MS. TESSA ELIZABETH SADIE CHARLESWORTH (Orcid ID : 0000-0001-5048-3088)

Article type      : Short Report

Children's implicit attitude acquisition: Evaluative statements succeed, repeated pairings fail

Tessa E.S. Charlesworth, Benedek Kurdi, & Mahzarin R. Banaji

Harvard University

Author Note

Correspondence concerning this article should be addressed to Tessa Charlesworth, Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138, email: tet371@g.harvard.edu.

Children's implicit attitude acquisition: Evaluative statements succeed, repeated pairings fail

## Research Highlights

- Understanding *how* and *when* implicit attitudes are acquired in childhood is important for theories of implicit social cognition and human learning.

- The separate and joint effects of acquiring implicit attitudes through *evaluative statements* (ES) and four variations of *repeated evaluative pairings* (REP) are compared.

- Children acquired implicit attitudes following ES and ES+REP, highlighting the surprisingly early-emerging power of verbal statements to create implicit attitudes.

- Children did *not* acquire implicit attitudes following REP variations, even those designed to maximize learning, suggesting an age-related limit in acquiring implicit attitudes from pairings.

**Abstract**

From the earliest ages tested, children and adults show similar overall magnitudes of implicit attitudes towards various social groups. However, such consistency in attitude magnitude may obscure meaningful age-related change in the ways that children (versus adults) acquire implicit attitudes. This experiment investigated children's implicit attitude acquisition by comparing the separate and joint effects of two learning interventions, previously shown to form implicit attitudes in adults. Children ($N = 280$, ages 7-11 years) were taught about novel social groups through either *evaluative statements* (ES; auditorily-presented verbal statements such as "Longfaces are bad, Squarefaces are good "), *repeated evaluative pairings* (REP; visual pairings of Longface/Squareface group members with valenced images such as a puppy or snake), or a combination of ES+REP. Results showed that children acquired implicit attitudes following ES and ES+REP, with REP providing no additional learning beyond ES alone. Moreover, children did not acquire implicit attitudes in four variations of REP, each designed to facilitate learning by systematically increasing verbal scaffolding to specify (*a*) the learning goal, (*b*) the valence of the unconditioned stimuli, and (*c*) the group categories of the conditioned stimuli. These findings underscore the early-emerging role of verbal statements in children's implicit attitude acquisition, as well as a possible age-related limit in children's acquisition of novel implicit attitudes from repeated pairings.


*Keywords:* associative learning, attitude acquisition, evaluative conditioning, implicit attitudes, propositional learning, social cognitive development

Children's implicit attitude acquisition: Evaluative statements succeed, repeated pairings fail

From the earliest ages tested, children show adult-like levels of automatic social group evaluations (*implicit attitudes*; Baron & Banaji, 2006; Dunham, Baron, & Banaji, 2008), suggesting that such attitudes come online early and may not change across the lifespan. This age-related similarity appears robust across attitudes toward Black vs. White Americans (Baron & Banaji, 2006; Newheiser & Olson, 2012), Hispanic vs. White Americans (Dunham, Baron, & Banaji, 2007), religious groups (Heiphetz, Spelke, Harris, & Banaji, 2013), and even minimal groups (Dunham, Baron, & Carey, 2011). This "developmental invariance" also generalizes across cultures (Cvencek, Meltzoff, & Kapur, 2014; Dunham, Newheiser, Hoosain, Merrill, & Olson, 2014; Qian et al., 2016; Rutland, Cameron, Milne, & Mcgeorge, 2005; Steele, George, Williams, & Tay, 2018). Such consistency suggests that "developmental invariance may represent a core property of the implicit associative system" (p. 51, Baron, 2015).

However, similarity observed in children and adults' mean-levels of implicit attitudes need not imply underlying similarity in how they *acquire* implicit attitudes (Baron, 2015). While much has been learned about adult's implicit attitude acquisition (for a review see Cone, Mann, & Ferguson, 2017), no study has systematically compared multiple learning interventions to examine *children's* implicit attitude acquisition. The present research addresses this question, yielding insights into fundamental processes of developing implicit attitudes, a core component of social cognition.

**Mechanisms of Implicit Attitude Acquisition.**

Much recent work on learning and attitude acquisition has made a distinction between two fundamentally different modalities of learning: (1) repeated *pairings* of evaluative images and category exemplars (evaluative conditioning), and (2) explicit rule-based evaluative *statements* about categories. Leading associative and dual-process theories posit that implicit attitudes are uniquely, or at least primarily, acquired by the former process (pairings) and therefore only represent stimulus associations (i.e., *Laapians – Good*, Baeyens, Eelen, Crombez, & van den Bergh, 1992; Rydell & McConnell, 2006; Strack & Deutsch, 2004). In contrast, propositional theories emphasize that implicit attitudes can be acquired not only through pairings, but *also* from rule-based statements and, as such,

can represent explicit relational information (i.e., *"Laapians are good"*, De Houwer, 2014, 2018; De Houwer & Hughes, 2016; Zanon, De Houwer, Gast, & Smith, 2014).

Adult learners can acquire novel implicit attitudes from both types of learning interventions: repeated evaluative pairings (REP; visually-presented pairings of group exemplars with valenced images) and verbal evaluative statements (ES; brief statements about which group goes with which valenced image), as well as their combination (ES+REP), consistently produce learning above baseline (Kurdi & Banaji, 2017, 2019). Additionally, although all interventions produce learning, ES (and ES+REP) produce stronger implicit attitudes than REP.

Thus, if developmental invariance is indeed a core property of implicit social cognition, then children, like adults, should acquire implicit attitudes following *all* learning interventions, with superiority of ES over REP. If, however, age-related similarity emerges only in attitude magnitude but not in modalities of learning, children may show unique patterns of acquisition (not acquiring attitudes from REP and/or ES). Specifically, children may differ from adults in not acquiring implicit attitudes from either intervention, perhaps requiring more diagnostic information to make an inductive inference about a group (Schulz, 2012). Alternatively, children may acquire implicit attitudes from only one intervention: only from REP (not ES), perhaps due to verbal and executive function demands of maintaining the statement in mind (Munakata, Snyder, & Chatham, 2012); or only from ES (not REP), perhaps due to additional inferential demands in learning from pairings (Kurdi & Banaji, 2017; Van Dessel, Hughes, & De Houwer, 2019). Research supporting each of these possibilities is elaborated below.

