



SHORT REPORT

Malleability of implicit associations across development

Antonya Marie Gonzalez,¹ William L. Dunlop² and Andrew Scott Baron¹

1. Department of Psychology, University of British Columbia, Canada

2. Department of Psychology, University of California, Riverside, USA

Abstract

The prevalence of implicit intergroup bias in adults underscores the importance of knowing when during development such biases are most amenable to change. Although research suggests that implicit intergroup bias undergoes little change across development, no studies have directly examined whether developmental differences exist in the capacity for novel implicit associations to form or change. The present study examined this issue among children ages 5–12. Results from over 800 children provided evidence that novel implicit associations formed quickly, regardless of child age, association type (evaluative or non-evaluative) or the target of the association (social or non-social). Moreover, the magnitude of these changes was comparable across conditions. Coupled with similar findings among adults, these data underscore the importance of first impressions in shaping implicit intergroup bias and provide further evidence that the acquisition of implicit associations is governed by a domain-general mechanism that may be fully in place by age 5.

Research highlights

- Although researchers have argued that implicit intergroup bias emerges by age 5 and remains stable across development (no change in magnitude), no studies have examined whether developmental differences exist in the *capacity to form or change* implicit associations.
- The present study tested over 800 children ages 5–12 to determine whether developmental differences exist in the capacity to form and change novel implicit associations.
- Novel implicit associations were quick to form, regardless of age, underscoring the importance of first impressions in the formation of intergroup bias.
- Finally, no differences in the target (social group and non-social group) and type (evaluative and non-evaluative) were observed in either the capacity to form or change novel implicit associations, suggesting common constraints governing the acquisition of implicit bias.

Introduction

Gordon Allport noted that, ‘no corner of the world is free from group scorn’. This realization led him and

many others to identify factors that contribute to intergroup conflict. One such factor is the prevalence of implicit, or unconscious, prejudice against out-groups. Substantial work with adults points to the possibility that these implicit cognitions can be modified, but the magnitude of this change is quite small (for a review, see Lai, Hoffman & Nosek, 2013; Gawronski & Bodenhausen, 2006). For example, the average effect size of successful implicit bias change in a recent meta-analysis examining the malleability of implicit biases in adults was $d = .36$ (Lai, Marini, Lehr, Cerruti, Shin *et al.*, 2014). Furthermore, even though studies with adults have provided evidence that implicit bias can be changed, it appears to be easier to form these biases than it is to change them, suggesting that first impressions are particularly important (Gregg, Seibt & Banaji, 2006). Changing implicit biases in adults also appears to be more difficult than changing explicit biases, even if these implicit biases are newly acquired (Rydell, McConnell, Strain, Claypool & Hugenberg, 2007). Thus, the limited success with changing adults’ implicit intergroup bias has led researchers to question whether more flexibility would be observed earlier in development, perhaps before such cognitions have been extensively reinforced (Greenwald & Banaji, 1995).

Studies with children across numerous social categories and methodologies suggest that implicit intergroup bias undergoes little, if any, age-related change from age 5 onward (Baron, 2015; Baron & Banaji, 2006; Baron, Schmader, Cvencek & Meltzoff, 2013; Dunham, Baron & Banaji, 2008; Dunham, Chen & Banaji, 2013; Heiphetz, Spelke & Banaji, 2013; McGlothlin & Killen, 2006; Rutland, Cameron, Milne & McGeorge, 2005). However, this reported stability does not speak directly to the possibility of developmental differences in the capacity for such biases to be formed or changed. These findings may simply demonstrate that children and adults are similarly attuned to the prevailing cultural attitudes toward and stereotypes about social groups (Olson & Fazio, 2004). Though it has been shown that young children can form implicit attitudes relatively quickly (Dunham, Baron & Carey, 2011), no studies have directly examined whether age-related differences exist in the capacity for implicit associations to be formed or changed (Baron, 2015). This work notwithstanding, a recent report on the development of implicit gender attitudes does point to the potential that implicit attitudes may change in magnitude across development (Dunham, Baron & Banaji, 2016), underscoring the need to better understand when in development such implicit associations are most amenable to change. Our study sought to address this gap in the literature by investigating the malleability of implicit associations across childhood.

There are several hypotheses that posit different developmental periods when implicit associations might be most amenable to change. On one view, implicit attitudes and stereotypes are acquired slowly, the result of accumulated experience over the lifespan (Devine, 1989; Greenwald & Banaji, 1995). This view supports the hypothesis that retraining implicit associations will become increasingly more difficult as these associations are cumulatively reinforced throughout the lifespan. A second view suggests that the ease of implicit change should coincide with development of the prefrontal cortex (Amodio, 2008; Gabrieli, 1998). For example, older children should be better equipped to shift the context of evaluation (Jones, Rothbart & Posner, 2003) and to control the activation and/or application of their associations (Amodio & Ratner, 2011; Davidson, Amso, Anderson & Diamond, 2006; Halim, Ruble & Amodio, 2011). Accordingly, this view supports the hypothesis that efforts to change implicit bias will be more successful later in development (i.e. adolescence) due to age-related improvement in the capacity to revise earlier thoughts. A third view suggests that implicit biases are particularly sensitive to early life experiences and first impressions more generally (Rudman, 2004; Rudman,

Phelan & Heppen, 2007). In this way, there are no specific predictions about age-related differences in the capacity for change due to cognitive development – what matters most is the first impression that forms, and not when it forms. Consequently, this view would predict that implicit associations are easier to form than they are to change, regardless of the age of acquisition.

In addition to investigating age-related differences in the formation and change of novel implicit associations, we also sought to explore potential differences based on association type. Doing so allows us to address whether some kinds of associations are easier (albeit faster) to learn (or change) compared with other kinds of associations. For example, past research shows that both children and adults appear to reason differently about social and non-social categories. Most notably, many social categories are frequently treated as natural kinds, and viewed as having an underlying ‘essence’ that might explain category similarities (Diesendruck & Gelman, 1999; Gelman & Hirschfeld, 1999; Rhodes & Gelman, 2009a; Rothbart & Taylor, 1992). Related work reveals that when a social group is essentialized, implicit biases about that group are stronger, and it can be more difficult to change implicit associations (Diesendruck & Menahem, 2015; Walton & Banaji, 2004). In contrast, non-social inanimate categories (e.g. artifacts) are not afforded the same internal (essentialized) structure. For this reason, revising one’s beliefs about inanimate categories does not require updating their commitments to an essence for that kind (Gelman, 1988; Gelman & Markman, 1986; Keil, 1995; Rhodes & Gelman, 2009b). Thus, it might be more difficult for children to change implicit associations about social groups as opposed to non-social artifact categories. To test this possibility, we also manipulated whether children learned implicit associations toward either a social or a non-social group.

