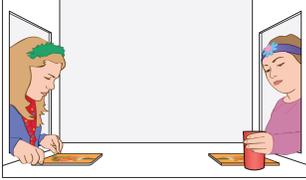


Fig. 2: Looking times at the test events, by condition, in Experiments 1 (N = 32), 2 (N = 32), and 3 (N = 32). Coral triangles indicate means. Pairs of connected grey dots indicate data from a single participant. Error bars represent 95%-CIs.

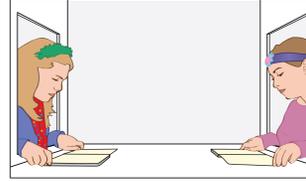
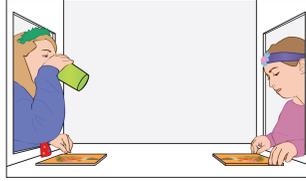
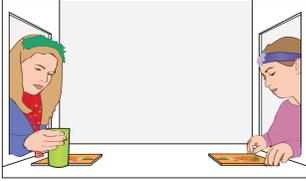
Experiment 2 Two-Group Condition

Familiarization Trials

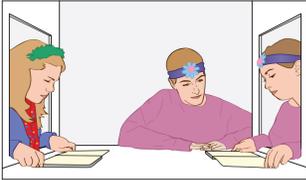
SimilarE Event



DifferentE Event



Test Trials



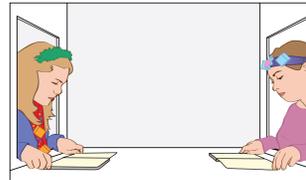
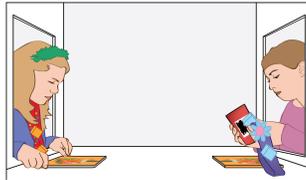
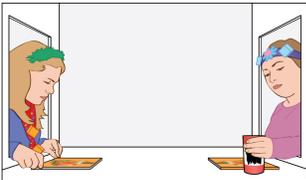
Approach-SimilarE Event



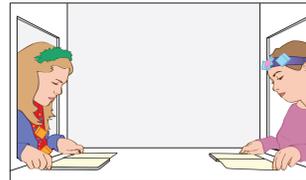
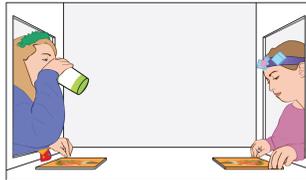
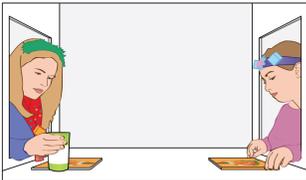
Approach-DifferentE Event

Familiarization Trials

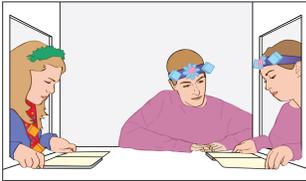
SimilarE Event



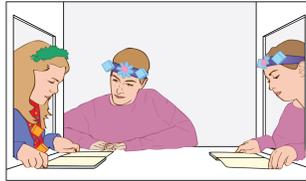
DifferentE Event



Test Trials



Approach-SimilarE Event

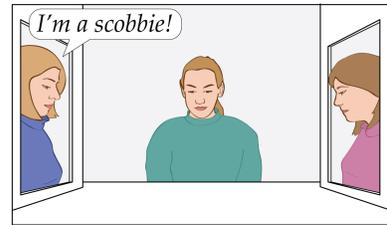
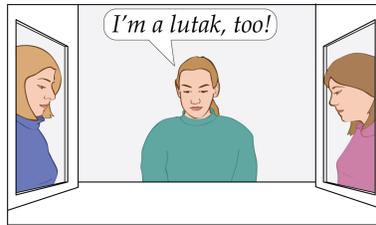
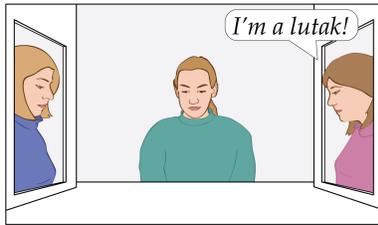


Approach-DifferentE Event

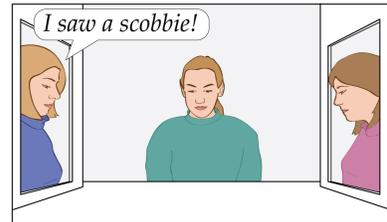
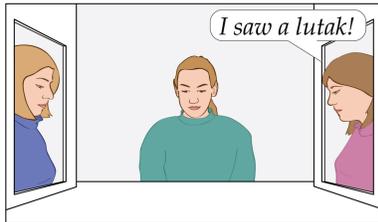
Fig. 3: Experiment 2 familiarization and test trials in the two-group and no-group conditions.

Experiment 3

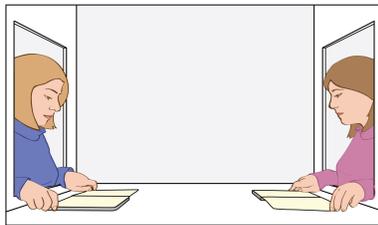
Familiarization Trials Two-Group Condition



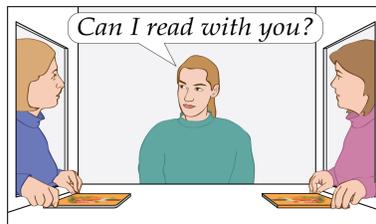
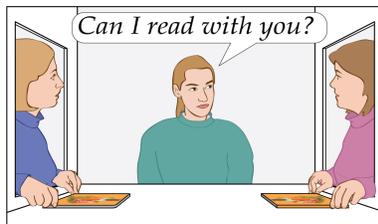
No-Group Condition



Pretest Trial



Test Trials



Approach-SimilarE Event

Approach-DifferentE Event

Fig. 4: Experiment 3 familiarization trials in the two-group and no-group conditions, followed by the pretest and test trials in both conditions.