**Age-related Similarity or Change in Implicit Attitude Acquisition.**

In an initial test of age-related similarity in implicit attitude acquisition, children (ages 5-12) were found to acquire novel implicit attitudes following a detailed verbal story (Gonzalez, Dunlop, & Baron, 2016). Similarly, children in the present experiment may acquire implicit attitudes following verbal evaluative statements (ES).

However, past research used extensive stories that increase rehearsal, immersion in the vivid narrative, and evidence of diagnostic evaluative properties (a group's character traits or behaviors), often argued to be especially conducive to implicit attitude acquisition (e.g., Cone et al., 2017; Mann, Kurdi, & Banaji, in press). By contrast, the ES intervention in the present work provides only

contingency information about the co-occurrence between groups and evaluative stimuli (positive/negative pictures). Thus, ES is notably minimal in that it (a) consists of only a single sentence that limits rehearsal and vividness, and (b) provides no diagnostic evaluative evidence. ES may therefore be insufficient for children's acquisition and expression of implicit attitudes.

Furthermore, children may have difficulty in ES because the ability to form, maintain, and express rules requires executive control and prefrontal cortex development (Davidson, Amso, Cruess Anderson, & Diamond, 2006; Munakata et al., 2012). This is especially true for competing rules, such as remembering the evaluative statement (e.g., "group X is paired with *good*") versus the requirements to complete the Implicit Association Test (IAT) (e.g., "group X is sorted with *bad* in this part of the task"). Thus, children may show weak or no implicit attitude acquisition following ES.

To date, no study has examined developmental patterns of implicit attitude acquisition from pairings. However, associative theories would suggest that pairings will be equally effective for children and adults since implicit attitude acquisition from pairings relies on low-level stimulus-driven mechanisms rather than high-level reasoning (Rydell & McConnell, 2006; Strack & Deutsch, 2004). Additionally, even young infants track co-occurrences between stimuli to acquire information ranging from causality (e.g., one ball hitting co-occurs with another ball moving, Gopnik, Sobel, Schulz, & Glymour, 2001) to word boundaries (Saffran, Aslin, & Newport, 1996; for a review see Santolin & Saffran, 2017). Finally, 3-year-old children appear to acquire explicit attitudes from pairings (Halbeisen, Walther, & Schneider, 2016). However, that study measured attitudes towards single *individuals*, not groups, on *explicit*, not implicit measures, which could arise from children simply reporting back what was presented to them but without acquiring an underlying attitude. Nevertheless, such evidence suggests that children may acquire implicit attitudes following REP.

On the other hand, learning from pairings appears to improve with age in visual statistical learning (Arciuli & Simpson, 2011), classical conditioning (O'Donnell & Brown, 1973), and evaluative conditioning (Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010). Indeed, children's meta-analytic learning effect from evaluative conditioning is small ($d = .11$) and not significant, unlike the robust effect for adults ($d = .52$; Hofmann et al., 2010). Thus, children may show weak or no implicit attitude acquisition following REP.

A final possibility is that children may require additional information to scaffold their learning from REP. Recent perspectives argue that learning from pairings requires multiple inferential steps (Kurdi & Banaji, 2017; Van Dessel et al., 2019). For instance, in the current experiment, the learner must infer that: the goal is to learn the relationship between stimuli (*goal inference*); unconditioned stimuli (US) are organized by positive/negative valence (*valence inference*); conditioned stimuli (CS) are organized into two group categories (*group inference*, e.g., Baron, Dunham, Banaji, & Carey, 2014); and CS and US categories have a meaningful relationship (*relational inference,* e.g., Zanon et al., 2014). Facilitating these inferential steps may be crucial for children's implicit attitude acquisition from pairings.

In sum, competing predictions can be offered on when and how children acquire implicit social group attitudes from evaluative statements and/or pairings. This experiment provides the first test directly comparing multiple learning manipulations to begin to isolate the conditions for children's implicit attitude acquisition, with implications for attitude theory and interventions.

**Method**

**Participants.**

A sample of 316 children was recruited from a children's museum and local parks in Boston and Cambridge, MA. The age range (7-11 years) ensured that children could complete the child-adapted IAT (Baron & Banaji, 2006) and detect group boundaries without explicit labels (Baron et al., 2014). Data were omitted for: failure to complete the IAT (4 participants); experimenter/computer error (1); misunderstanding instructions or random responding (6); and being outside the age range (25).

The final sample consisted of 280 children, $M_{age}$ = 9.07 years [7.00; 10.95 years], $SD_{age}$ = 1.14 years, half of whom were identified by their parents as female ($N$ = 155) and White ($N$ = 146). The remaining participants were Asian ($N$ = 76), other race ($N$ = 46), or Black/African-American ($N$ = 12). Children were distributed across 7 conditions, defined below (39 Control, 39 ES, 44 ES+REP, 42 $REP_{min.instruct}$, 39 $REP_{max.instruct}$, 38 $REP_{val.label}$, 39 $REP_{group.label}$). Explicit attitude data was incomplete for 4 participants, leaving 276 participants for explicit attitude analyses.

**Materials.**

Stimuli were chosen to be identical to those used previously with adults (Kurdi & Banaji, 2017). Conditioned stimuli (CS) were morphed photographs of White men transformed to create two distinct groups: "Longfaces" (length-to-width ratio 2:1) or "Squarefaces" (length-to-width ratio 4:3; Figure 1a). Valenced stimuli (unconditioned stimuli, US) were cartoons of intrinsically positive and negative objects (Figure 1b).

**Procedure.**

*Learning Phase.* Participants completed the experiment individually on a computer, with an experimenter present to read instructions and ensure attention. Participants were assigned to a control condition or one of six learning conditions, described below[1]. Because (1) the groups were novel and assumed to be neutral, (2) previous research showed no baseline implicit preference for Squarefaces (a finding also replicated in this experiment), and (3) the *relative* effect of each condition was of primary interest, only one pairing was implemented and all children were taught that Squarefaces are good and Longfaces are bad.