A second dimension along which we examined the malleability of implicit associations concerned the content of the association. One perspective argues that evaluative associations (e.g. a negative association with African Americans) tap into different mental processes subserved by distinct neural systems compared to non-evaluative associations (e.g. an association between African Americans and athleticism) (Amodio & Devine, 2006). In particular, Amodio and Devine (2006) demonstrated that these two association types are uncorrelated and predict different types of behaviors. In addition, evaluative and non-evaluative judgments are known to lead to differential activation of the amygdala and prefrontal cortex, which suggests that different processes subserve reasoning about both forms of information (Cunningham, Johnson, Gatenby, Gore & Banaji, 2003).

An outstanding issue in the field of implicit intergroup cognition is whether the capacity to form and change such associations differs across these two types of representations. For example, it is possible that evaluative associations might be easier to form and more difficult to change as a result of increased attention to emotionally valenced information. In our study, we distinguish association types between evaluations of a target group (e.g. good or bad) and non-evaluative associations that co-occur with a target group (e.g. consumption of a particular food; see Baron, Dunham, Banaji & Carey, 2014, for a similar approach when examining constraints on the formation of explicit attitudes and inductive reasoning for novel social categories).

Our method is a modification of the one used by Gregg *et al.* (2006), which found that implicit associations in adults formed quickly, and were slower to change. Gregg and colleagues observed that implicit biases could be formed after brief exposure to stories describing a novel group with positive traits and another novel group with negative traits. After participants had acquired these implicit biases, they were exposed to stories suggesting that each group had changed over time, and was now associated with the opposite characteristics. Results showed that participants did not reverse their initial implicit biases after reading the counter-information, suggesting that establishing a novel association may be easier than changing it. Our study extends this work by investigating whether children's capacity to form and change novel associations undergoes developmental change across early to late childhood.

We applied a similar methodology with children ages 5–12, as this is the age range commonly used in studies on implicit bias in children (for a review see Dunham *et al.*, 2008). Specifically, we employed a 2×2 design manipulating whether participants learned either novel evaluative associations or novel non-evaluative associations, which were paired with either social or non-social categories. Because studying familiar groups presents a confound between age and prior experience, both of which may independently shape the possibility of change, we employed a Novel Groups paradigm (Baron *et al.*, 2014; Foroni & Mayr, 2005; Gregg *et al.*, 2006). Similar to Gregg and colleagues (2006), children were taught an initial association, and this association was measured implicitly to assess the strength of formation. Subsequently, an intervention designed to change the initial association was introduced, and the resulting degree of association change was measured. Our goal was to investigate (a) possible developmental differences in the capacity to form novel implicit associations, (b) possible

developmental differences in the capacity to change novel implicit associations and (c) possible effects of association and group type.

Method

Participants

The full sample consisted of 1265 participants, ages 5–12 years old (674 males and 589 females, $M_{age} = 8.26$ years, $SD = 2.03$), recruited from a community-based science center. From these participants, 306 children were excluded for failing to complete the task¹ ($N = 245$), language barriers ($N = 11$), interference by a parent or sibling ($N = 28$), developmental issues² ($N = 13$) or experimenter/computer error ($N = 9$). For each of our study's four conditions, our goal was to recruit 150–200 participants spread across our age range in order to conduct meaningful age comparisons. The resulting sample consisted of 959 participants ($M_{age} = 8.59$).³ Out of these participants, 456 participants were female. In addition, the ethnic breakdown of our sample was as follows: 63.8% ($N = 612$) identified as Caucasian, 19.1% ($N = 183$) of our sample identified as Asian, 7.9% ($N = 76$) identified as a race besides Asian or Caucasian, 7.7% ($N = 74$) identified as mixed race, and 14 participants did not report their race or ethnicity. Participants were recruited from a population with a mean household income of \$58,125 (USD). All participants were recruited from a local children's museum and tested

¹ Participants who failed to complete the task were excluded for one or more of the following three reasons: (a) they did not respond to the questions asked or did not comply with the instructions of the task, (b) they verbally expressed that they no longer wanted to continue the study, or (c) parents intervened to end the study early due to time restrictions.

² We excluded any child whose legal guardian informed us that they had any sort of documented developmental delay that might affect their performance in the task.

³ We recognize that our drop-out rates appear large (306 initial exclusions), but this is not an abnormal percent of participants to exclude, particularly from a community sample with a variety of testing constraints (e.g. Workshop on Research and Museum Partnerships, Cognitive Development Society Meeting, October 2015; Gonzalez, Steele & Baron, 2016). Most of the participants who failed to complete the current study were younger children (77% of all participants who were excluded were between the ages of 5 and 8). Higher rates of attrition are to be expected among younger children, and are consistent with past studies (e.g. Baron & Banaji, 2006). Despite these exclusions, our study has greater than 90% power to detect a medium effect size across the different groups and associations (Cohen, 1988). Further, whether or not we include the participants excluded at Time 2 ($N = 78$), IAT means are nearly identical (pre: .13 and .13, post: .01 and .02).

onsite in an area dedicated to behavioral research.⁴ A legal guardian provided informed consent for all participants. No incentive was described to participants before the study. After completing the task or after withdrawing from the study, each participant received a sticker.

Procedure

Participants were tested individually on a computer running Inquisit™ version 4.0. An experimenter was present throughout the duration of the task and read the instructions to each participant. Participants were randomly assigned to one of four conditions, within a 2 (Target of Association: Social group, Non-social group) × 2 (Content of Association: Evaluative information, Non-evaluative information) study design: Social Group Evaluative, Social Group Non-Evaluative, Non-Social Group Evaluative, and Non-Social Group Non-Evaluative. Participants were first presented with a story describing two novel groups (either social agents, called the Lups and the Nifs, or non-social inanimate objects, see Appendix A) in which one group co-occurred with evaluative behaviors (e.g. hurting someone) or non-evaluative behaviors (e.g. eating cookies). For example, in the Social Evaluative condition, children would read a vignette stating:

Lups are very nice. Let me tell you about some nice things Lups did. First, Lups listen to what their parents tell them to do. They clean their room when asked and they play nice with their brothers and sisters. Second, when Lups meet other people they are very friendly and are very helpful. For example, when they play with other people they like to share their toys and snacks. Lups are very nice.

