

Children's Spontaneous Associations With Targets Who Differ by Race and Emotional Expression

Corey Lipman¹, Amanda Williams², Kerry Kawakami¹, and Jennifer R. Steele¹

¹ Department of Psychology, York University

² School of Education, University of Bristol

Across three studies, we examined non-Black children's spontaneous associations with targets who differed by both race and emotional expression. Children aged 5 to 10 years ($N = 419$; 215 girls; 58% White; 65% of household incomes $> \$75,000/\text{year}$) completed Implicit Association Tests (IAT; Greenwald et al., 2003) containing smiling Black and neutral White target faces. In all three studies, when children categorized these faces by emotional expression, they showed more positive associations with smiling Black targets over neutral White targets, as compared with when they categorized by race. This was the case when children were shown how to categorize these faces (Studies 1 and 2) and when they spontaneously categorized by race or emotional expression on an ambiguous-categorization IAT that allowed for categorization by race and/or emotion (Studies 2 and 3). In Study 3, after watching an adult explain that she was categorizing racially diverse faces by emotional expression in a seemingly unrelated card-sorting task, children were also relatively faster to pair smiling Black faces with pleasant images and neutral White faces with unpleasant images in this ambiguous-categorization IAT compared with children in a control condition. Older children were more likely to spontaneously categorize primarily by race (Studies 2 and 3) but were also more likely to categorize by emotion following the intervention (Study 3) compared with younger children. Together, these studies provide insight into children's social categorization processes and spontaneous associations with targets who differ systematically across multiple perceptually salient categories.

Keywords: categorization, emotion, implicit, intersectionality, race

Supplemental materials: <https://doi.org/10.1037/dev0001199.supp>

Researchers have been increasingly making use of implicit measures to examine the development of children's race-based associations. A number of these studies have specifically focused on non-Black children's associations with the members of White and Black racial groups. For example, in an initial study with children, Rutland and colleagues (2005) found that White British children (aged 6 to 16 years) were faster to pair racially prototypical White faces with pleasant stimuli and Black faces with unpleasant stimuli relative to the reverse pairing. This pattern of associations did not differ by age and was not affected by the salience of antiprejudice social norms. Baron and Banaji (2006) found a similar pattern of bias

among 6-year-old White American children that was equal in magnitude to that of 10-year-olds and adults. Since these initial studies, child-friendly versions of the most popular implicit measure, the Implicit Association Test (IAT; Greenwald et al., 2003), have been used to assess spontaneous associations with racial groups among diverse school-age children in a variety of locations around the world. The findings from these studies suggest that when categorizing targets by race, non-Black children show relatively more positive associations with White over Black target faces from at least 6 years of age (Dunham et al., 2006, 2008; Gonzalez et al., 2017; Newheiser & Olson, 2012; Steele, George, Williams, & Tay, 2018; Williams & Steele, 2019). Given the prevalence of these biases, there are both theoretical and practical benefits to understanding whether these associations can be reduced or even reversed (Gonzalez et al., 2017, 2021; Kawakami et al., 2020).¹ Research conducted with children to date has focused almost exclusively on children's

Jennifer R. Steele  <https://orcid.org/0000-0003-2275-0446>

This research was supported by grants from the Social Sciences and Humanities Research Council of Canada (SSHRC Insight Grant 435-2014-1247) and the Canada Foundation for Innovation (CFI-LOF Grant 12800) awarded to the senior author, Jennifer R. Steele, as well as an Ontario Graduate Scholarship awarded to Corey Lipman. We thank members of the Interpersonal Perception and Social Cognition (IPSC) Laboratory for their help at each stage of the research process. We also extend our appreciation to the Ontario Science Center, the children who participated in these studies, and their parents for affording us the opportunity to conduct this research.

Correspondence concerning this article should be addressed to Jennifer R. Steele, Department of Psychology, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada. Email: steeleje@yorku.ca

¹ There is ongoing discussion in the field about *definitional disparities* and potential *conceptual ambiguities* (Corneille & Hütter, 2020) that exist around the use of the term implicit. We have chosen to use this term because of its familiarity and ongoing use in the field and acknowledge that in the present article we make use of a procedural definition. Specifically, we use the term *implicit measure* to describe the indirect nature of our measures that are, consistent with previous definitions, not a "direct, deliberate, controlled, and intentional self-assessment" and that assess "mental content without requiring awareness of the relation between the response and the measured content." (Nosek et al., 2011, p. 153; see also Greenwald & Banaji, 2017).

spontaneous affective biases toward a single social category (cf. Perszyk et al., 2019), typically race. However, the people who children encounter in daily life are multifaceted, each belonging to multiple social and perceptual categories. In the current research we examined how social categorization processes impact children's associations with White and Black targets who differed not only by race, but also by emotional expression. Specifically, across three studies we examined children's spontaneous association with smiling Black and neutral White targets using child-friendly versions of the Implicit Association Test (Child IAT; Baron & Banaji, 2006). Using these targets allowed us to test the hypothesis that dynamic perceptual information, specifically emotional expression, can act as a competing social cue and influence children's spontaneous affective responses toward targets who simultaneously differ by the stable social category of race.

In addition, because very little is known about children's categorization of targets who differ in multiple ways, we were particularly interested in the extent to which children *spontaneously* categorize by race, as opposed to emotional expression (Studies 2 and 3), and the relation between categorization and children's associations when they are presented with smiling Black and neutral White target faces. Finally, in Study 3, we examined whether seeing an adult categorize others by emotional expression would change how children subsequently categorized smiling Black and neutral White targets on the Implicit Association Test and their ensuing affective reactions to these faces. The overarching goal of this research was to increase our understanding of implicit social cognition in childhood and the role that social categorization plays in informing children's spontaneous affective responses toward others who can be categorized in multiple ways.

Social Categorization and Spontaneous Associations

When encountering new people, we are faced with a multitude of social categories and attributes, sometimes with competing associations. Theory and research with adults suggest that how we categorize others has important consequences for the stereotypes and attitudes that are activated, and ultimately the impressions that we form (Barden et al., 2004; Fiske et al., 1999; Gawronski & Bodenhausen, 2006; Macrae et al., 1995; Mitchell et al., 2003; Rhodes & Baron, 2019; Sinclair & Kunda, 1999; Williams & Steele, 2019). Research also provides evidence that both category labels and category exemplars can impact adults' responses on implicit measures (Bluemke & Friese, 2006; Govan & Williams, 2004; Steffens & Plewe, 2001; cf. De Houwer, 2001). For example, Mitchell et al. (2003) examined adults' associations with familiar liked Black athletes and disliked White politicians using an Implicit Association Test. When participants were asked to categorize these targets by race, they were relatively faster to pair the disliked White politicians with positive (vs. negative) stimuli as compared with when participants were asked to categorize these same familiar targets by occupation. By changing the basis of categorization, participants were able to reconstrue targets from members of racial groups to liked and disliked members of a profession, leading to a different affective response to the same targets.

Although few studies have examined children's attitudes toward targets who can be categorized in multiple perceptually salient ways, there is mounting evidence to suggest that children's spontaneous associations as measured by child-friendly implicit measures, can be malleable and context-dependent (Gonzalez et al.,

2017; Neto et al., 2016; Qian et al., 2019; Vezzali et al., 2012; Xiao et al., 2015). For example, Gonzalez et al. (2017) found that the pro-White (vs. Black) biases of older school-age children, as measured by a child-friendly Implicit Association Test, were reduced following exposure to positive Black exemplars (see also Gonzalez et al., 2021). In addition, studies that have examined children's self-reported attitudes toward targets with multiple salient features suggest that children's expressed preferences are sometimes guided by social categories other than race. For example, Kinzler et al. (2009) found that children expressed a preference for playing with other-race children who spoke with their native accent over same-race children with a foreign accent. Research by Shutts and colleagues further suggests that gender is often prioritized over race when young children aged three- to six-years spontaneously categorize others, express peer preferences, or use cues to guide their preferences for objects or activities (Shutts, 2015; Shutts et al., 2010, 2013; Weisman et al., 2015).

Emotional Expression and Racial Attitudes

In the present research we extended previous findings by examining the impact of social categorization processes on children's associations with targets who differed by both race and emotional expression. We chose to examine the impact of emotional expression on children's spontaneous associations for several reasons. First, emotional expression can provide important information about an individual's willingness to engage in friendly interactions (Keltner & Haidt, 1999). Research by Schultz et al. (2008) suggests that even children as young as three- to five-years recognize the value of attending to emotional expression when selecting children to play with; children in their study were more likely to select smiling children over children with neutral expressions as potential playmates. Second, there is some evidence that emotional expression can be used as a basis for social categorization in a way that is similar to other more stable social categories (Cañadas et al., 2016; Freeman & Ambady, 2011). As such, we felt that emotional expression would provide important social information to which children would readily attend. In addition, there is evidence that both children and adults show positive associations with smiling faces on implicit measures (Donges et al., 2012; Rohr et al., 2012; Williams et al., 2016). However, limited research has examined the impact of emotional expression on children's attitudes toward members of different racial groups.