Learning conditions differed in the amount and type of information provided to participants (see Table 1): (1) exposure to CS–US pairings, (2) learning goal (explicit instruction to learn the relationship between CSs and USs), (3) valence labels (verbally labeling USs as positive/negative), (4) group labels (verbally labeling CSs as Longfaces/Squarefaces), and (5) relational description (verbal instruction that Squarefaces will be paired with positive images and Longfaces with negative images). Verbatim instruction text is provided in supplemental materials.

In the control condition, children viewed 16 US–US pairings[2] (i.e., "good" pictures with "good" pictures, "bad" pictures with "bad" pictures) and received no further information. In the ES condition, participants were verbally informed about upcoming pairings (providing valence, group, and relational information) but did not view any actual pairings. In the ES+REP condition, participants received verbal statements (as in ES) and exposure to pairings (as in REP). All REP conditions provided 16 simultaneous presentations of Squarefaces/positive US images and Longfaces/negative US images, randomized for each participant. Each pairing was presented for 2,500 ms, followed by a 1,000-ms intertrial blank screen.

REP conditions systematically increased the information provided to participants: (1) $REP_{min.instruct}$ provided no instructions beyond the general statement that pictures of groups would

appear with other images; (2) $REP_{max.instruct}$ provided the explicit learning goal to learn the relationship between two types of groups and two kinds of images; (3) $REP_{labels}$ additionally provided valence labels for the US ("good images" vs. "bad images"); and (4) $REP_{group.labels}$ additionally provided group labels for the CS ("Squarefaces" vs. "Longfaces"). Notably, no REP condition provided explicit relational information because, unlike in ES and ES+REP, children were never told that Squarefaces *are* positive or Longfaces *are* negative.

      *Test Phase.* After learning, children completed tests of implicit attitudes, explicit attitudes, and memory for the learning manipulation.

      *Implicit attitudes.* Implicit attitudes were assessed using the child-adapted Implicit Association Test (Baron & Banaji, 2006). Like the adult-IAT, the Child-IAT measures the strength of association between two groups (Longfaces/Squarefaces) and two attributes (good/bad). Response latencies to categorize group and attribute stimuli are compared across two critical blocks: (1) a learning-congruent block, where Squarefaces/good words (and Longfaces/bad words) are assigned the same buttons; and (2) a learning-incongruent block, where Longfaces/good words (and Squarefaces/bad words) are assigned the same buttons. Attribute stimuli are presented auditorily through headphones to reduce reading demands, and large color-coded categorization buttons are used to reduce dexterity demands. Faster responses are assumed to reflect stronger associations between the group and attribute (stronger Squarefaces-good/Longfaces-bad association), yielding higher IAT D-scores.

      *Explicit attitudes and memory check.* After the Child-IAT, children completed two forced-choice explicit attitude measures in randomized order ("Which of these groups do you think is better?" "Which of these groups would you like to play with?"). Additionally, to measure retrospective recall for the instructions or pairing contingencies, children completed a forced-choice memory check ("Which of these groups was paired with good pictures?"). For all questions, pictures of Longfaces/Squarefaces were presented side-by-side and children pointed to their answer. Children and their parent/guardian were debriefed and offered a handstamp or sticker.

      **Analytic Strategy.**

      Primary analyses were conducted in a Bayesian framework[3], offering advantages over more typical frequentist approaches in the ability to (1) quantify evidence for (or against) the null, (2) provide interpretable estimates of the weight of this evidence as percentages (rather than non-intuitive

p-values and confidence intervals), and (3) accommodate smaller sample sizes, typical in developmental research. These advantages are achieved because Bayesian models provide a probability distribution over all potential values of coefficients given a certain sample (e.g., mean differences between conditions), known as the *posterior distribution*. A *region of practical equivalence* (ROPE) can then be established with a range of values a priori deemed consistent with the null hypothesis (Kruschke, 2015). The percentage of the posterior distribution that falls within (or outside) the ROPE is then calculated to provide an interpretable index of the weight of evidence for (or against) the null.

In this paper, if more than 95% of the posterior distribution fell *outside* the ROPE, we inferred that the alternative hypothesis ($H_1$) was supported. Analogously, if more than 95% of the posterior fell *within* the ROPE, we inferred that the null hypothesis ($H_0$) was supported. Finally, if less than 95% of the posterior fell outside/within the ROPE, we used the percentage of the posterior falling outside (or within) the ROPE as a relative index of evidence in favor of $H_1$ (or $H_0$) without reaching firm conclusions.

In addition, the 95% *highest density interval* (HDI) provided a secondary criterion for inference. The HDI specifies the upper and lower bounds of 95% of the posterior distribution, indicating the most likely values of the coefficient. If the HDI does not include zero, then the coefficient can be interpreted as meaningfully different from zero (support for $H_1$).

*Implicit attitudes*. Child-IAT D-scores were computed using the improved scoring algorithm (Greenwald, Nosek, & Banaji, 2003). For implicit attitudes, the ROPE was the IAT D-score interval of [-0.15, 0.15], typically reported as an interval of "no bias" (http://implicit.harvard.edu). A Bayesian linear regression was fit using *rstanarm* (Goodrich, Gabry, Ali, & Brilleman, 2018) in R (R Core Team, 2017). Default (uninformative) priors and 10,000 iterations of the Markov Chain Monte Carlo sampler were used.

*Explicit attitudes*. Responses on the two forced-choice explicit questions were analyzed using a Bayesian generalized linear mixed-effects model with random intercepts for participants and items to account for dependencies. For the intercept (control condition), the probability estimate indicates the chance of providing an explicit Squareface=good response. For all slopes (all learning conditions), the estimates indicate an increase in the probability of providing a Squareface=good response, relative

to control. The ROPE of [0.35, 0.65] was centered around chance (0.50), with a range of +/- 0.15 points, similar to the implicit attitude null interval.