The stories for the other conditions paralleled this example, but differed based on the type of actions as well as the individuals associated with those actions (see the Supplemental Materials for the text of the other stories). Next, participants completed an implicit measure assessing the strength of association between the target group and behaviors just learned. Subsequently, participants were introduced to a second story in which the same novel group as before was now associated with a different behavior (e.g. being mean or only eating French Fries) with an equal number of examples to that presented in the initial story. Finally, as before, participants

⁴ Unexpected frequencies in subject testing in our laboratory, which is located in a community-based science center where the number of participants can vary considerably day to day, resulted in us exceeding our stopping criteria in two of our conditions (Social Evaluative and Social Non-Evaluative).

completed an implicit measure of association strength. Such a design uniquely positions us to measure whether developmental differences exist in both the formation and change of implicit associations. In addition, we are able to assess whether this question is dependent upon the content (evaluative or non-evaluative) or upon the target of these associations (social or non-social groups).

Formation story manipulation

All participants began by hearing a story with accompanying illustrations. In the two Social Group conditions, participants were introduced to two novel social groups, the Lups and the Nifs. In the Social Group Evaluative condition, participants learned that one of these two groups engaged in either prosocial (e.g. helping others) or antisocial (e.g. pushing someone) behaviors. In the Social Group Non-Evaluative condition, participants were introduced to the same novel social groups (the Lups and the Nifs), and then observed individuals from one of those groups repeatedly exhibiting a preference for eating a particular food (e.g. either cookies or French fries).

In the Non-Social Group conditions, participants were introduced to two novel inanimate object categories (purple balls and red balls) and told that one group of balls was repeatedly (but non-causally) associated with either positive (e.g. a gust of wind blew some purple balls into a field where children were looking for toys to play with . . . leading to their delight) or negative events (e.g. a gust of wind blew a purple ball through a puddle on the street resulting in bystanders getting dirty from the splash).⁵ In the Non-Social Group Non-Evaluative condition, these novel inanimate groups of balls were described such that one group was repeatedly (but not causally or intentionally) associated with the consumption of a particular food item (e.g. a gust of wind blew some purple balls down a street past where some people were eating French fries; the balls then rolled past more people eating French fries). See the Supplemental Material available online for the full text of all story versions.

Implicit association formation

Following the presentation of the Formation Story, participants' implicit associations were measured with

⁵ Pretesting of the four versions of stories including purple and red balls ensured that children at no age viewed the balls as causal agents following any of the stories involving these objects. Pilot testing using a sorting task indicated that on over 80% of trials, children classified the purple and red balls from the story as non-agentic objects.

the Child Implicit Association Test (Child IAT; Baron & Banaji, 2006). The Child IAT assesses the strength of association between a target category (Social Group conditions: the Lups and the Nifs; Non-Social Group conditions: purple balls and red balls) and an attribute (Evaluative conditions: good and bad words; Non-Evaluative conditions: images of cookies and French fries). During the Child IAT, participants used two large buttons attached to a computer to classify pictures as belonging to one of two target categories (e.g. as Lups or Nifs or as purple balls and red balls). Next, participants used the same two buttons to classify words as either good or bad or to classify pictures as either cookies or French fries. As in previous studies (Baron & Banaji, 2006; Dunham, Baron & Banaji, 2006, 2007, 2008, 2016), for the Evaluative conditions, words were presented acoustically through headphones.

During the critical (test) blocks, participants were asked to use the same two buttons to classify stimuli from both the target and attribute categories. For half the trials ($N = 40$) the same response key was used to categorize images from one target group (e.g. Lups) and stimuli from one attribute group (e.g. good words) while the other response key was used to categorize images from the other target group (e.g. Nifs) and stimuli from the other attribute group (e.g. bad words). For the other half of trials, the pairings were switched such that now one response key was used to categorize a different group of stimuli (e.g. Nifs+good words) while the other key was used to categorize the other group of stimuli (e.g. Lups+bad words). Latency to categorize each stimulus and error rates were recorded. The established logic of this procedure is that the stronger the association between two concepts (e.g. Lups+good, Nifs+bad), the faster and more accurately participants will categorize those stimuli when they share a single response key compared with when they share separate response keys (e.g. Lups+bad, Nifs+good). The order of pairings was counterbalanced across participants. See Appendix B for an example screenshot of an IAT for each condition.

Change story manipulation

Following the first IAT, participants were presented with a second story in which contrasting information was presented. Thus, if participants first learned about the Lups engaging in antisocial behavior during the Formation Story, they now learned that Lups engaged in prosocial behavior during the Change Story. If participants previously learned that Lups consistently ate cookies, they now learned that Lups consistently ate French fries. If participants initially learned that purple balls (non-causally) covaried with negative events

involving people, they now learned that purple balls (non-causally) covaried with positive events involving people. And, if they initially learned that purple balls (non-causally) covaried with people consuming cookies, they now learned that purple balls (non-causally) covaried with people consuming French fries. The number of associative pairings between the target group and attribute were matched across each story type for each participant.

Implicit association change

Following the Change Story, participants completed the same IAT as before in order to assess the change in association strength following the second story.

Results

The following data were analyzed using the same scoring procedures outlined by Greenwald, Nosek and Banaji (2003) and previously employed with children (Baron & Banaji, 2006; Cvencek, Meltzoff & Greenwald, 2011; Cvencek, Meltzoff & Kapur, 2014; Dunham *et al.*, 2006, 2013, 2016; Olson, Key & Eaton, 2015), producing individual D-scores for each participant. Although it is beyond the scope of this paper to review the extensive literature documenting the validity and reliability of the IAT method, we recommend that interested readers begin with these papers by Greenwald and colleagues (Greenwald, Nosek & Sriram, 2006; Greenwald, Poehlman, Uhlmann & Banaji, 2009; Nosek, Greenwald & Banaji, 2005, 2007). IAT values (for the first and second IAT administered) were coded in accordance with the initial association children had been taught, such that higher D-scores corresponded with a stronger association congruent with the examples provided in the Formation Story. Thus, if the IAT scores at Time 2 were reduced compared with those observed at Time 1, then this meant there was a reduction in the strength of the initial implicit association. Coding the data this way allowed us to compare the magnitude of IAT scores across conditions and across measurement times.