Most studies that have examined the relationship between emotional expression and race have investigated their combined effect on categorization processes. For instance, research by Hugenberg and Bodenhausen suggests that White adults who show more racially biased associations on a race-attitude Implicit Association Test are more likely to categorize emotionally ambiguous Black compared with White faces as angry (Hugenberg & Bodenhausen, 2003) and to categorize racially ambiguous angry compared with happy faces as Black (Hugenberg & Bodenhausen, 2004; see also Freeman & Ambady, 2011). In related studies, White American, Taiwanese, and Chinese children were more likely to categorize racially ambiguous angry compared with smiling faces as racial outgroup members (Dunham et al., 2013; Xiao et al., 2015). These findings are consistent with recent theorizing suggesting a dynamic interactive model of person construal in which top-down social-cognitive processes interact with bottom-up perceptual

processes to influence categorization (Freeman & Ambady, 2011; Kawakami et al., 2017; see also Quinn et al., 2020 for research with infants). As such, it would seem that both social and perceptual categories such as race and emotional expression can influence social categorization processes and also spontaneously activated associations.

In support of this theorizing, recent evidence indicates that adults' spontaneous affective responses toward targets who differ both by race and emotional expression depend on categorization processes. In one study, Steele, George, Cease, et al. (2018) instructed non-Black adults to categorize smiling Black and neutral White targets by either race or emotional expression. When categorizing these targets by race, the speed of participants' associations was no different in magnitude to a condition where racially prototypical targets were all smiling. This finding suggests that, once categorized by race, the emotional expression had no impact on adults' associations. By contrast, when adult participants were instructed to categorize these faces by emotional expression, they showed the reverse pattern. Specifically, they were relatively faster to associate smiling Black faces with pleasant images and neutral White faces with unpleasant images. The magnitude of this effect was no different from the associations of participants who saw smiling White and neutral White targets, suggesting that once categorized by emotional expression, race had no impact on participants' associations.

In a subsequent study, Steele, George, Cease, et al. (2018) examined participants' spontaneous categorization of these faces using a modified Implicit Association Test. In this task participants were not given instructions to categorize faces by race or emotional expression but instead were asked to match images to a pictorial header that included both race and emotion information (i.e., a smiling Black face and a neutral White face). Their results indicated that while participants were more likely to categorize faces by race than emotional expression, those who spontaneously categorized by emotional expression were faster at pairing pleasant versus unpleasant stimuli with the smiling Black versus neutral White faces. Together, these findings suggest that associations with targets can be informed both by category labels (i.e., asking a participant to categorize smiling Black and neutral White faces by race or emotional expression) as well as spontaneous categorization. However, to our knowledge, no research has directly investigated the impact of emotional expression on children's spontaneous evaluative associations, with a focus on whether race-based associations will be activated when smiling Black and neutral White faces are categorized by emotional expression as opposed to race.

In addition, it is unclear whether differences would emerge in younger and older children's spontaneous categorization by race as opposed to emotional expression when categorizing smiling Black and neutral White targets on the child-friendly Implicit Association Test, and the direction of possible age-related differences. On the one hand, theory and research suggest that explicit racial bias is significantly greater in younger children relative to older children (Aboud, 2008; Rutland et al., 2005; Williams & Steele, 2019; see Raabe & Beelmann, 2011, for a meta-analysis), which could lead us to expect that younger children would spontaneously categorize these faces by race more often than older children. On the other hand, some research suggests that with increasing age the tendency to spontaneously attend to race during person perception increases, potentially because older children

have had more time to develop race-based associations (Degner & Wentura, 2010; Kawakami et al., 2020). Furthermore, the decrease in explicit racial bias among older children is at least partially attributable to learning antiprejudice social norms (Monteiro et al., 2009; Rutland et al., 2005) rather than a decreased tendency to spontaneously categorize by race. Therefore, it is also possible that older children would be equally or more likely to spontaneously categorize by race (over emotional expression) relative to younger children, even if they are less likely to use race-based information when explicitly expressing preferences.

The Present Research

The first goal of the present research was to examine the impact of social categorization on children's affective associations with targets who could be categorized in multiple ways. To accomplish this goal, children completed a child-friendly Implicit Association Test (Child IAT; Baron & Banaji, 2006) designed to measure their affective responses toward targets who differed simultaneously by race and emotional expression. Specifically, in each study, children were presented with smiling Black targets and neutral White targets matched for age and attractiveness. In Study 1, children were randomly assigned to either categorize these faces by race or emotional expression. Building on previous research using a similar paradigm with adults (Steele, George, Cease, et al., 2018), we hypothesized that children's associations with these targets would depend on the target category. Specifically, we predicted that children would be faster to associate neutral White targets with pleasant stimuli and smiling Black target faces with unpleasant stimuli when they were categorized by race (Baron & Banaji, 2006; Dunham et al., 2008; Newheiser & Olson, 2012; Rutland et al., 2005; Williams & Steele, 2019) relative to when they were categorized by emotional expression.

A second goal of the current research was to extend previous findings by examining children's *spontaneous* categorization of these targets and whether differences exist between younger (5–7 years) and older (8–10 years) children's tendency to spontaneously categorize these targets by race. As noted previously, given that younger children tend to express greater racial prejudice than older children (Raabe & Beelmann, 2011), one possibility was that younger children would be more likely to spontaneously categorize targets by race. By contrast, given that older children have had more time for race-related cognitions to be reinforced and for race to become a psychologically salient dimension for spontaneous categorization (Bigler & Liben, 2007), and that older children are more likely to see race, but not emotional expressions, as stable across time relative to younger children (Roberts & Gelman, 2016), another possibility was that older children would be more likely to categorize targets by race than younger children. We tested these possibilities in Studies 2 and 3.

Finally, in Study 3 we examined whether having an adult attend to emotional expression and not race would affect younger and older children's subsequent social categorization. Specifically, we examined whether an adult's subtle or blatant use of emotional expression as a basis for categorization in a seemingly unrelated card-sorting task would subsequently impact children's spontaneous social categorization of smiling Black and neutral White target faces on the IAT as well as children's associations with these targets.

Study 1

In Study 1 we examined whether children's spontaneous associations with smiling Black and neutral White faces differed depending on whether children were asked to categorize these faces by race or by emotional expression. We expected that when asked to categorize by race, children (aged 5 to 10 years) would be relatively faster to pair neutral White faces with pleasant images and smiling Black faces with unpleasant images than when asked to categorize these faces by emotional expression.

Method

Participants

One hundred and twenty-nine non-Black children between the ages of 5 and 10 years (74 female and 55 male) were recruited from the local science museum to take part in the study.² Of the parents who provided their annual household income ($n = 89$), the majority (76%; 68 of 89) reported earning at least \$75,000 per year. The data from eleven children were excluded prior to any analyses because of a technical error resulting in no data ($n = 1$), they chose not to complete the IAT ($n = 2$), or they had a response error rate of 35% or greater ($n = 8$; Cvencek et al., 2011), leading to a final sample of 118 participants. This included 53 younger children who ranged in age from 5 to 7 years ($M_{\text{age}} = 6$ years; 11 five-year-olds, 24 six-year-olds, 18 seven-year-olds) and 65 older children who ranged in age from 8 to 10 years ($M_{\text{age}} = 9$ years; 22 eight-year-olds, 24 nine-year-olds, and 19 ten-year-olds). The a priori decision to group participants into younger (5 to 7 years) and older (8 to 10 years) age groups was based both on social-cognitive developmental theories of prejudice development (e.g., Aboud, 2008) positing key developmental changes around 8 years of age and because other intergroup researchers have used similar age groupings (Baron & Banaji, 2006; Dunham et al., 2006; Gonzalez et al., 2017; Raabe & Beelmann, 2011; Williams & Steele, 2019). The racial makeup of the final sample consisted of 83 White, 20 East Asian, 10 South Asian, and 5 multiracial participants (who identified with more than one race that did not include Black).

Prior to each study, written parental consent for participation was obtained along with demographic information; children were also informed that they could stop the experiment at any time. In each study we aimed to recruit as many participants as possible within a set testing period, with the goal of recruiting at least 40 children per experimental condition. Each study received approval from the York University Research Ethics Board ("Understanding Intergroup Attitudes Across the Life Span" certificates #2010-302, #e2014-226, #e2020-200). A sensitivity power analysis using G*Power (Version 3.1.9.4) indicates that this sample is sufficient to detect effects as small as $\eta^2 = .06$ ($f = .26$) at .80 power.