**Results**

### Implicit Attitudes.

As shown in Table 2 and Figure 2, children showed no preference for Squarefaces or Longfaces in the control condition, reinforcing that these groups provide a neutral reference point. In line with prior data from adults (Kurdi & Banaji, 2017), both ES and ES+REP produced large learning effects compared to control, with 99% and 100% of the posterior distributions, respectively, falling outside of the null interval. Additionally, no meaningful difference was found between ES vs. ES+REP, also replicating adult data and highlighting the primary role of ES (see supplemental materials).

Unlike ES and ES+REP, no variant of REP provided evidence for implicit attitude acquisition (Table 2). Strong evidence in favor of the null hypothesis was obtained in the most minimal REP variant ($REP_{min.instruct}$), in the REP variant providing additional instructions about learning goals ($REP_{max.instruct}$), and even in the REP variant providing additional valence (US) distinctions ($REP_{labels}$). The REP variant providing information about group (CS) distinctions (in addition to valence and learning goals, $REP_{group.labels}$) provided suggestive evidence in favor of the null (albeit more ambiguous): the HDI included zero, but only 76% of the posterior distribution fell within the null interval. Across all REP variants, it therefore appears that children did not acquire novel implicit attitudes following exposure to stimulus pairings, even with additional scaffolding.

Implicit attitude acquisition was further examined among only those children who passed the memory check (see supplemental materials). Crucially, the pattern of results was unchanged, with no learning from REP even among children who explicitly recalled the pairings. This important finding rules out the possibility that children's lack of implicit attitude acquisition from REP was due to a lack of contingency memory or to some failure of the manipulations. Moreover, the result is notable given theorizing that contingency memory predicts adult's learning from pairings (Hofmann et al., 2010). By contrast, children may be less able or willing than adults to use contingency information to acquire and express implicit attitudes.

### Explicit Attitudes.

Children in the control condition revealed a small baseline explicit preference for Squarefaces (Table 3). Compared to control, children showed explicit attitude acquisition in both ES and ES+REP, although the percentage of posterior distributions outside the ROPE was slightly below the prespecified criterion (91% and 94%, respectively).

As with implicit attitudes, children in all REP variants showed little evidence of explicit attitude acquisition above control, with all 95% HDIs including zero. However, for all REP variants, the percentage of the posterior falling in the ROPE suggested inconclusive evidence for or against the null (Table 3). Nevertheless, it appears that children generally acquired strong explicit attitudes following ES and ES+REP but not following any REP condition.

**Memory Checks.**

Full results of memory checks are reported in supplemental materials. Compared to a chance (50%) baseline, children in ES and ES+REP were significantly above chance in recalling the accurate pairings (i.e., Squarefaces paired with good pictures). Accurate recall in the most minimal REP ($REP_{min.instruct}$) did not deviate from chance. Surprisingly, however, participants in all other REP variants were significantly above chance. Thus, while additional scaffolding in REP was not sufficient for children to acquire implicit attitudes, it nevertheless appeared to successively improve recall for contingencies. Again, this reinforces that the REP learning manipulation did not fail since most children recalled the contingencies. Instead, it suggests the novel result that contingency memory alone may not be sufficient for the acquisition and expression of implicit attitudes among children.

Additional analyses were conducted with memory performance in each condition compared to the control condition (supplemental materials). However, these analyses are difficult to interpret because, at baseline, children were above chance in reporting that Squarefaces were paired with positive images, despite having never seen any pairings. The most likely explanation for this result is that, in the absence of pertinent information, recall was driven by explicit preferences ("affect as information" effect; Hütter, Sweldens, Stahl, Unkelbach, & Klauer, 2012). Thus, these comparisons remain inconclusive[4].

**General Discussion**

This research explored the conditions that enable children's implicit attitude acquisition. In so doing, this experiment provided the first developmental test of ongoing theoretical debates between

associative and propositional models of attitude acquisition (Baeyens, Eelen, Crombez, & van den Bergh, 1992; De Houwer, 2014; 2018; Hughes & De Houwer, 2016; Rydell & McConnell, 2006; Strack & Deutsch, 2004). Moreover, the experiment offers insights into theories of social cognitive development and developmental change, and reinforces that understanding age-related similarity or change in attitudes requires investigating not only attitude magnitude but also attitude acquisition (Baron, 2015). Indeed, this experiment makes a novel contribution in revealing both age-related similarity and change in children's implicit attitude acquisition. Beyond their theoretical import, these findings also have implications for interventions: identifying the timing and mechanisms of acquisition can inform the optimal moments and methods to shift implicit attitudes toward neutrality (Gonzalez, Steele, & Baron, 2017).

**Mechanisms of Children's Implicit Attitude Acquisition.**

While children have been shown to acquire implicit attitudes from elaborate stories (Gonzalez et al., 2016), this experiment shows the surprising result that children also acquired attitudes from statements that were minimal both in their brevity and lack of diagnostic evaluative evidence (e.g., character traits, behaviors), instead providing only contingency information about groups and evaluative stimuli. Moreover, implicit attitude acquisition was not improved by experience with evaluative pairings (ES+REP), implying that statements drive learning. This power of explicit statements is particularly impressive considering that implicit attitudes were measured on an arguably "associative" test (the Implicit *Association* Test).

Crucially, the present experiment also showed that children at these ages do *not* yet acquire implicit attitudes from repeated evaluative pairings (REP). This result may be surprising given (1) the assumption that learning from pairings relies on low-level associative learning (Rydell & McConnell, 2006), as well as (2) children's capacities for tracking co-occurrences (Santolin & Saffran, 2017). On the other hand, this result aligns with emerging perspectives that learning from REP requires multiple inferences (Kurdi & Banaji, 2017; Van Dessel et al., 2019). Successful learning from REP may require as many as four inferences: the goal of the task (*goal inference*); the valence of the unconditioned stimuli (US) (*valence inference*); the group structure of the conditioned stimuli (CS) (*group inference*); and the relationship between CS and US (*relational inference*). While adults

appear to make these inferences spontaneously, this experiment tested whether children require additional scaffolding for learning from REP.