IAT exclusion criteria

Our initial sample included 959 subjects between the ages of 5 and 12 years ($M = 8.59$ years, $SD = 1.94$). Following convention, participants who made errors on 25% or more of total trials or who had response latencies occurring under 300 ms on 25% or more of trials were excluded (Andrews, Hampson, Greenwald, Gordon &

Widdop, 2010; Baron & Banaji, 2006; Cao & Banaji, 2016; Cvencek *et al.*, 2011; Dunham *et al.*, 2006; Cvencek, Greenwald & Meltzoff, 2016). These exclusions were performed separately for the two administered IATs since we were testing two different effects (one concerning formation and one concerning change). To examine questions concerning formation, we used only exclusion criteria from the first IAT. To examine questions concerning change, we necessarily had to apply these exclusion criteria to both IATs.

Implicit association formation (Time 1)

Of the children who completed the Time 1 IAT, 173 were excluded based on the criteria described above, resulting in a sample of 835 participants (ranging from 154 to 257 in each condition). To first test whether participants exhibited an implicit association that was consistent with the Formation Story, we collapsed the data across all conditions and tested this value against chance. Children formed an implicit association congruent with what was presented in the Formation Story ($M = 0.13$) and the strength of this association was significant, $t(834) = 7.94$, $p < .001$, Cohen's $d = .55$, $CI_{95} = [0.10, 0.16]$. See Figure 1 for means by condition at Time 1.

In addition to testing the strength of the initial association, we wanted to investigate possible differences in IAT scores at Time 1 based on (a) Group Type (social or non-social), (b) Association Type (evaluative or non-evaluative) and (c) age of the child. We tested these potential differences using a 2×2 ANCOVA (with age entered as a covariate in order to control for possible age-related differences). There was no significant main effect of Group Type, $F(1, 830) = 0.24$, $p = .62$, $\eta_p^2 < .001$, or Association Type, $F(1, 830) = 0.01$, $p = .93$, $\eta_p^2 < .001$, indicating that there were no overall differences

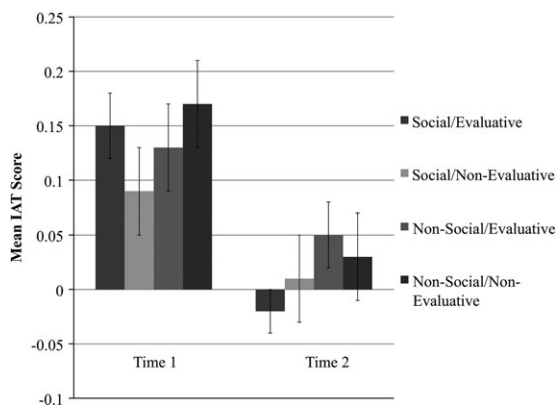


Figure 1 Mean IAT scores after formation story (Time 1) and change story (Time 2).

between Social and Non-Social Groups or Evaluative and Non-Evaluative associations.

There was a significant interaction between Group Type and Association Type, suggesting possible differences in the strength of association formation based on these two category types, $F(1, 830) = 3.73$, $p = .05$, $\eta_p^2 = .004$. Simple main effects analyses revealed a marginally significant difference between participants in the Non-Social Non-Evaluative ($M = .18$) and Social Non-Evaluative ($M = .09$) conditions ($p = .08$, $\eta_p^2 = .004$). Overall, participants across conditions formed implicit associations of similar magnitudes, but associations were weakest when social targets were paired with a non-evaluative association, and strongest when non-social targets were paired with a non-evaluative association. Given the relatively small effect size ($\eta_p^2 < .01$) and our large sample size, we believe these data support the broader conclusion that implicit associations are quick to form and do so similarly for a variety of association types.

IAT scores did not vary based on child age, as this variable was not a significant covariate in the ANCOVA, $F(1, 830) = 1.09$, $p = .30$, $\eta_p^2 = .001$. To further explore potential age-related differences, we performed a bivariate correlation between age and IAT score (collapsing across conditions), and again no relationship with age was observed, $r(833) = -.04$, $p = .27$. Thus, these data further suggest that across this wide age range, implicit associations are quick to form regardless of association type or child's age, providing perhaps the clearest evidence to date that the development of implicit associations is acquired via a domain-general learning process in place by age 5.

Implicit association change (Time 2)

In order to investigate change in implicit association strength following the Change Story, we applied exclusion criteria for both IATs (as before, more than 25% errors or with greater than 25% of response latencies occurring under 300 ms), resulting in a total of 757 participants (from the 1008 who completed both IATs) for these analyses (ranging from 137 to 242 in each condition). In order to compare IAT scores at Time 1 and Time 2, we first collapsed the data across all conditions and performed a paired samples t -test. Results showed that there was a significant difference between the Time 1 IAT score ($M = .13$) and the Time 2 IAT score ($M = .01$), indicating that children's associations were significantly reduced after hearing the change story, $t(756) = 6.00$, $p < .001$, Cohen's $d = 0.44$, $CI_{diff} = [0.08, 0.16]$.

In order to test whether change in implicit association strength differed as a function of condition or age, we

conducted a mixed factorial ANCOVA. IAT Test Time was entered as a within-subjects variable in order to look at differences between the Time 1 IAT score and the Time 2 IAT score. In order to explore potential differences by condition and age, Group Type and Association Type were once again entered as between-subjects variables and age was entered as a covariate. As before, there was no significant main effect of Group Type, $F(1, 752) = 1.02, p = .31, \eta_p^2 = .001$, or Association Type, $F(1, 752) = 0.00, p = .98, \eta_p^2 < .001$. However, a three-way interaction between IAT Test Time, Group Type and Association Type was once again observed, indicating potential differences in the degree of association change based on condition, $F(1, 752) = 3.74, p = .05, \eta_p^2 = .005$. Simple main effects analyses revealed that in the Social Evaluative ($p < .001, \eta_p^2 = .03$), Social Non-Evaluative ($p = .03, \eta_p^2 = .007$) and the Non-Social Non-Evaluative ($p = .002, \eta_p^2 = .01$) conditions, change in association strength from Time 1 to Time 2 was still significant. In contrast, in the Non-Social Evaluative condition ($p = .10, \eta_p^2 = .004$), association change was only marginally significant. These results suggest that the degree of association change differed somewhat depending on condition, but in most conditions, there was change from Time 1 to Time 2 (see Figure 1). However, as indicated by the reported effect sizes ($\eta_p^2 < .03$), these changes were quite small and should be interpreted cautiously. Although the magnitude of change differed by condition, all effects reflected a reduction in implicit association strength. These results, coupled with the analyses from both IAT test times, suggest that the Change Story was sufficient to reduce the newly formed associations to chance levels.