Measures and Procedure

After providing verbal assent to participate, each child was brought to a quiet area of the museum and was seated in front of a laptop computer in one of four individual testing cubicles. Parents also provided written consent and completed a brief demographics questionnaire in a waiting area that allowed them to see their child in the testing area from a distance. Each child was randomly

assigned to complete one of two child-friendly Implicit Association Tests (Child IAT; Baron & Banaji, 2006) that was presented to children as a matching game. The IAT is a computer-based reaction time (RT) measure designed to test the strength of associations with specific categories. The Child IAT used in this study was comparable to the adult IAT (Greenwald et al., 2003) with a few minor modifications designed to make the task child-friendly (Steele, George, Williams, & Tay, 2018; Williams & Steele, 2019). Unlike most adult IATs, the stimuli used in the Child IAT were all pictorial and the number of trials in each block was reduced as outlined below. In addition, a non-Black experimenter sat to the side and slightly behind the child throughout the task and provided scripted verbal instructions.

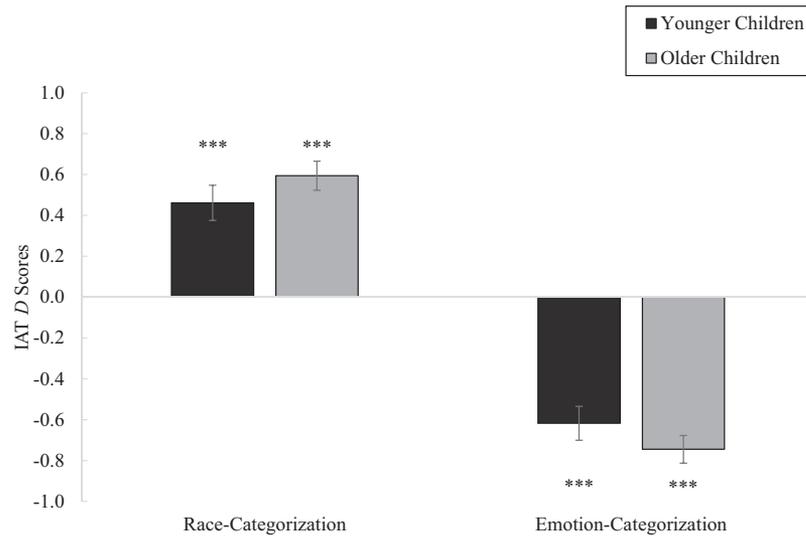
In both the Race-Categorization and Emotion-Categorization conditions, children were asked to sort the same target stimuli. Each face presented was either a Black boy with a smiling expression ($n = 4$) or a White boy with a neutral expression ($n = 4$). These target faces, which differed both by race and emotional expression, were matched for age and attractiveness. The attribute dimension for both IAT conditions was represented by eight pictures from the International Affective Picture System (IAPS; Lang et al., 2005) and included four pleasant (i.e., sunflowers, kittens, puppies, beautiful landscape) and four unpleasant (i.e., garbage, oil spill, demolished house, landfill) images.

Children in the Race-Categorization condition saw a header containing a cartoon image of a White boy on one side of the screen and a Black boy on the other side of the screen (see Appendix). The cartoon faces in the header had the mouth removed so that no emotional expression was visible. Children in the Emotion-Categorization condition saw a header containing a line drawing of a smiling face on one side of the screen and a neutral face on the other side of the screen. For both IAT conditions, in blocks where participants categorized the attribute concept (pleasant and unpleasant pictures), the header showed a sun with the word 'pleasant' written above it on one side of the screen and a storm cloud with the word 'unpleasant' written above it on the other side of the screen.

In Block 1 of each Child IAT, participants were introduced to the target concept. For eight trials, participants were presented with a single smiling Black or neutral White target face on the computer screen. Using the header as their guide, they were asked to sort pictures by matching them to the headers, which led them to categorize faces either by race (in the Race-Categorization condition) or emotional expression (in the Emotion-Categorization condition) using two computer keys. In Block 2, participants were asked to categorize the pleasant and unpleasant images with the same two computer keys, using the header as a guide. Blocks 3 and 4 contained one set of critical trials (8 practice and 16 real). In this set of critical trials, participants categorized neutral White faces and pleasant images using one computer key and categorized smiling Black faces and unpleasant images using another computer key. In Block 5, participants completed 8 trials sorting only the target faces once again, but this time with the response keys reversed. Finally, in Blocks 6 and 7, participants completed the second set of 24 critical trials, in which they now categorized

² Children whose parent identified their race as Black, multiracial with Black ancestry, or who did not indicate their race were excluded a priori from all analyses. This included six children in Study 1, nine children in Study 2, and twelve children in Study 3.

Figure 1
D Scores by Age and Categorization Condition (Study 1)



Note. Higher D scores indicate a relatively faster association between neutral White faces and pleasant images versus smiling Black faces and unpleasant images on the Child Implicit Association Test (IAT), whereas lower scores indicate relatively faster associations between smiling Black faces and pleasant images versus neutral White faces and unpleasant images. Error bars represent the SE . Asterisks indicate whether the mean was significantly different from 0. *** $p < .001$.

smiling Black faces and pleasant images using one computer key and neutral White faces and unpleasant images using the other key. Importantly, participants were randomly assigned to either complete the blocks as described above or to complete this measure with the order of the critical blocks reversed (i.e., with Blocks 6 and 7 described above as the first set of critical trials and Blocks 3 and 4 as the second set). The side on which each pairing appeared was also counterbalanced across participants. Images presented within blocks appeared in random order.

Throughout the task, a correct response was required on each trial; if the participant pressed the incorrect response key, an “X” appeared underneath the stimulus, and remained until the participant provided the correct response. We therefore made use of the natural response times and did not penalize incorrect responses. At the end of the study, each child was debriefed and given a certificate and a sticker as a token of appreciation.

Results and Discussion

For each participant, an IAT D score was created following the scoring algorithm recommended by Greenwald et al. (2003). Specifically, we calculated the difference between the mean response latency for the trials in each critical block and divided this number by the associated pooled standard deviation for each participant. Higher D scores indicate a relatively faster association between neutral White faces and pleasant images versus smiling Black faces and unpleasant images on the Child IAT. D scores for all three studies were calculated using the same scoring algorithm.

A 2 (Condition: Race-Categorization, Emotion-Categorization) \times 2 (Age Group: Younger, Older) ANOVA using D scores as the dependent variable revealed a significant main effect of Condition, $F(1,$

114) = 244.01, $p < .001$, $\eta_p^2 = .68$. Children in the Race-Categorization condition ($D = .54$, $SD = .44$), were faster when pairing neutral White faces with pleasant images and smiling Black faces with unpleasant images than when pairing smiling Black faces with pleasant images and neutral White faces with unpleasant images, relative to participants in the Emotion-Categorization condition ($D = -.69$, $SD = .39$), see Figure 1.³ The main effect of Age Group, $F(1, 114) = .001$, $p = .97$, $\eta_p^2 < .001$, and the interaction, $F(1, 114) = 2.83$, $p = .10$, $\eta_p^2 = .02$, were not significant.⁴

The results of Study 1 provide evidence that how children categorize people has consequences for the affective associations that are elicited. When children categorized targets by race, they showed more positive associations with neutral White faces relative to smiling Black faces in comparison to when these faces were categorized by emotional expression. This finding is significant for two reasons. First, this finding suggests that children can attend to and categorize targets according to different perceptually salient categories when instructed to do so. More importantly, we found that these different bases for categorization had consequences for the affective reactions elicited by these faces among our child participants. To our knowledge, these are the first results to demonstrate non-Black children being faster when associating pleasant stimuli with Black target faces

³ In each study, we also compared the IAT D scores to 0 using one-sampled t -tests. As there has been debate as to whether a comparison to 0 on the IAT provides a meaningful, or arbitrary, metric (e.g., Blanton & Jaccard, 2006; Mitchell & Tetlock, 2017, cf. Jost, 2019), we provide these analyses in the supplement and suggest that caution should be used in the interpretation of these scores.

⁴ See the online supplemental materials for an additional control condition in Studies 1 and 2.

and unpleasant stimuli with White targets faces on a child-friendly IAT. It is important to recognize that this pattern *only* occurred when children categorized smiling Black and neutral White faces by emotion and *not* by race. This finding suggests that although children show pro-White affective responses when White and Black targets are categorized by race, their associations differ significantly when these same targets are categorized by an alternative perceptually salient category.

Study 2

The goal of Study 2 was to replicate and extend the findings from Study 1 by examining children's *spontaneous* categorization and the ensuing consequences for their associations with smiling Black and neutral White faces. Specifically, in Study 2, children were randomly assigned to a Race-Categorization, Emotion-Categorization, or Ambiguous-Categorization condition. The first two conditions were identical to Study 1. In the new Ambiguous-Categorization condition, participants were again asked to categorize smiling Black and neutral White target faces, but unlike the other two conditions, the header did not instruct children to categorize by race or emotional expression. Instead, they were presented with a header that contained both pieces of information, specifically, a smiling Black and neutral White cartoon image.