Surprisingly, even after providing assistance to encourage goal, valence, and group inferences, children failed to acquire implicit attitudes from pairings. Thus, in line with propositional theories (De Houwer, 2014, 2018; De Houwer & Hughes, 2016), the data highlight the importance of the final *relational* inference, where the learner must infer an equivalence relationship (i.e., that group X *is* good). The present REP paradigm may not be sufficient for children to make and use this inference in forming and expressing their attitudes. Rather, providing explicit relational information (as in ES) may be necessary for implicit attitude acquisition.

**Age-related Similarity or Change in Implicit Attitudes.**

Discussions of age-related similarity versus change in implicit attitudes (Baron, 2015) can be informed by comparing the present data with published data from adults using similar stimuli and designs (Kurdi & Banaji, 2017, 2019). Children's robust learning following ES and ES+REP implies age-related similarity in acquiring implicit attitudes from verbal instruction (see also Gonzalez et al., 2016). Yet children's absence of implicit attitude acquisition following REP suggests a possible source of age-related change. As discussed, this change may be due to concurrent developments in the ability to spontaneously infer stimuli relationships. Future research could examine older children to identify when such capacities come online, as well as directly measure age-related capacities presumed to scaffold acquisition, such as the development of relational inferences, inferences about experimenter/teacher intentions or accuracy (Koenig & Harris, 2005; Schulz, 2012), and categorization abilities (Bigler & Liben, 1992).

Finally, given that age-related change was identified by offering the first comparison across learning manipulations (ES vs. REP variants), additional comparisons may reveal when and why age-related similarity/change emerges in implicit social cognition. Given the many interventions that have been compared in adults (Lai et al., 2014), extending systematic comparisons into childhood seems warranted.

**Limitations.**

This experiment compared four variants of repeated evaluative pairings (REP), providing both internal replications of our results and, more important, identifying possible inferential prerequisites

for acquiring implicit attitudes from pairings. Although adequate for a first test, these variants do not exhaust all possibilities. For example, although the number of pairings does not appear to affect *adults'* learning from REP (Kurdi & Banaji, 2019), it is possible that, for children, effective learning requires more than 16 co-occurrences. Furthermore, the pairings' base-rate may moderate learning, with slightly variable base-rates (e.g., 75% Squarefaces=bad/Longfaces=good) resulting in greater learning (Brauer, Er-rafiy, Kawakami, & Phills, 2012). Finally, learning may be facilitated when pairings are repeated across multiple occasions, shown to more effectively alter implicit evaluations in both adults (McNulty, Olson, Jones, & Acosta, 2017) and children (Qian et al., 2017).

An additional feature was the use of novel fictional groups (Longfaces/Squarefaces), which help control for baseline variations in previous experiences and preferences and isolate the psychological mechanisms of interest. Nevertheless, it is also important to understand whether patterns of age-related change in implicit attitude acquisition extend to real-world social groups or age-relevant stimuli. In particular, real group boundaries have societal significance for children's interactions and are repeatedly reinforced for young children (e.g., gender boundaries are reinforced in peer, parent, and teacher interactions, Maccoby, 1990). This additional repetition and significance may be pivotal for children's implicit attitude acquisition.

**Conclusion.**

Despite its relevance to both theory and application, the question of how and when children acquire implicit attitudes has only begun to be explored. In this experiment, children (ages 7-11) rapidly acquired novel implicit attitudes from minimal verbal evaluative statements but not from repeated evaluative pairings. Even with additional scaffolding of learning goals, valence, and group categories, children showed little evidence of learning from purportedly low-level associative information. Thus, while age-related similarity emerges in acquiring implicit attitudes from statements, age-related change is suggested in acquiring implicit attitudes from stimulus pairings. Such results underscore the importance of comparing multiple learning interventions to identify nuanced developmental patterns in implicit social cognition. Finally, this research can help shape recommendations regarding how and when certain interventions are optimal for shifting implicit attitudes.

# Endnotes

Due to data resource limitations, collection first involved random assignment to Control, ES, ES+REP, or $REP_{min.instruct}$ and proceeded sequentially through each REP variant within a span of 6 months.

[2] Note that Kurdi and Banaji (2017) presented adult participants with 37 trials of pairings. A smaller number of pairings was used to increase children's engagement, reduce participant fatigue, and conform to task length requirements at the data collection sites. Moreover, recent work with adults has indicated that exposure to fewer trials in this paradigm is unlikely to reduce the magnitude of learning in REP, with any effects of learning emerging within the first 5 trials (Kurdi & Banaji, 2019).

[3] Frequentist analyses (providing identical conclusions) are reported in the supplemental materials.

[4] To address the possibility that children were relying on explicit attitudes for their memory checks in the control condition, future research may benefit from including a third "don't know" answer option.

# Data Availability Statement

All materials, data, and analysis scripts for this project are available from the Open Science Framework (https://osf.io/c49fg/).

## References

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science*, *14*(3), 464–473. https://doi.org/10.1111/j.1467-7687.2009.00937.x

Baeyens, F., Eelen, P., Crombez, G., & van den Bergh, O. (1992). Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behaviour Research and Therapy*, *30*(2), 133–142. https://doi.org/10.1016/0005-7967(92)90136-5

Baron, A. S. (2015). Constraints on the Development of Implicit Intergroup Attitudes. *Child Development Perspectives*, *9*(1), 50–54. https://doi.org/10.1111/cdep.12105

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

Baron, A. S., Dunham, Y., Banaji, M., & Carey, S. (2014). Constraints on the Acquisition of Social Category Concepts. *Journal of Cognition and Development*, *15*(2), 238–268. https://doi.org/10.1080/15248372.2012.742902

Bigler, R. S., & Liben, L. S. (1992). Cognitive Mechanisms in Children's Gender Stereotyping : Theoretical and Educational Implications of a Cognitive-Based Intervention. *Child Development*, *63*(6), 1351–1363. https://doi.org/http://www.jstor.org/stable/1131561