Once again age was not a significant covariate, $F(1, 752) = 1.21, p = .27, \eta_p^2 = .002$, meaning that the degree of association change did not differ across our age range. In addition to this analysis, we performed a bivariate correlation between age and an IAT difference score (which was calculated by subtracting the Time 2 IAT score from the Time 1 IAT score). This analysis further showed that there was no relationship between age and the degree of association change, $r(755) = -.01, p = .83$.

Order effects

To further investigate the sensitivity of these novel associations, we looked at potential order effects within each IAT. Our IATs comprised two critical blocks of trials. For example, in the Social Evaluative condition, one block of trials prompted participants to first categorize images of Lups and good words using the same response key (and images of Nifs and bad words using a second response key). In the second critical

block, participants were instructed to categorize images of Lups and bad words using the same response key (and images of Nifs and good words using the second response key). Depending on the content of the story, the first IAT block may reflect the initial association formed (at Time 1). For Time 2, the first IAT block may represent what was first taught in the initial formation story or it may reflect what was just taught in the change story. Thus, it is important to examine whether the order of IAT blocks at Time 1 and Time 2 affects participants' IAT scores.

To look at basic order effects, we first analyzed the data collapsed across conditions. At Time 1, participants who completed the compatible block of the IAT first ($M = .17$) had significantly stronger initial implicit associations than participants who completed the incompatible block first ($M = .09, t(833) = -2.42, p = .02$, Cohen's $d = 0.17, CI_{diff} = [-0.14, -.02]$). However, further analyses indicate that regardless of block order, participants formed an implicit association significantly different from chance (and congruent with the Formation Story; compatible block first: $t(410) = 7.26, p < .001$; incompatible block first: $t(423) = 4.00, p < .001$), underscoring our broader conclusion that novel implicit associations form relatively quickly. These results also indicate that completing the compatible block of the IAT first most likely reinforced the initial learned association, but regardless of block order, all participants acquired an implicit association consistent with the one they were taught at Time 1.

For the IAT at Time 2, participants who completed the compatible block first (meaning the first block was consistent with the association taught in the Change Story at Time 2, and inconsistent with the initial learned association at Time 1) exhibited significantly higher difference scores from Time 1 to Time 2 ($M = .23$) than participants who completed the incompatible block first ($M = .00; t(755) = -6.27, p < .001$, Cohen's $d = 0.46, CI_{diff} = [-0.31, -0.16]$). Thus, participants who first completed a block of trials that reinforced the association learned in the Change Story at Time 2 experienced a significant reduction in the strength of the initial learned association. In contrast, participants who first completed a block reinforcing the initial association learned during the Formation Story experienced little to no reduction in association strength between Time 1 and Time 2.

Furthermore, participants who first completed a block reinforcing the Change Story association had IAT scores that were significantly different from chance in the direction of the change story ($M = -.15, t(382) = -7.10, p < .001$, Cohen's $d = 0.73, CI_{95} = [-0.20, -0.11]$), while participants who completed a block reinforcing the initial association first learned at Time

1 had IAT scores that were significantly different from chance in the direction of the formation story ($M = .18$, $t(374) = 7.97$, $p < .001$, Cohen's $d = 0.82$, $CI_{95} = [0.14, 0.23]$). These results suggest that block order influenced whether participants' initial associations were reduced at Time 2. Specifically, following the Change Story, the particular association assessed via the first block of the IAT may have served to reinforce which of the two competing associations was more dominant. The significance of this finding for theories of implicit association change will be discussed further in the discussion section.

Discussion

This study provides the first evidence that the capacity to form and change novel implicit associations does not vary across childhood, suggesting that the mechanism(s) governing implicit associative learning may be fully intact by age 5. Consistent with previous work with adults (Gregg *et al.*, 2006; Rydell *et al.*, 2007), we observed that implicit associations form quickly, following just a few examples pairing category exemplars with events. Moreover, these data show that novel implicit associations can form quickly and be of similar magnitude regardless of whether those associations contain evaluative or non-evaluative content or whether the target of those associations are social or non-social categories. Surprisingly, it appears that children can form an implicit association between purple balls and cookies simply based on object co-occurrence as quickly and robustly as they form an association about a novel social group engaging in intentional antisocial behavior. This finding showcases the domain-general and automaticity of this system, and complements research demonstrating that both implicit and explicit attitudes can be induced in children and adults through associative learning (e.g. classical conditioning; Field, 2006; Olson & Fazio, 2001).

We also found that age did not predict the capacity to form novel implicit associations. This finding is noteworthy, as the period of development we tested is marked by substantial cognitive maturation. Consequently, these data support the view articulated by Rudman and colleagues that novel implicit associations are particularly sensitive to first impressions regardless of the age of first exposure (Rudman *et al.*, 2007). Indeed, the present results underscore how rapidly these first impressions can form across childhood, and emphasize the importance of shaping these associations at the point they are first acquired. Although this finding is conceptually similar to other reported demonstrations of developmental invariance in the magnitude of implicit bias

(Baron & Banaji, 2006; Dunham *et al.*, 2008, 2013), this study provides the first empirical test of whether developmental differences exist in the *capacity* to form novel implicit associations and begin to shed light on some emerging evidence of developmental changes in the magnitude of implicit intergroup bias (e.g. Dunham *et al.*, 2016).

Certain aspects of implicit association change were found to be similar to that of formation. For example, the capacity to change novel associations appears to be relatively similar for social and non-social groups as well as for evaluative and non-evaluative associations, once again implicating a domain-general learning mechanism. Moreover, like implicit association formation, age was not found to be a significant predictor of implicit association change, further pointing to the likelihood that the implicit associative system may be fully formed by age 5.