To determine how participants were primarily categorizing target faces throughout the task, we included five additional trials at the end of each IAT. The five stimuli were randomly selected from a set of six new images that included three Black targets with neutral expressions (rather than smiling expressions in the previous blocks) and three White targets with smiling expressions (rather than neutral expressions in the previous blocks). All images were again matched for age and attractiveness. We selected five trials to have a sufficient number to determine which aspect (race or emotional expression) children were using to categorize these new images and to sufficiently limit the number to ensure that these trials appeared to be seamlessly part of the task. We were interested in examining whether differences would emerge between younger (5–7 years) and older (8–10 years) children in their tendency to spontaneously categorize these multiply categorizable targets by race. We assessed this by examining whether children sorted the majority of these final five trials by race (e.g., matching smiling White targets with the neutral White header) or by emotional expression (e.g., matching smiling White targets with the smiling Black header).

In addition, we examined whether children's score on the ambiguous-categorization IAT would differ depending on the categorization of these final five trials. We hypothesized that regardless of age, children who spontaneously categorized these final trials primarily by race would be relatively faster at pairing neutral White targets with pleasant images and smiling Black targets with unpleasant images during the IAT relative to those who categorized these final trials primarily by emotional expression.

Method

Participants

One hundred ten non-Black children between the ages of 5 and 10 years (57 female, 53 male) were recruited from and tested in a local science museum. Of the parents who provided their annual

household income ($n = 72$), the majority (57%; 41 of 72) reported earning at least \$75,000 per year. The data from thirteen children were excluded prior to analyses because there was a technical error resulting in no data ($n = 4$) or because they had greater than 10% of response latencies less than 300 ms ($n = 1$), had D scores that were numerical outliers exceeding three standard deviations from the mean ($n = 2$), or had a response error rate of 35% or greater ($n = 6$; Cvencek et al., 2011). The final sample consisted of 97 participants, which included 44 younger children who ranged in age from 5 to 7 years ($M_{\text{age}} = 6$ years; 14 five-year-olds, 20 six-year-olds, and 10 seven-year-olds) and 53 older children who ranged in age from 8 to 10 years ($M_{\text{age}} = 9$ years; 23 eight-year-olds, 16 nine-year-olds, and 14 ten-year-olds). The racial makeup of the sample consisted of 53 White, 18 East/South East Asian, 15 South Asian/Arab, and 11 multiracial participants. A sensitivity power analysis using G*Power (Version 3.1.9.4) indicates that this sample is sufficient to detect effects as small as $\eta^2 = .09$ ($f = .32$) at .80 power.

Measures and Procedure

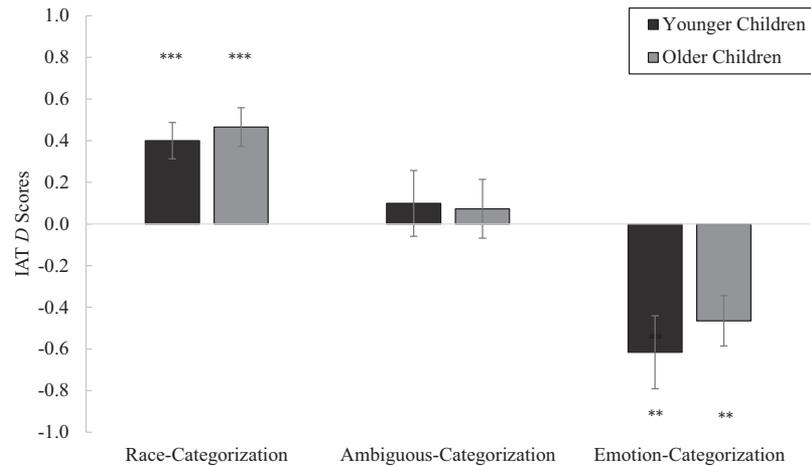
After providing verbal assent to participate, each child was seated in front of a laptop computer in one of four testing cubicles located in a quiet area of the museum and was randomly assigned by a non-Black experimenter to one of the three conditions. The target images used in all three conditions were identical and included faces of smiling Black boys and neutral White boys. Children in the Race-Categorization and Emotion-Categorization conditions completed the same Child IATs described in Study 1. Children in the Ambiguous-Categorization condition were presented with the same stimuli but saw a header that contained both race and emotional expression information. Specifically, this header included a cartoon image of a White boy with a neutral expression on one side of the screen and a cartoon image of a smiling Black boy on the other side of the screen (see the Appendix). All children were asked to press one computer key if the face that appeared in the center of the computer screen looked similar to the picture provided in the left header and to press a different key if the face in the center looked similar to the picture on the right header. The order and side for each of the critical blocks was again counterbalanced between participants.

To determine whether children were categorizing targets primarily by race or emotional expression, all participants completed five additional *spontaneous categorization* trials at the end of their IAT. During these trials, the header from the last IAT block of critical trials remained on the screen, but participants were now required to categorize smiling White targets and neutral Black targets. Upon completion of the study, each child was debriefed and given a certificate and a sticker.

Results and Discussion

To determine whether children's associations differed by condition, we first conducted a 3 (Condition: Race-Categorization, Emotion-Categorization, Ambiguous-Categorization) \times 2 (Age Group: Younger, Older) ANOVA using D scores as the dependent variable. The anticipated main effect of condition was significant, $F(2, 91) = 25.37, p < .001, \eta_p^2 = .36$ (see Figure 2). Simple effects analyses revealed that children in the Race-Categorization condition ($D = .44, SD = .37$) were relatively faster at pairing neutral

Figure 2
D Scores by Age and Categorization Condition (Study 2)



Note. Higher D scores indicate a relatively faster association between neutral White faces and pleasant images versus smiling Black faces and unpleasant images on the Child Implicit Association Test (IAT), whereas lower scores indicate relatively faster associations between smiling Black faces and pleasant images versus neutral White faces and unpleasant images. Error bars represent the SE . Asterisks indicate whether the mean was significantly different from 0. ** $p < .01$. *** $p < .001$.

White faces with pleasant images and smiling Black faces with unpleasant images when compared with children in both the Emotion-Categorization ($D = -.52$, $SD = .53$; $p < .001$) and the Ambiguous-Categorization ($D = .09$, $SD = .63$; $p = .02$) conditions. In addition, children in the Emotion-Categorization condition were relatively faster at pairing smiling Black faces with pleasant images and neutral White faces with unpleasant images compared with children in the Ambiguous-Categorization condition ($p < .001$). The main effect of Age Group, $F(1, 91) = .34$, $p = .56$, $\eta_p^2 = .004$, and the interaction, $F(2, 91) = .22$, $p = .80$, $\eta_p^2 = .005$, were not significant.

Next, we examined how children categorized the majority of the final five *spontaneous categorization* trials across each of the three conditions (see Table 1). Specifically, we examined whether the majority of these trials were categorized by race or by emotional expression. As expected, almost all of the children (97%; 32 of 33) in the Race-Categorization condition sorted the majority (at least three) of the final five trials by race. By contrast, almost all children (96%; 27 of 28) in the Emotion-Categorization condition sorted the majority of the trials by emotional expression. After following specific instructions on how to categorize faces in these two conditions, most children continued to use this strategy in the final five trials of the task.

Of greatest interest was how children in the Ambiguous-Categorization condition would categorize these final target faces. Specifically, we examined whether they sorted these faces primarily by race (i.e., neutral Black targets with smiling Black header and smiling White targets with neutral White header) or by emotional expression (i.e., neutral Black targets with neutral White header and smiling White targets with smiling Black header). The results demonstrated that more children in this condition categorized these final trials primarily by race (67%; 24 of 36) than by

emotional expression (33%; 12 of 36), $\chi^2(1) = 4.00$, $p = .046$. Results from additional exploratory analyses revealed that older children were more likely to categorize the majority of these final trials by race (89%) than were younger children (56%), $\chi^2(1) = 8.00$, $p = .005$. However, owing to the relatively small sample size, caution must be taken in the interpretation of these results.

Importantly, and as expected, children in the Ambiguous-Categorization condition who categorized the final *spontaneous categorization* trials primarily by race were relatively faster at pairing smiling Black faces with unpleasant images and neutral White faces with pleasant images ($D = .31$, $SD = .53$) in comparison with children who primarily sorted these final trials by emotional expression ($D = -.35$, $SD = .58$), $t(34) = -3.38$, $p = .002$, $d = -1.20$. In addition, the magnitude of the IAT effect was related to categorization; the number of the final five trials that were categorized by race (0–5) was positively correlated with the relative degree of positive associations with neutral White, as compared with smiling Black, targets, $r_p(36) = .48$, $p = .003$. These findings are consistent with the possibility that children who sorted these final trials primarily by race were more likely to be categorizing by race throughout the task, leading them to be faster to pair neutral White faces with pleasant images and smiling Black faces with unpleasant images relative to the reverse pairing. By contrast, those children who categorized the final five trials primarily by emotion were likely categorizing by emotion throughout the task, leading to relatively faster responding when pairing smiling Black targets with pleasant images and neutral White targets with unpleasant images.