Brauer, M., Er-rafiy, A., Kawakami, K., & Phills, C. E. (2012). Describing a group in positive terms reduces prejudice less effectively than describing it in positive and negative terms. *Journal of Experimental Social Psychology*, *48*(3), 757–761. https://doi.org/10.1016/j.jesp.2011.11.002

Cone, J., Mann, T. C., & Ferguson, M. J. (2017). Changing Our Implicit Minds: How, When, and Why Implicit Evaluations Can Be Rapidly Revised. In *Advances in Experimental Social Psychology* (Vol. 56, pp. 131–199). https://doi.org/10.1016/bs.aesp.2017.03.001

Cvencek, D., Meltzoff, A. N., & Kapur, M. (2014). Cognitive consistency and math-gender

stereotypes in Singaporean children. *Journal of Experimental Child Psychology*, *117*(1), 73–91. https://doi.org/10.1016/j.jecp.2013.07.018

Davidson, M., Amso, D., Cruess Anderson, L., & Diamond, A. (2006). Development of cognitive control and executive fun citons from 4 to 13 years: Evidence from manipulations of memory, inhibition and task switching. *Neuropsychologia*, *44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, *8*(7), 342–353. https://doi.org/10.1111/spc3.12111

De Houwer, J. (2018). Propositional Models of Evaluative Conditioning. *Social Psychological Bulletin*, *13*(3), 1–21. https://doi.org/10.5964/spb.v13i3.28046

De Houwer, J., & Hughes, S. (2016). Evaluative Conditioning as a Symbolic Phenomenon: On the Relation between Evaluative Conditioning, Evaluative Conditioning via Instructions, and Persuasion. *Social Cognition*, *34*(5), 480–494. https://doi.org/10.1521/soco.2016.34.5.480

Dunham, Y., Baron, A. S., & Banaji, M. R. (2007). Children and social groups: A developmental analysis of implicit consistency in Hispanic Americans. *Self and Identity*, *6*(2–3), 238–255. https://doi.org/10.1080/15298860601115344

Dunham, Y., Baron, A. S., & Banaji, M. R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences*, *12*(7), 248–253. https://doi.org/10.1016/j.tics.2008.04.006

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

Dunham, Y., Newheiser, A.-K., Hoosain, L., Merrill, A., & Olson, K. R. (2014). From a Different Vantage: Intergroup Attitudes Among Children from Low- and Intermediate-Status Racial Groups. *Social Cognition*, *32*(1), 1–21. https://doi.org/10.1521/soco.2014.32.1.1

Fazio, R. H. (2007). Attitudes as Object–Evaluation Associations of Varying Strength. *Social*

*Cognition*, *25*(5), 603–637. https://doi.org/10.1521/soco.2007.25.5.603

Gonzalez, A. M., Dunlop, W. L., & Baron, A. S. (2016). Malleability of implicit associations across development. *Developmental Science*, 1–13. https://doi.org/10.1111/desc.12481

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*(1), 123–130. https://doi.org/10.1111/cdev.12582

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2018). rstanarm: Bayesian applied regression modeling via Stan. R package version 2.17.4.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, *37*(5), 620–629. https://doi.org/10.1037/0012-1649.37.5.620

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Halbeisen, G., Walther, E., & Schneider, M. (2016). Evaluative Conditioning and the Development of Attitudes in Early Childhood. *Child Development*, *00*(0), 1–8. https://doi.org/10.1111/cdev.12657

Heiphetz, L., Spelke, E. S., Harris, P. L., & Banaji, M. R. (2013). The development of reasoning about beliefs: Fact, preference, and ideology. *Journal of Experimental Social Psychology*, *49*(3), 559–565. https://doi.org/10.1016/j.jesp.2012.09.005

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, *136*(3), 390–421. https://doi.org/10.1037/a0018916

Koenig, M. A., & Harris, P. L. (2005). The role of social cognition in early trust. *Trends in Cognitive Science*, *9*(10), 457–459. https://doi.org/10.1016/j.tics.2005.08.006

Kruschke, J. K. (2015). *Doing Bayesian Data Analysis*. *Doing Bayesian Data Analysis* (2nd Edition). Elsevier. https://doi.org/10.1016/B978-0-12-405888-0.09999-2

Kurdi, B., & Banaji, M. R. (2017). Repeated Evaluative Pairings and Evaluative Statements : How Effectively Do They Shift Implicit Attitudes ? *Journal of Experimental Psychology: General*, *146*(2), 194–213. https://doi.org/10.1037/xge0000239

Kurdi, B., & Banaji, M. R. (2019). Attitude change via repeated evaluative pairings versus evaluative statements: Shared and unique features. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspa0000151

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., … Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, *143*(4), 1765–1785. https://doi.org/10.1037/a0036260

Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist*, *45*(4), 513–520. https://doi.org/10.1037/0003-066X.45.4.513

Mann, T. C., Kurdi, B., & Banaji, M. R. (*in press*). How effectively can implicit evaluations be updated? Using evaluative statements after aversive repeated evaluative pairings. *Journal of Experimental Psychology: General*.

McNulty, J. K., Olson, M. A., Jones, R. E., & Acosta, L. M. (2017). Automatic Associations Between One's Partner and One's Affect as the Proximal Mechanism of Change in Relationship Satisfaction: Evidence From Evaluative Conditioning. *Psychological Science*, *28*(8), 1031–1040. https://doi.org/10.1177/0956797617702014

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing Cognitive Control: Three Key Transitions. *Current Directions in Psychological Science*, *21*(2), 71–77.

https://doi.org/10.1177/0963721412436807

Newheiser, A.-K., & Olson, K. R. (2012). White and Black American children's implicit intergroup bias. *Journal of Experimental Social Psychology*, *48*(1), 264–270. https://doi.org/10.1016/j.jesp.2011.08.011

O'Donnell, J. M., & Brown, M. J. (1973). The classical conditioning of attitudes: A comparative study of ages 8 to 18. *Journal of Personality and Social Psychology*, *26*(3), 379–385. https://doi.org/10.1037/h0034431