The nature of implicit association change was found to be decidedly more complex than its formation. For example, in contrast to our findings on implicit association formation, the block order of the second IAT (i.e. the one administered following the presentation of the Change Story) played a key role in the magnitude of implicit association change. This finding demonstrates that novel implicit associations are particularly sensitive to additional reinforcement. Specifically, after the Formation Story, all children formed an implicit association consistent with what had been presented, regardless of the IAT block order (compatible trials first or incompatible trials first). However, after the Change Story, which presented children with information that was contrary to the initial association formed, the block order of the subsequent IAT influenced which association children maintained (Time 1 or Time 2). If children were exposed to the association learned in the Formation Story in the first block of the second IAT, their implicit associations were consistent with their initial association at Time 1. In contrast, if children were exposed to the association learned in the Change Story in the first block of the second IAT, their implicit associations moved in the direction consistent with the Change Story. In summary, the association that was reinforced by the first block of the second IAT was the association children maintained after the Change Story. This finding contrasts with other studies on the malleability of novel implicit associations in adults, which suggest that novel implicit associations are difficult to change (Gregg *et al.*, 2006). Our results showed that under certain conditions, novel implicit associations can be reversed, a result that may be particular to our child sample.

One possible explanation for the above finding is that, after hearing the Formation and Change stories,

participants held both associations in mind, and the first block of the second IAT simply served to activate one of the two learned associations. If this interpretation is correct, then an IAT conducted at a later point in time might be expected to exhibit similar order effects, such that whichever association was presented in the first block would be activated in participants' minds. Another possibility is that the first block of the IAT at Time 2 served to reinforce whichever association was presented by pairing those concepts together immediately after children learned the association from the Change Story, and thus cemented this particular association in their mind. If this interpretation is correct, then an IAT conducted at a later point in time should exhibit a similar magnitude association regardless of block order. Future research will need to examine whether such associations are changing or whether the IAT might be priming certain associations over others.

One limitation of this study, as well as of the study with adults upon which the present study was modeled (Gregg *et al.*, 2006), is the lack of a control group that took both IATs without receiving information about different groups. Such a control would have allowed us to make more direct claims about the effect of taking the same IAT twice. However, research with adults has found reasonable test–retest reliability upon multiple administrations of the IAT, suggesting that after repeated testing, participant bias does not change significantly (see Lane, Banaji, Nosek & Greenwald, 2007). Thus, it is unlikely that taking the IAT twice would explain the pattern of results reported in this manuscript. However, future research on interventions to decrease implicit bias should include a control group to more precisely report on the efficacy of an intervention to change bias, especially if multiple interventions are deployed in a single study.

Thus far, studies with adult populations that have successfully reduced implicit bias have focused on short-term association formation and change (Gregg *et al.*, 2006; Lai *et al.*, 2014), while studies investigating more long-term change have been relatively unsuccessful (Lai, Skinner, Cooley, Murrar, Brauer *et al.*, 2016). Our study only assessed association strength immediately after association formation and change. Future studies should explore the effect of a larger temporal gap between learning novel associations and assessing them. Examining association strength after a longer period of time would more closely mimic the conditions under which individuals learn and change implicit associations and may help to reveal whether developmental differences exist in the capacity to sustain long-term association change. Moreover, in order to appropriately

simulate the conditions in which stereotypes are formed, future studies examining the formation and malleability of novel implicit associations should consider varying the base rates of occurrences (e.g. having Lups participate in more good actions than bad actions), to see how this might impact the magnitude of implicit bias change (see Cao & Banaji, 2016). Recent work has shown that implicit stereotypes rely on base rates of statistical occurrence even after individuals are exposed to counterstereotypical facts. Thus, future studies varying the number of incidences used to form an association might shed light on the malleability of implicit associations in relation to their degree of reinforcement.

While this study specifically examined the malleability of novel implicit associations, future work could also examine possible developmental differences in the malleability of familiar attitudes and stereotypes, such as attitudes toward categories of race, gender or religion, especially considering that implicit associations about social groups have been found to predict prejudice and discrimination in these and other domains (Greenwald *et al.*, 2009). Although the age of participants and their prior knowledge about a group will present a confound, researchers can still examine whether developmental differences exist in the amount of evidence needed to revise familiar implicit associations as well as in the duration of that change. Such work would also more directly test the hypothesis that implicit associations are more difficult to change among older individuals because they have to overcome more accumulated experiences reinforcing the initial association (Greenwald & Banaji, 1995).

Taken together, the results of this study suggest that implicit associations form rapidly, regardless of the age of the learner (at least between ages 5–12). The formation of these associations appears to be supported by a domain-general mechanism that remains stable across development, a finding that emphasizes our proclivity to detect co-occurring events and to form congruent associations regardless of their content or the causal relationship between those co-occurring events. The implications of these findings are that we have a tendency to form implicit associations quickly, even following minimal exposure to those co-occurring events.

The results from this study underscore the likelihood that seemingly minimal pairings of certain social groups with stereotypical traits or evaluative behaviors will lead to rapid formation of implicit associations. Therefore, efforts to reduce negative stereotypes and attitudes toward stigmatized social groups should necessarily begin to focus on instilling individuals with positive

associations at a young age, as well as attending to the reinforcement of these early forming associations.

Acknowledgements

This research was supported by a Social Sciences and Humanities Research Council of Canada grant to ASB (# 435-2013-0286). We would also like to thank the Living Lab at Science World at TELUS World of Science in Vancouver and the participating parents and children.