The results of Study 2 provide additional evidence that children's associations with smiling Black and neutral White targets depend on how they are categorized. Consistent with the findings in Study 1, children assigned to the Race-Categorization condition

Table 1
Number of Final Trials Categorized by Race Separated by Condition (Study 2)

Condition	Majority sorted by race						Majority sorted by emotion					
	All by race		Mixed				All by emotion					
	5/5		4/5		3/5		2/5		1/5		0/5	
	N	%	N	%	N	%	N	%	N	%	N	%
Race-Categorization condition ($n = 33$)	19	58%	11	33%	2	6%	1	3%	0	0%	0	0%
Emotion-Categorization condition ($n = 28$)	0	0%	1	3%	0	0%	2	7%	3	11%	22	79%
Spontaneous-Categorization condition ($n = 36$)	17	47%	5	14%	2	6%	5	14%	3	8%	4	11%
All participants ($n = 97$)	36	37%	17	18%	4	4%	8	8%	6	6%	26	27%

Note. All participants sorted five additional *spontaneous categorization* trials that followed seamlessly at the end of their IAT. During these trials, the header from the last IAT block of critical trials remained on the screen and participants were now required to categorize smiling White targets and neutral Black targets. The number (N) and percentage (%) refers to how many participants sorted a specific number of the final five trials by race. Participants either sorted the majority of these trials by race (i.e., for 5/5, 4/5, or 3/5 of the trials they paired a neutral Black target with the smiling Black header or a smiling White target with the neutral White header) or by emotional expression (i.e., pairing a neutral Black target with the neutral White header or a smiling White target with the smiling Black header resulting in 2/5, 1/5, or 0/5 of the trials being sorted by race).

were relatively faster when associating neutral White targets with pleasant images and smiling Black targets with unpleasant images, whereas children who categorized these same targets by emotional expression showed relatively more positive associations with the smiling Black targets. These findings replicate the results from Study 1 that non-Black children can show more positive associations with Black over White targets on the IAT; *however*, this pattern only emerged when children were asked to categorize these targets by emotional expression rather than race.

Notably, children in the Ambiguous-Categorization condition showed a different pattern of associations from children in either of the other two conditions, with the average IAT D score falling between both other conditions. The final *spontaneous categorization* trials provide additional insight into these findings. Children who categorized the final five trials primarily by race were relatively faster when associating neutral White faces with pleasant images and smiling Black faces with unpleasant images compared with children who categorized the final five trials primarily by emotional expression. These results suggest that children who sorted the final trials by race were categorizing primarily by race throughout the task, and their association with these targets reflected more positive spontaneous evaluative responses toward White versus Black faces. By contrast, children who sorted the final trials primarily by emotion were categorizing by emotional expression throughout the task, and their association with these same targets reflected more positive associations with smiling as opposed to neutral faces.

Finally, these *spontaneous categorization* trials provide some initial evidence, albeit with a small sample size, that older children are more likely to spontaneously categorize targets by race than by emotional expression as compared with younger children. This finding is consistent with other research suggesting that the tendency to spontaneously categorize by race during person perception increases as children age (Degner & Wentura, 2010; Kawakami et al., 2020; cf. Williams & Steele, 2019). It seems possible that this effect is due to children's increasing exposure across development to race-based categorizations and biases in daily life, which may lead them, like adults, to be more likely to automatically attend to race, a point we return to in the General Discussion.

Study 3

The results from Studies 1 and 2 suggest that how children categorize others has a direct impact on their affective associations. Both when we manipulated the way that children categorized target faces in Studies 1 and 2, and when we examined individual differences in children's spontaneous categorization in Study 2, children's associations differed. Children who categorized by race were relatively faster to pair neutral White faces with pleasant stimuli and smiling Black faces with unpleasant stimuli in comparison to those who categorized by emotional expression. In Study 3, our goal was to replicate the results of the Categorization condition with a larger sample of children. We also aimed to extend the results by examining whether a brief intervention, in which an adult modeled categorization by emotional expression, would increase children's tendency to also use this feature as a basis for categorization, thereby influencing their affective responses.

Given the complexities of daily life, there are many factors that can influence the dimensions that children spontaneously use to categorize others. According to the Developmental Intergroup Theory (DIT; Bigler & Liben, 2006, 2007), children are more likely to spontaneously categorize others by social categories that are psychologically salient. One way to make a dimension salient to children is to label and explicitly use that dimension. For example, by referring to children in a classroom as "boys and girls" or by separating boys and girls for gym class, teachers may increase the psychological salience of gender for children. However, even in the absence of explicit labeling, theory and research suggest that children may also infer the psychological importance of a social or perceptual category if they see it being used in their environment (Bigler & Liben, 2007; Rhodes & Baron, 2019).

In the current study we investigated whether watching an adult categorize faces by emotional expression prior to completing an ambiguous-categorization IAT would impact children's spontaneous associations with others who differ by both emotional expression and race. To accomplish this goal, children were presented with the ambiguous-categorization IAT used in Study 2, containing smiling Black and neutral White target faces. Before completing this task, they were randomly assigned to either a Blatant Intervention, Subtle Intervention, or Control condition.

In both Intervention conditions, children watched as an adult experimenter sorted laminated cards with pictures of racially diverse children by their emotional expression. In the Blatant Intervention condition, the adult explicitly stated that they were sorting the images based on their emotional expression. In the Subtle Intervention condition, the adult simply sorted the laminated pictures by emotional expression but made no explicit comment as to their strategy for sorting. By contrast, children in the Control condition watched an adult sort a nonsocial set of images by shape. The inclusion of these three conditions allowed us to examine whether modeling by an adult would influence children's subsequent associations and whether explicit labeling was critical for the efficacy of the intervention.

We predicted that children in both Intervention conditions would be (a) relatively faster to pair smiling Black targets with pleasant stimuli and neutral White targets with unpleasant stimuli on the ambiguous-categorization IAT relative to children in the control condition, and (b) more likely to sort the final five categorization trials primarily by emotional expression as opposed to race. In addition, we anticipated that these effects would be more pronounced for children in the Blatant Intervention condition, where an adult's focus on emotion as a basis for social categorization would be reinforced by explicit labeling (Bigler & Liben, 2007).

Method

Participants

One hundred eighty non-Black children between the ages of 5 and 10 years (84 female and 96 male) were recruited from and tested in a local science museum. Of the parents who provided their annual household income ($n = 128$), the majority (62%; 79 of 128) reported earning at least \$75,000 per year. The data from twelve participants were excluded prior to analyses because there was a technical error resulting in incomplete IAT data ($n = 3$), or because more than 10% of response latencies were less than 300 ms ($n = 1$), the D score was a numerical outlier exceeding three standard deviations from the mean ($n = 1$), or response error rates were 35% or greater ($n = 7$; Cvencek et al., 2011), leading to a final sample of 168 participants. The final sample included 83 younger children who ranged in age from 5 to 7 years ($M_{\text{age}} = 6$ years; 26 five-year-olds, 27 six-year-olds, and 30 seven-year-olds) and 85 older children who ranged in age from 8 to 10 years ($M_{\text{age}} = 9$ years; 31 eight-year-olds, 36 nine-year-olds, and 18 ten-year-olds). The racial makeup of the sample consisted of 88 White, 39 East/South East Asian, 20 South Asian/Arab, 15 multi-racial, four Latinx, and two Aboriginal Canadian participants. A sensitivity power analysis using G*Power (Version 3.1.9.4) indicates that this sample is sufficient to detect effects as small as $\eta^2 = .06$ ($f = .24$) at .80 power.

Materials and Procedure

After receiving parental consent, each participant was brought to an individual testing cubicle in a quiet location of the museum and was randomly assigned to one of three conditions: Blatant Intervention, Subtle Intervention, or Control. Participants in all three conditions first completed a phase in which a non-Black experimenter sorted a set of eight laminated cards on a table in front of the child. In the two Intervention conditions, the cards consisted of eight color

photographs of child faces that varied by race (four White, two Black, one South Asian, one East Asian), sex (four male, four female), and emotional expression (four smiling, four neutral). In the Control condition, stimuli consisted of eight color pictures of shapes that varied by color (four green, four orange), size (four big, four small), and shape (four triangles, four circles).