Qian, M. K., Heyman, G. D., Quinn, P. C., Messi, F. A., Fu, G., & Lee, K. (2016). Implicit Racial Biases in Preschool Children and Adults From Asia and Africa. *Child Development*, *87*(1), 285–296. https://doi.org/10.1111/cdev.12442

Qian, M. K., Quinn, P. C., Heyman, G. D., Pascalis, O., Fu, G., & Lee, K. (2017). A Long-term effect of perceptual individuation training on reducing implicit racial bias in preschool children. *Child Development*, 1–16. https://doi.org/http://dx.doi.org/10.1037/dev0000290

R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rutland, A., Cameron, L., Milne, A., & Mcgeorge, P. (2005). Social Norms and Self-Presentation : Children ' s Implicit and Explicit Intergroup Attitudes. *Child Development*, *76*(2), 451–466. https://doi.org/10.1111/j.1467-8624.2005.00856.x

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*(6), 995–1008. https://doi.org/10.1037/0022-3514.91.6.995

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Santolin, C., & Saffran, J. R. (2017). Constraints on Statistical Learning Across Species. *Trends in*

*Cognitive Sciences*, *22*(1), 52–63. https://doi.org/10.1016/j.tics.2017.10.003

Schulz, L. (2012). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2012.06.004

Steele, J. R., George, M., Williams, A., & Tay, E. (2018). A cross-cultural investigation of children's implicit attitudes toward White and Black racial outgroups. *Developmental Science*, *21*(6). https://doi.org/10.1111/desc.12673

Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review*, *8*(3), 220–247.

Van Dessel, P., Hughes, S., & De Houwer, J. (2019). How Do Actions Influence Attitudes? An Inferential Account of the Impact of Action Performance on Stimulus Evaluation. *Personality and Social Psychology Review*, *23*(3), 267–284. https://doi.org/10.1177/1088868318795730

Zanon, R., De Houwer, J., Gast, A., & Smith, C. T. (2014). When does relational information influence evaluative conditioning? *Quarterly Journal of Experimental Psychology*, *67*(11), 2105–2122. https://doi.org/10.1080/17470218.2014.907324
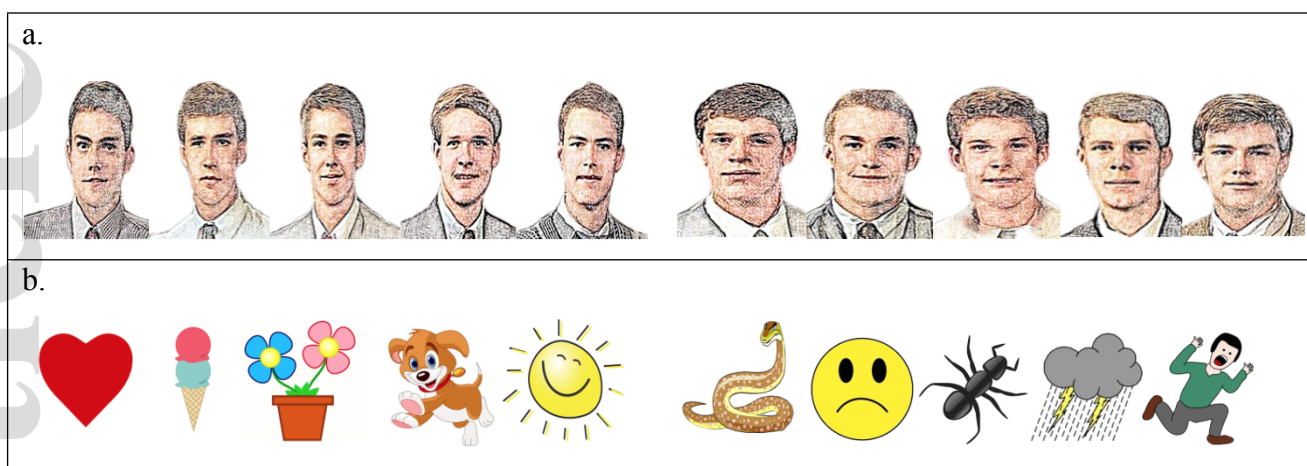
**Figure 1.** *Stimuli Presented to Participants.* a. Longfaces and Squarefaces group category stimuli; b. Positive and negative valence attribute stimuli used in all repeated evaluative pairing (REP) learning conditions.
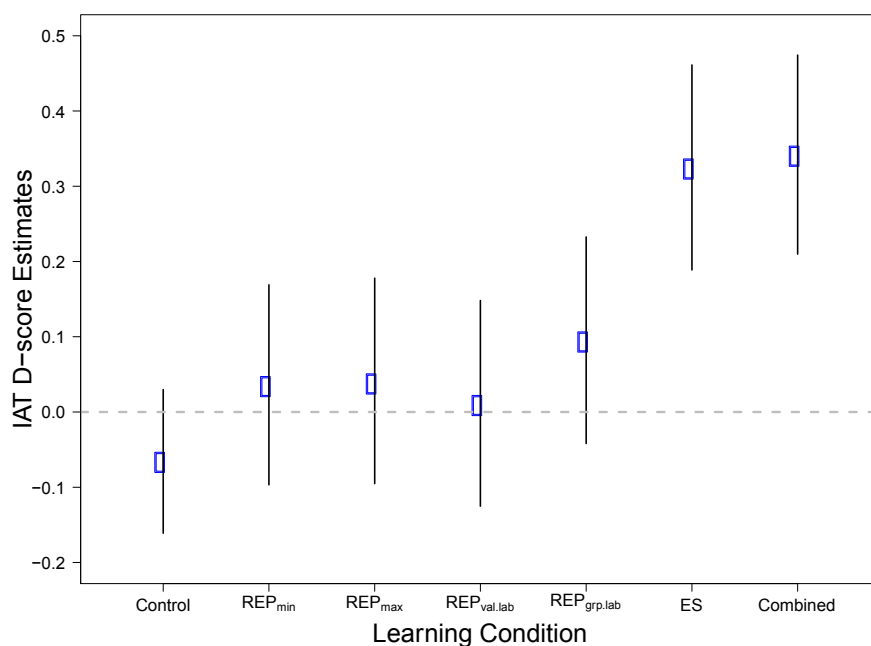


**Figure 2.** *Implicit Preferences by Learning Condition.* Error bars indicate 95% Highest Density Intervals (HDIs) from Bayesian linear models, with control condition as the dummy-coded intercept. $REP_{min}$ indicates repeated evaluative pairings with minimal instructions ($REP_{min.instruct}$); $REP_{max}$ indicates $REP_{max.instruct}$ condition with additional task information; $REP_{val.lab}$ indicates $REP_{labels}$

condition with valence US labels; REP$_{grp.lab}$ indicates REP$_{group.labels}$ condition with both group CS and valence US labels; ES indicates evaluative statements condition; and Combined indicates the combination of ES+REP.