References

- Amodio, D.M. (2008). The social neuroscience of intergroup relations. *European Review of Social Psychology*, **19**, 1–54. doi:10.1080/10463280801927937
- Amodio, D.M., & Devine, P.G. (2006). Stereotyping and evaluation in implicit race bias: evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, **91**, 652–661. doi:10.1037/0022-3514.91.4.652
- Amodio, D.M., & Ratner, K.G. (2011). A memory systems model of implicit social cognition. *Current Directions in Psychological Science*, **20**, 143–148. doi:10.1177/0963721411408562
- Andrews, J.A., Hampson, S.E., Greenwald, A.G., Gordon, J., & Widdop, C. (2010). Using the implicit association test to assess children's implicit attitudes toward smoking. *Journal of Applied Social Psychology*, **40** (9), 2387–2406. doi:10.1111/j.1559-1816.2010.00663.x
- Baron, A.S. (2015). Constraints on the development of implicit intergroup attitudes. *Child Development Perspectives*, **9**, 50–54. doi: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**, 53–58. doi:10.1111/j.1467-9280.2005.01664.x
- Baron, A.S., Dunham, Y., Banaji, M.R., & Carey, S. (2014). Constraints on the acquisition of social category concepts. *Journal of Cognition and Development*, **15**, 238–268. doi:10.1080/15248372.2012.742902
- Baron, A.S., Schmader, T., Cvencek, D., & Meltzoff, A.N. (2013). The gendered self-concept: how implicit gender stereotypes and attitudes shape self-definition. In P.J. Lehman & H.R. Tenenbaum (Eds.), *Gender and development* (pp. 109–132). Hove, East Sussex: Psychology Press.
- Cao, J., & Banaji, M.R. (2016). The base rate principle and the fairness principle in social judgment. *Proceedings of the National Academy of Sciences, USA*, **113**, 7475–7480. doi:10.1073/pnas.1524268113
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edn.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cunningham, W.A., Johnson, M.K., Gatenby, J.C., Gore, J.C., & Banaji, M.R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology*, **85**, 639–649. doi:10.1037/0022-3514.85.4.639
- Cvencek, D., Greenwald, A.G., & Meltzoff, A.N. (2016). Implicit measures for preschool children confirm self-esteem's role in maintaining a balanced identity. *Journal of Experimental Social Psychology*, **62**, 50–57. doi:10.1016/j.jesp.2015.09.015
- Cvencek, D., Meltzoff, A.N., & Greenwald, A.G. (2011). Math-gender stereotypes in elementary school children. *Child Development*, **82**, 766–779. doi:10.1111/j.1467-8624.2010.01529.x
- Cvencek, D., Meltzoff, A.N., & Kapur, M. (2014). Cognitive consistency and math-gender stereotypes in Singaporean children. *Journal of Experimental Child Psychology*, **117**, 73–91. doi:10.1016/j.jecp.2013.07.018
- Davidson, M.C., Amso, D., Anderson, L.C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, **44**, 2037–2078. doi:10.1016/j.neuropsychologia.2006.02.006
- Devine, P.G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology*, **56**, 5–18. doi:10.1037/0022-3514.56.1.5
- Diesendruck, G., & Gelman, S.A. (1999). Domain differences in absolute judgments of category membership: evidence for an essentialist account of categorization. *Psychonomic Bulletin & Review*, **6**, 338–346. doi:10.3758/BF03212339
- Diesendruck, G., & Menahem, R. (2015). Essentialism promotes children's inter-ethnic bias. *Frontiers in Psychology*, **6**, 1180. doi:10.3389/fpsyg.2015.01180
- Dunham, Y., Baron, A.S., & Banaji, M.R. (2006). From American city to Japanese village: a cross-cultural investigation of implicit race attitudes. *Child Development*, **77**, 1268–1281. doi:10.1111/j.1467-8624.2006.00933.x
- 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**, 238–255. doi:10.1080/15298860601115344
- Dunham, Y., Baron, A.S., & Banaji, M.R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences*, **12**, 248–253. doi:10.1016/j.tics.2008.04.006
- Dunham, Y., Baron, A.S., & Banaji, M.R. (2016). The development of implicit gender attitudes. *Developmental Science*, **19** (5), 781–789. doi:10.1111/desc.12321
- Dunham, Y., Baron, A.S., & Carey, S. (2011). Consequences of 'minimal' group affiliations in children. *Child Development*, **82**, 793–811. doi:10.1111/j.1467-8624.2011.01577.x
- Dunham, Y., Chen, E., & Banaji, M.R. (2013). Two signatures of implicit intergroup bias: developmental invariance and early enculturation. *Psychological Science*, **24**, 860–868. doi:10.1177/0956797612463081
- Field, A.P. (2006). I don't like it because it eats sprouts: conditioning preferences in children. *Behaviour Research and Therapy*, **44**, 439–455. doi:10.1016/j.brat.2005.03.006