In the *Blatant Intervention* condition, the experimenter showed participants the eight cards and said "Here are a bunch of faces. I'm going to sort them by their emotion. The faces that are smiling go together in one group over here and the faces that are not smiling go together in a different group over here." The experimenter then proceeded to sort the laminated images based on their emotional expression, while explicitly labeling each picture as "smiling" or "not smiling." In the *Subtle Intervention* condition, the experimenter showed participants the eight cards and said, "Here are a bunch of faces. I'm going to sort them into two groups." The experimenter then proceeded to sort them by emotional expression, but simply indicated for each face that it "goes here." In the *Control* condition, the experimenter used language comparable to the Subtle Intervention condition but sorted the pictures of triangles and circles according to their shape. In each condition, after the cards were sorted, the experimenter said, "Now take a look at how I sorted these faces/cards because I might ask you about this later," and allowed the child to examine the cards for 3 seconds. Following this intervention, each child was seated in front of a laptop computer and completed the same ambiguous-categorization IAT described in Study 2, which included the five final spontaneous categorization trials. Upon completion of the study, children were debriefed and given a certificate and sticker.

Results and Discussion

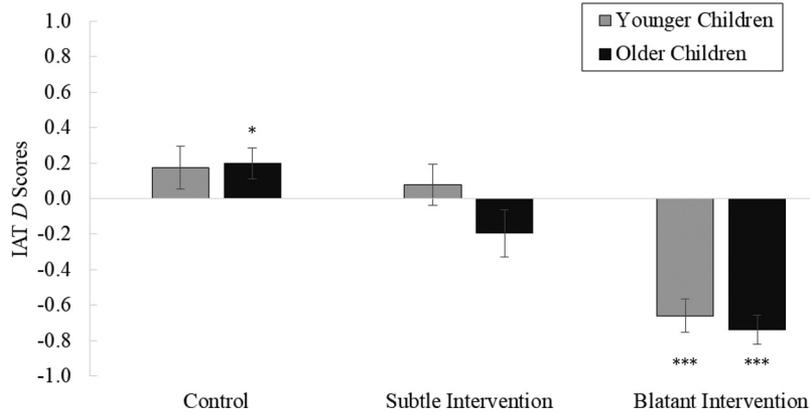
To determine whether children's associations differed by condition, we first conducted a 3 (Intervention Condition: Blatant Intervention, Subtle Intervention, Control) \times 2 (Age Group: Younger, Older) ANOVA using D scores as the dependent variable. As expected, the main effect of Intervention Condition was significant, $F(2, 162) = 36.03$, $p < .001$, $\eta_p^2 = .31$. Simple effects analyses revealed that children in the Blatant Intervention condition were faster at associating smiling Black faces with pleasant images and neutral White faces with unpleasant images relative to the reverse pairing ($D = -.74$, $SD = .45$), when compared with children in the Subtle Intervention ($D = -.05$, $SD = .66$), and the Control ($D = .19$, $SD = .55$) conditions, $ps < .001$ (see Figure 3). Although children in the Subtle Intervention condition were relatively faster at pairing smiling Black targets with pleasant images and neutral White targets with unpleasant images than children in the Control condition, $p = .07$, this difference was not significant. The main effect of Age Group, $F(1, 162) = 1.56$, $p = .21$, $\eta_p^2 = .01$, and the interaction, $F(2, 162) = 1.01$, $p = .37$, $\eta_p^2 = .01$, were not significant.

Next, we examined how children categorized the majority of the final five *spontaneous categorization* trials across each of the three conditions (see Table 2). First, we examined the Control condition. Replicating the results from Study 2, more children categorized the majority (i.e., three or more of the final five trials) of these trials by race (78%; 47 of 60) as opposed to emotional expression (22%; 13 of 60), $\chi^2(1) = 19.27$, $p < .001$. Furthermore, we replicated the age difference in categorization with this larger

F3

T2

Figure 3
D Scores by Age and Intervention Condition (Study 3)



Note. Higher *D* scores indicate a relatively faster association between neutral White faces and pleasant images versus smiling Black faces and unpleasant images on the Child Implicit Association Test (IAT), whereas lower scores indicate relatively faster associations between smiling Black faces and pleasant images versus neutral White faces and unpleasant images. Error bars represent the *SE*. Asterisks indicate whether the mean was significantly different from 0.
 * $p < .05$. *** $p < .001$.

T3 sample. Older children (88%) were again more likely than younger children (67%) to categorize the majority of these trials by race, $\chi^2(1) = 3.94, p = .047$ (see Table 3). In the Subtle Intervention condition, children did not significantly differ in their tendency to categorize the majority of the final trials primarily by race (62%; 34 of 55) as opposed to emotional expression (38%; 21 of 55), $\chi^2(1) = 3.07, p = .08$, and no age difference emerged, $\chi^2(1) = .36, p = .55$. By contrast, in the Blatant Intervention condition, where an adult explained that they were sorting racially diverse laminated cards by emotional expression in a seemingly unrelated task, children were *less* likely to categorize the majority of the final trials on the Child IAT by race (25%; 13 of 52) compared with emotional expression (75%; 39 of 52), $\chi^2(1) = 13.00, p < .001$. Older children (8% by race) were more influenced by this blatant intervention than younger children (41% by race), $\chi^2(1) = 7.42, p < .006$.

To further determine the downstream consequences of categorization, we first examined whether, regardless of Intervention

Condition, children who categorized the final five trials primarily by race showed a different pattern of associations from those who categorized these trials primarily by emotional expression. A 2 (Final Categorization Trials: Majority by Race, Majority by Emotion) \times 2 (Age Group: Younger, Older) ANOVA using *D* scores as the dependent variable revealed a significant main effect of Final Categorization Trials, $F(1, 163) = 61.68, p < .001, \eta_p^2 = .28$. As expected, when collapsing across conditions, participants who spontaneously categorized the final five trials primarily by race ($D = .14, SD = .59$) were faster to pair neutral White faces with pleasant images and smiling Black faces with unpleasant images relative to the reverse pairing when compared with children who categorized the final trials primarily by emotional expression ($D = -.56, SD = .55$). The main effect of age, $F(1, 163) = .62, p = .43, \eta_p^2 = .004$, and the interaction, $F(1, 163) = 3.78, p = .054, \eta_p^2 = .02$, were not significant. However, additional analyses confirmed that the effects of categorization, based on these final trials, emerged both for younger, $F(1, 81) = 13.74,$

Table 2
Number of Final Trials Categorized by Race Separated by Condition (Study 3)

Condition	Majority sorted by race						Majority sorted by emotion					
	All by race		Mixed				All by emotion					
	5/5		4/5		3/5		2/5		1/5		0/5	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Control condition ($n = 60$)	19	31%	18	30%	10	17%	4	7%	3	5%	6	10%
Subtle Intervention condition ($n = 55$)	23	42%	5	9%	6	11%	4	7%	5	9%	12	22%
Blatant Intervention condition ($n = 52$)	7	13%	2	4%	4	8%	6	12%	8	15%	25	48%
All participants ($n = 167$)	49	29%	25	15%	20	12%	14	8%	16	10%	43	26%

Note. The number (*N*) and percentage (%) refers to how many participants sorted a specific number of the final five *spontaneous categorization* trials by race. Participants either sorted a specific number of the majority of these trials by race (i.e., for 5/5, 4/5, or 3/5 of the trials they paired a neutral Black target with the smiling Black header or a smiling White target with the neutral White header) or by emotional expression (i.e., pairing a neutral Black target with the neutral White header or a smiling White target with the smiling Black header resulting in 2/5, 1/5, or 0/5 of the trials being sorted by race).

Table 3

Number and Percentage of Participants, by Age Group and Condition, Who Categorized the Final Spontaneous Categorization Trials Primarily by Race as Opposed to Emotional Expression (Study 3)

Condition	Younger children		Older children	
	<i>N</i>	%	<i>N</i>	%
Control condition (<i>n</i> = 60)	18/27	67%	29/33	88%
Subtle Intervention condition (<i>n</i> = 55)	19/29	66%	15/26	58%
Blatant Intervention condition (<i>n</i> = 52)	11/27	41%	2/25	8%
All participants (<i>n</i> = 167)				

Note. The number (*N*) and percentage (%) refers to participants who sorted the majority (3/5 or more) of the final five *spontaneous categorization* trials by race.

$p < .001$, $\eta_p^2 = .15$, and older, $F(1, 82) = 64.93$, $p < .001$, $\eta_p^2 = .44$, children.

Next, we examined whether, regardless of Intervention Condition, children who categorized the final five trials by a combination of race and emotional expression showed a different pattern of bias than participants who categorized all targets by either race or emotional expression. A 3 (Final Categorization Trials: All by Race, Mixed, All by Emotional Expression) \times 2 (Age Group: Younger, Older) ANOVA on IAT *D* scores revealed a main effect of Final Categorization Trials, $F(1, 161) = 22.32$, $p < .001$, $\eta_p^2 = .25$. Simple effect tests confirmed that children who categorized all of the final five trials by race ($D = .17$, $SD = .58$) were faster at associating neutral White faces with pleasant images and smiling Black faces with unpleasant images relative to the reverse pairing as compared with children who categorized all of the final trials by emotional expression ($D = -.67$, $SD = .50$; $p < .001$) and children who categorized the final trials by both race and emotional expression ($D = -.10$, $SD = .66$; $p = .04$). The IAT *D* scores of those who categorized all of the final five trials by emotional expression also differed in the anticipated direction from those who showed mixed categorization ($p < .001$). The main effect of age, $F(1, 161) = .44$, $p = .51$, $\eta_p^2 = .003$, and the interaction, $F(1, 161) = 2.69$, $p = .07$, $\eta_p^2 = .03$, were not significant.