**Table 1.**

*Comparison of the information provided to participants across learning conditions.*

| | CS–US pairings | Learning goal | Valence labels | Group labels | Relational description |
|---|---|---|---|---|---|
| Control | X | X | X | X | X |
| REP$_{min.instruct}$ | **Y** | X | X | X | X |
| REP$_{max.instruct}$ | **Y** | **Y** | X | X | X |
| REP$_{labels}$ | **Y** | **Y** | **Y** | X | X |
| REP$_{group.labels}$ | **Y** | **Y** | **Y** | **Y** | X |
| ES | X | **Y** | **Y** | **Y** | **Y** |
| ES+REP | **Y** | **Y** | **Y** | **Y** | **Y** |

*Note*: "Y" and dark shading indicates that the information was provided to participants; "X" and light shading indicates that the information was not provided to participants.

[a.] *Learning goal* refers to whether the learning goal was highlighted by explicitly instructing children to learn the relationship between groups and good/bad pictures (i.e., "Your job is to learn which group goes with which kind of picture") [b.] *Valence labels* refers to whether US (good/bad) categories were highlighted by explicitly introducing the positively valenced images as "good things" and the negatively valenced images as "bad things.". [c.] *Group labels* refers to whether CS

(Squarefaces/Longfaces) categories were highlighted by explicitly introducing the two kinds of groups as "Longfaces" and "Squarefaces." [d.] *Relational description* refers to whether the relationship among CS and US was highlighted by explicitly stating that this group is good [points to the Squarefaces], and this group is bad [points to the Longfaces]. Verbatim instruction text is provided in supplemental materials.

**Table 2.**

*Descriptives and Bayesian Regression of Implicit Preferences (IAT D-scores) by Learning Condition*

| | Descriptives | | Bayesian Posterior Estimates | | | | |
|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *95% HDI[a]* | *P $H_{01}$[b]* | *P $H_{10}$[b]* |
| *Control* | -0.07 | 0.28 | -0.06 | 0.05 | [-0.16, 0.03] | 0.96 | 0.04 |
| *REP$_{min.instruct}$* | -0.03 | 0.31 | 0.03 | 0.07 | [-0.10, 0.17] | 0.95 | 0.05 |
| *REP$_{max.instruct}$* | -0.03 | 0.30 | 0.04 | 0.07 | [-0.10, 0.17] | 0.94 | 0.06 |
| *REP$_{labels}$* | -0.05 | 0.27 | 0.01 | 0.07 | [-0.13, 0.14] | 0.97 | 0.03 |
| *REP$_{group.labels}$* | 0.03 | 0.31 | 0.09 | 0.07 | [-0.04, 0.23] | 0.76 | 0.24 |
| *ES* | 0.26 | 0.30 | 0.32 | 0.07 | [0.19, 0.46] | 0.01 | 0.99 |
| *ES+REP Combined* | 0.28 | 0.36 | 0.34 | 0.07 | [0.21, 0.47] | 0.00 | 1.00 |

[a.] Highest Density Intervals (HDIs) of the posterior indicate the range of values that include 95% of the posterior model estimates. [b.] P ($H_{01}$) indicates the proportion of the posterior density that falls inside the pre-specified null interval of [-0.15, 0.15], while P ($H_{10}$) indicates the proportion of the posterior density that falls outside the null interval.

**Table 3.**

*Descriptives and Bayesian Regression of Explicit Preference Questions by Learning Condition*

|  | P(better) | P(play) | Bayesian Posterior Estimates | | | | |
|---|---|---|---|---|---|---|---|
|  |  |  | $M^a$ | $SD^a$ | 95% HDI$^b$ | P $H_0^c$ | P $H_1^c$ |
| Control | 0.71 | 0.74 | 2.27 | 0.83 | [0.72, 3.89] | 0.02 | 0.98 |
| REP$_{min.instruct}$ | 0.64 | 0.67 | -0.73 | 0.86 | [-2.45, 0.93] | 0.40 | 0.60 |
| REP$_{max.instruct}$ | 0.64 | 0.59 | -1.05 | 0.85 | [-2.74, 0.57] | 0.28 | 0.72 |
| REP$_{labels}$ | 0.68 | 0.71 | -0.34 | 0.87 | [-2.04, 1.36] | 0.50 | 0.50 |
| REP$_{group.labels}$ | 0.62 | 0.72 | -0.61 | 0.87 | [-2.35, 1.07] | 0.44 | 0.56 |
| ES | 0.90 | 0.87 | 1.87 | 0.96 | [0.08, 3.82] | 0.09 | 0.91 |
| ES+REP Combined | 0.86 | 0.91 | 2.00 | 0.93 | [0.26, 3.92] | 0.06 | 0.94 |

*Note:* Models were Generalized Linear Mixed Effects Model with random intercepts of subject and question ("better" or "play with"). Descriptives indicate the probability of selecting Squarefaces as "better" or the group to "play with."

[a.] Mean and standard deviation of the posterior are in untransformed log-odds from the logistic regression results. [b.] Highest Density Intervals (HDIs) of the posterior indicate the range of values that include 95% of the posterior model estimates, in log-odds. [c.] P ($H_0$) indicates the proportion of the posterior density, transformed into probability estimates, that falls inside the pre-specified null interval centered around chance (0.50), [-0.35, 0.65], while P ($H_1$) indicates the proportion of the posterior density that falls outside the null interval.