- Feroni, F., & Mayr, U. (2005). The power of a story: new, automatic, associations from a single reading of a short scenario. *Psychonomic Bulletin & Review*, **12**, 139–144. doi:10.3758/BF03196359
- Gabrieli, J.D. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology*, **49**, 87–115. doi:10.1146/annurev.psych.49.1.87
- Gawronski, B., & Bodenhausen, G.V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, **132**, 692–731. doi:10.1037/0033-2909.132.5.692
- Gelman, S.A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology*, **20**, 65–96. doi:10.1016/0010-0285(88)90025-4
- Gelman, S.A., & Hirschfeld, L.A. (1999). How biological is essentialism. In D.L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 403–446). Cambridge, MA: The MIT Press.
- Gelman, S.A., & Markman, E.M. (1986). Categories and induction in young children. *Cognition*, **23**, 183–209. doi:10.1016/0010-0277(86)90034-X
- Gonzalez, A.M., Steele, J.R., & Baron, A.S. (2016). Reducing children's implicit racial bias through exposure to positive out-group exemplars. *Child Development*. doi:10.1111/cdev.12582
- Greenwald, A.G., & Banaji, M.R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, **102**, 4–27. doi:10.1037/0033-295X.102.1.4
- 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**, 197–216. doi:10.1037/0022-3514.85.2.197
- Greenwald, A.G., Nosek, B.A., & Sriram, N. (2006). Consequential validity of the Implicit Association Test: comment on the article by Blanton and Jaccard. *American Psychologist*, **61**, 56–61. doi:10.1037/0003-066X.61.1.56
- Greenwald, A.G., Poehlman, T.A., Uhlmann, E.L., & Banaji, M.R. (2009). Understanding and using the Implicit Association Test: III. *Meta-analysis of predictive validity*. *Journal of Personality and Social Psychology*, **97**, 17–41. doi:10.1037/a0015575
- Gregg, A.P., Seibt, B., & Banaji, M.R. (2006). Easier done than undone: asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, **90**, 1–20. doi:10.1037/0022-3514.90.1.1
- Halim, M., Ruble, D.N., & Amodio, D.M. (2011). From pink frilly dresses to 'one of the boys': a social-cognitive analysis of gender identity development and gender bias. *Social and Personality Psychology Compass*, **5**, 933–949. doi:10.1111/j.1751-9004.2011.00399.x
- Heiphetz, L., Spelke, E.S., & Banaji, M.R. (2013). Patterns of implicit and explicit attitudes in children and adults: tests in the domain of religion. *Journal of Experimental Psychology: General*, **142**, 864–879. doi:10.1037/a0029714
- Jones, L.B., Rothbart, M.K., & Posner, M.I. (2003). Development of executive attention in preschool children. *Developmental Science*, **6**, 498–504. doi:10.1111/1467-7687.00307
- Keil, F.C. (1995). The growth of causal understandings of natural kinds. In D. Sperber, D. Premack & A.J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 234–267). Cambridge, MA: The MIT Press.
- Lai, C.K., Hoffman, K.M., & Nosek, B.A. (2013). Reducing implicit prejudice. *Social and Personality Psychology Compass*, **7**, 315–330. doi:10.1111/spc3.12023
- Lai, C.K., Marini, M., Lehr, S.A., Cerruti, C., Shin, J.L., et al. (2014). Reducing implicit racial preferences: I A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, **143**, 1765–1785.
- Lai, C.K., Skinner, A.L., Cooley, E., Murrar, S., Brauer, M., et al. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, **145**, 1001–1016.
- Lane, K.A., Banaji, M.R., Nosek, B.A., & Greenwald, A.G. (2007). Understanding and using the implicit association test: IV. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes*, (pp. 59–102). New York: The Guilford Press.
- McGlothlin, H., & Killen, M. (2006). Intergroup attitudes of European American children attending ethnically homogeneous schools. *Child Development*, **77**, 1375–1386. doi:10.1111/j.1467-8624.2006.00941.x
- Nosek, B.A., Greenwald, A.G., & Banaji, M.R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, **31** (2), 166–180. doi:10.1177/0146167204271418
- Nosek, B.A., Greenwald, A.G., & Banaji, M.R. (2007). The Implicit Association Test at age 7: a methodological and conceptual review. In J.A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265–292). New York: Psychology Press.
- Olson, M.A., & Fazio, R.H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, **12**, 413–417. doi:10.1111/1467-9280.00376
- Olson, M.A., & Fazio, R.H. (2004). Trait inferences as a function of automatically activated racial attitudes and motivation to control prejudiced reactions. *Basic and Applied Social Psychology*, **26**, 1–11. doi:10.1207/s15324834basps2601_1
- Olson, K.R., Key, A.C., & Eaton, N.R. (2015). Gender cognition in transgender children. *Psychological Science*, **26** (4), 467–474. doi:10.1177/0956797614568156
- Rhodes, M., & Gelman, S.A. (2009a). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, **59**, 244–274. doi:10.1016/j.cogpsych.2009.05.001
- Rhodes, M., & Gelman, S.A. (2009b). Five-year-olds' beliefs about the discreteness of category boundaries for animals and artifacts. *Psychonomic Bulletin & Review*, **16**, 920–924. doi:10.3758/PBR.16.5.920
- Rothbart, M., & Taylor, M. (1992). Category labels and social reality: do we view social categories as natural kinds?. In G.R. Semin & K. Fiedler (Eds.), *Language, interaction, and social cognition* (pp. 11–36). Thousand Oaks, CA: Sage Publications.

- Rudman, L.A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science*, **13**, 79–82. doi:10.1111/j.0963-7214.2004.00279.x
- Rudman, L.A., Phelan, J.E., & Heppen, J. (2007). Developmental sources of implicit attitudes. *Personality and Social Psychology Bulletin*, **33**, 1700–1713. doi:10.1177/0146167207307487
- Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social norms and self-presentation: children's implicit and explicit intergroup attitudes. *Child Development*, **76**, 451–466. doi:10.1111/j.1467-8624.2005.00856.x
- Rydell, R.J., McConnell, A.R., Strain, L.M., Claypool, H.M., & Hugenberg, K. (2007). Implicit and explicit attitudes respond differently to increasing amounts of counterattitudinal information. *European Journal of Social Psychology*, **37**, 867–878. doi:10.1002/ejsp.393

- Walton, G.M., & Banaji, M.R. (2004). Being what you say: the effect of essentialist linguistic labels on preferences. *Social Cognition*, **22**, 193–213. doi:10.1521/soco.22.2.193.35463
- Workshop on Research and Museum Partnerships (2015). 9th Biennial Meeting of the Cognitive Development Society, Columbus, OH.

Received: 17 September 2014

Accepted: 7 July 2016

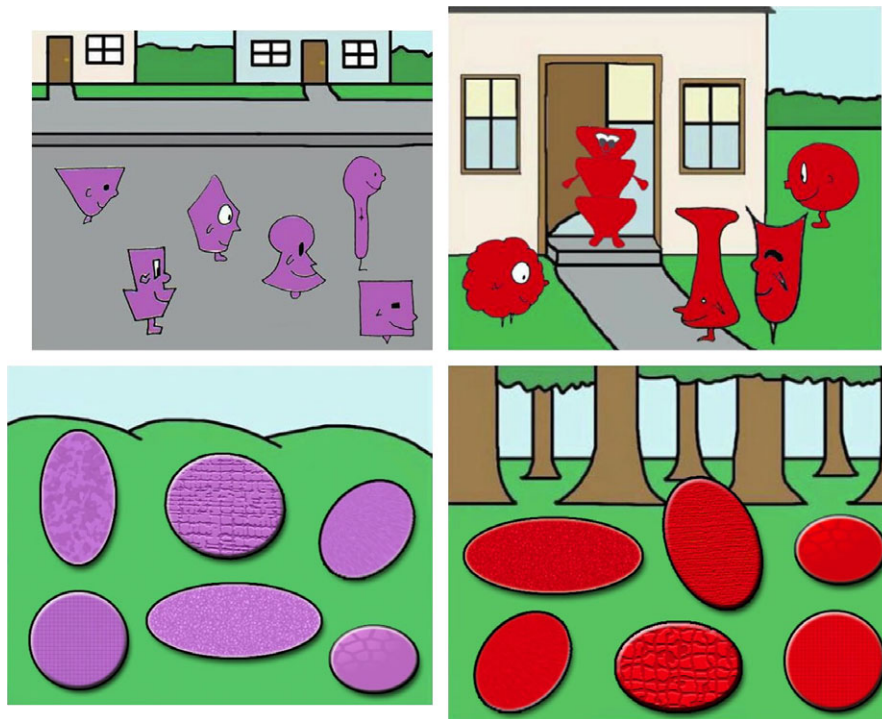
Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Data S1. Story Text

Appendix A

Example images of the Lups, Nifs, purple balls and red balls



Appendix B

Example screenshots for each of the four conditions (from top left to bottom right: Social Evaluative, Social Stereotype, Non-Social Attitude, Non-Social Stereotype)