Finally, we examined whether children's IAT effect was again related to categorization of the final trials. The number of the final five trials that were categorized by race (0–5) was significantly correlated with the speed with which children associated neutral White faces with pleasant images and smiling Black faces with unpleasant images relative to the reverse pairing, $r_p(167) = .39$, $p < .001$. Taken together, these findings suggest that the more children consistently categorized by race as opposed to emotional expression, the more their IAT *D* scores reflected race-based, as opposed to emotion-based, associations.

In Study 3, after watching and listening to an experimenter explain that emotional expression was being used to categorize racially diverse laminated images in the Blatant Intervention condition, children were more likely to categorize faces in a subsequent computer task by emotional expression as opposed to race. This was particularly true for older children. Whereas 88% of older children categorized the majority of the final five trials by race in the control condition, only 8% did so after watching an

adult sort racially diverse pictures by emotional expression and explicitly describe their strategy. This finding suggests that, like adults (Steele, George, Cease, et al., 2018), older children are more likely to spontaneously categorize targets by race as opposed to emotional expression. However, when exposed to an alternative way of categorizing others, children adopted this approach. In turn, this categorization strategy had consequences for spontaneous associations elicited by smiling Black and neutral White targets. Specifically, children in the Blatant Intervention condition were relatively faster at pairing smiling Black faces with pleasant stimuli and neutral White faces with unpleasant stimuli when compared with children in the two other conditions. The findings also suggest that the Subtle Intervention condition, where an adult made use of emotional expression but did not explain the basis of categorization, was less effective in changing children's subsequent categorization processes.

General Discussion

Research has repeatedly demonstrated that non-Black children are faster at associating White faces with pleasant images and Black faces with unpleasant images relative to the reverse pairing on child-friendly versions of the IAT (Baron & Banaji, 2006; Dunham et al., 2006, 2008; Gonzalez et al., 2017, 2021; Newheiser & Olson, 2012; Rutland et al., 2005; Steele, George, Williams, & Tay, 2018; Williams & Steele, 2019). In the current research, we examined children's associations with Black and White targets who also differed by emotional expression. By presenting children with smiling Black and neutral White targets, we were able to examine how social categorization processes can impact children's associations with targets who differ systematically across multiple perceptually salient categories.

We found that when children were asked to categorize targets by race in Studies 1 and 2, both younger and older children were faster at pairing pleasant images with neutral White faces and unpleasant images with smiling Black faces relative to the reverse pairing, as compared with when they were asked to categorize targets by emotional expression. This effect of categorization was not moderated by the age. This finding is consistent with two studies by Steele, George, Cease, et al. (2018) in which adults showed a similar pattern when categorizing targets by race on the IAT when presented with smiling Black and neutral White target faces as in the current studies. However, to our knowledge, this is the first research to show relatively more positive responding toward Black targets on an implicit measure with non-Black child participants. Importantly, this *only* occurred when children categorized these smiling Black faces and neutral White faces by emotional expression. Our results suggest that facilitated responding to positive associations with smiling Black, over neutral White, target faces occurred because children were categorizing these faces using a dynamic perceptual category, specifically emotional expression, rather than a stable social category like race.

To test this assumption, we examined how children spontaneously categorized faces in the final five trials in Studies 2 and 3. When presented with smiling Black and neutral White targets, both younger and older children were more likely to spontaneously categorize the majority of the final trials by race than by emotional expression. Furthermore, those who categorized these final trials primarily by race showed more positive associations with neutral

White, compared with smiling Black, targets than those who categorized the trials primarily by emotional expression. This was the case in both the Ambiguous-Categorization condition in Study 2 and in the Control condition in Study 3. Together these results provide evidence that spontaneous categorization processes have consequences for the evaluation of targets.

We also found evidence that the consistency with which children categorized faces according to race or emotional expression in the final five trials predicted the speed with which they associated neutral White, or smiling Black, targets with pleasant versus unpleasant images on the IAT. In Studies 2 and 3, the more children categorized the final five trials by race, the faster they were at associating White faces with pleasant images and Black faces with unpleasant images relative to the reverse pairing during the IAT. In addition, in Study 3, children who categorized some of the final trials by race and some by emotional expression showed less bias in either direction on the IAT relative to those who either categorized all of the final trials by race or by emotional expression.

Together, these findings provide some initial insight into the dynamic nature of categorization and the influence that this process can have on affective responses during person perception. It seems possible that children who showed a mixed response pattern in the final five trials were either classifying according to both categories or even a subcategory throughout the task or their classification strategy switched and/or alternated throughout the task, either intentionally or unintentionally. Continuing to understand the conditions under which children prioritize certain perceptually salient categories is a worthy avenue for future research to better understand how categorization processes and person perception change across development.

In addition, across both studies, older children were more likely than younger children to spontaneously categorize the final trials by race. These findings provide further insight into the results of a study by Gonzalez et al. (2017), in which older, but not younger, children's racial biases were attenuated relative to control conditions following exposure to positive Black exemplars. In their study, White and Asian children were presented with four stories, each depicting a positive role model in the community. In the experimental condition, a picture of a Black adult accompanied each story. It is possible that older children in the Gonzalez et al. (2017) study were more likely to spontaneously categorize the positive Black role models by their race and this increased the efficacy of the intervention leading to decreased racial bias for only the older children.

In Study 3 we also examined the impact of modeling on children's tendency to spontaneously use race as a basis for social categorization. Children who watched an adult explicitly label and categorize people according to their emotional expression were subsequently more likely to spontaneously categorize the final trials of a Child IAT by emotional expression rather than race and were also faster to associate smiling Black faces with pleasant images and neutral White faces with unpleasant images relative to the reverse pairing on the Child IAT. These findings suggest that exposing children to an adult modeling a different categorization strategy has the potential to decrease the psychological salience of race (Bigler & Liben, 2007) and increase the use of other perceptual cues as a basis for categorization. This intervention was particularly effective for older children, providing additional evidence that late childhood might be an ideal developmental period for

interventions aimed at challenging children's racial biases (Gonzalez et al., 2017; Gonzalez et al., 2021). Older children have the benefit of additional cognitive flexibility and awareness of social norms relative to younger children yet may have had less exposure to cultural messages inculcating racial biases relative to adults.

Implications and Future Directions

Although our results may initially seem to suggest that adults should downplay the salience and importance of race with children by taking a "colorblind" approach (Apfelbaum et al., 2010; Holoien & Shelton, 2012; Purdie-Vaughns et al., 2008), we believe some caution is warranted in drawing this conclusion. There is evidence that adopting a colorblind approach can have negative consequences. For example, when compared with *multiculturalism*—an approach that acknowledges and celebrates diversity—*colorblindness* has been found to be associated with greater levels of racial bias on both implicit and explicit measures (Richeson & Nussbaum, 2004), more negative nonverbal behavior during interracial interactions (Apfelbaum et al., 2008; Holoien & Shelton, 2012), and more negative perceptions of interracial interactions (Karmali et al., 2019) among White participants, as well as greater disengagement in the workplace among ethnic minority coworkers (Plaut et al., 2009). Children exposed to colorblind messages are also less likely to detect blatant racial discrimination and to report discriminatory behavior in a way that would prompt intervention from an adult (Apfelbaum et al., 2010).

Additional research suggests that discussing racism with children can have positive outcomes. Children who were explicitly taught about historical racial discrimination had more positive explicit attitudes toward people who are Black and valued racial fairness to a greater degree than children who were not explicitly taught about racial discrimination (Hughes et al., 2007). Taken together, we believe that it is worth acknowledging that there are many contexts in which it is important to draw children's attention to race to teach children to appreciate the benefits of living in a multicultural society as well as the detrimental impact of prejudice and discrimination. However, there are other situations when encouraging children to categorize others by race would be inappropriate and could increase negative evaluations or lead children to form inaccurate impressions of others based on stereotypes or perceived outgroup membership. Consistent with this possibility, recent research suggests that exposing adults to messages promoting multiculturalism can increase their race essentialist beliefs relative to being exposed to a colorblind ideology (Wilton et al., 2019), suggesting an important potential caveat to previous findings. Future research is needed to better understand the conditions under which it is helpful to draw children's attention to stable social categories, such as race, and when it is not. At the least, it seems likely that there can be social benefits from reminding children that people have multiple aspects to their identities.

In addition, it is important to note that although this research demonstrates that how children categorize others affects their spontaneous associations with members of different racial groups, this research does not provide direct insight into how to challenge racial biases. There is a growing body of literature suggesting that both adults' (Forscher et al., 2019; Lai et al., 2014) and children's (Gonzalez et al., 2017; Gonzalez et al., 2021; Kawakami et al., 2018; Neto et al., 2016; Qian et al., 2019; Vezzali et al., 2012;

Xiao et al., 2015) spontaneous associations with racial groups can be altered. For example, as noted earlier, research suggests that exposure to positive exemplars can decrease children's racial biases on implicit measures in both the short- and long-term (Gonzalez et al., 2017, 2021; Neto et al., 2016). In addition, recent research has found that having preschool- and school-age children perceptually individuate target faces by teaching children to learn and memorize specific outgroup faces can have a lasting impact on spontaneous racial biases (Qian et al., 2019; Xiao et al., 2015), presumably by disrupting the process of generalizing racial bias to all outgroup members.

A large body of research also suggests that cross-race friendships are associated with more positive racial attitudes, on both implicit and explicit measures (Aberson et al., 2004; Aboud et al., 2003; Page-Gould et al., 2010; Pettigrew, 1998; for a meta-analysis, see Pettigrew & Tropp, 2006). For example, White British children with South Asian friends were found to have more positive associations on implicit measures with people who are South Asian (Turner et al., 2008). Although speculative, given that our research focused on the temporary salience of a competing perceptual feature, the current research may provide one additional, albeit indirect, route through which spontaneously activated racial attitudes can be challenged. Given the pervasiveness of racial biases in childhood and that children in our studies were faster to respond to positive stimuli when paired with smiling Black (relative to neutral White) targets when they categorized them by emotional expression, it is possible that children may be more likely to approach and befriend Black peers if they initially view them through a nonracial lens. Over time, if these friendships evolve and mature, children will have more opportunities to develop positive associations with their friends and in turn their friends' racial group, which may ultimately lead to the automatization of more positive attitudes toward other members of these groups (Al Ramiah & Hewstone, 2013). The results of Study 3 further suggest that adults might be in a particularly good position to direct children to psychologically useful categories, such as shared interests or identities (Dovidio et al., 1998). However, given the dynamic nature of social categorization processes (Kawakami et al., 2020; Pauker et al., 2016), the contextual nature of bias (Payne et al., 2017), the complexity of racial bias development (Baron, 2015; Degner & Calanchini, 2020), and reduction (Paluck et al., 2021), as well as the potential for ironic effects to emerge in the context of intergroup contact (Al Ramiah & Hewstone, 2013; Hayward et al., 2018; Vorauer & Sasaki, 2009), additional real-world data would be useful in examining this possibility.

Other limitations to this research include our use of only one measure of racial attitudes, the IAT (Greenwald et al., 2003), which is a relative category-based reaction time measure, our inclusion of only male targets, and our focus on target stimuli who differed only both by race and emotional expression but not other social or perceptual categories. Although this approach allowed us to examine the role that categorization can play in informing spontaneous affective associations with others who differ in multiple ways, making use of additional implicit measures and social categories can broaden our understanding of the downstream consequences of social categorization processes in childhood. For example, a recent study by Perszyk et al. (2019) used a child-friendly Affect Misattribution Procedure (Payne et al., 2005; Williams et al., 2016; Williams & Steele, 2019) to examine responses

to intersectional targets who varied by race and gender. They found evidence of a pro-White racial bias among four-year-old children. However, this effect was moderated by target gender, such that children showed more positive responses following exposure to Black girls relative to Black boys. Given this finding, it seems possible that a Black child expressing a neutral or negative emotional expression might elicit greater negativity than a White child expressing more neutral emotions, and this is a question worthy of further investigation.

It would be beneficial for future research to investigate whether similar effects emerge when children categorize targets using other stable social categories, such as gender or age. It is also interesting to consider whether emotional expression would have been more salient to children had the White children been frowning. Similarly, it is worth noting that in the current research we did not include a full factorial design to assess whether children's racial biases would be magnified if children were instructed to categorize smiling White and neutral (or angry) Black faces (see Raissi & Steele, *in press*, for related work with adults). One additional avenue for further study is to assess the truly dynamic nature of racial biases across a longer period of time in which a person might be categorized by race at one moment and by a different dimension in the next. These findings do, however, contribute to a growing body of literature suggesting that evaluations of targets are online constructions that integrate group- and emotion-based associations as well as contextual cues in a fluid way (Freeman & Ambady, 2011; Kawakami et al., 2017, 2018).

Conclusions

In conclusion, the current research adds to a growing body of literature that provides a deeper understanding of the conditions under which children show race-based affective associations in response to targets who can be categorized in multiple ways. In the current research, categorization by race was a precondition for bias (see also Williams & Steele, 2019). When children were categorizing smiling Black and neutral White target faces by emotional expression, they showed relatively more positive associations with smiling Black faces as compared with when they were categorizing these faces by race. The results further suggest that adults' actions can inform the social categories that children attend to as well as ultimately children's spontaneously activated attitudes toward others. In our study, this was particularly the case when the adults explicitly labeled the alternate category that they were attending to. Through future research it is hoped that we will continue to increase our understanding of social categorization processes in childhood as well as the malleability and contextual nature of children's spontaneous associations and racial biases.

References

- Aberson, C. L., Shoemaker, C., & Tomolillo, C. (2004). Implicit bias and contact: The role of interethnic friendships. *The Journal of Social Psychology, 144*(3), 335–347. <https://doi.org/10.3200/SOCP.144.3.335-347>
- Aboud, F. E. (2008). A social-cognitive developmental theory of prejudice. In S. M. Quintana & C. McKown (Eds.), *Handbook of race, racism, and the developing child* (pp. 55–71). Wiley and Sons. <https://doi.org/10.1002/9781118269930.ch4>

AQ:

AQ: 11

- Aboud, F. E., Mendelson, M., & Purdy, K. T. (2003). Cross-race peer relations and friendship quality. *International Journal of Behavioral Development, 27*(2), 165–173. <https://doi.org/10.1080/01650250244000164>
- Al Ramiah, A., & Hewstone, M. (2013). Intergroup contact as a tool for reducing, resolving, and preventing intergroup conflict: Evidence, limitations, and potential. *American Psychologist, 68*(7), 527–542. <https://doi.org/10.1037/a0032603>
- Apfelbaum, E. P., Pauker, K., Sommers, S. R., & Ambady, N. (2010). In blind pursuit of racial equality? *Psychological Science, 21*(11), 1587–1592. <https://doi.org/10.1177/0956797610384741>
- Apfelbaum, E. P., Sommers, S. R., & Norton, M. I. (2008). Seeing race and seeming racist? Evaluating strategic colorblindness in social interaction. *Journal of Personality and Social Psychology, 95*(4), 918–932. <https://doi.org/10.1037/a0011990>
- Augoustinos, M., & Rosewarne, D. L. (2001). Stereotype knowledge and prejudice in children. *British Journal of Developmental Psychology, 19*(1), 143–156. <https://doi.org/10.1348/026151001165912>
- Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology, 87*(1), 5–22. <https://doi.org/10.1037/0022-3514.87.1.5>
- 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>
- Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. *Advances in Child Development and Behavior, 34*, 39–89. [https://doi.org/10.1016/S0065-2407\(06\)80004-2](https://doi.org/10.1016/S0065-2407(06)80004-2)
- Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory. *Current Directions in Psychological Science, 16*(3), 162–166. <https://doi.org/10.1111/j.1467-8721.2007.00496.x>
- Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist, 61*(1), 27–41. <https://doi.org/10.1037/0003-066X.61.1.27>
- Bluemke, M., & Friese, M. (2006). Do features of stimuli influence IAT effects? *Journal of Experimental Social Psychology, 42*(2), 163–176. <https://doi.org/10.1016/j.jesp.2005.03.004>
- Cañadas, E., Lupiáñez, J., Kawakami, K., Niedenthal, P. M., & Rodríguez-Bailón, R. (2016). Perceiving emotions: Cueing social categorization processes and attentional control through facial expressions. *Cognition and Emotion, 30*(6), 1149–1163. <https://doi.org/10.1080/02699931.2015.1052781>
- Corneille, O., & Hütter, M. (2020). Implicit? What do you mean? A comprehensive review of the delusive implicitness construct in attitude research. *Personality and Social Psychology Review, 24*(3), 212–232. <https://doi.org/10.1177/1088868320911325>
- Cvencek, D., Greenwald, A. G., & Meltzoff, A. N. (2011). Measuring implicit attitudes of 4-year-olds: The preschool implicit association test. *Journal of Experimental Child Psychology, 109*(2), 187–200. <https://doi.org/10.1016/j.jecp.2010.11.002>
- De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37*(6), 443–451. <https://doi.org/10.1006/jesp.2000.1464>
- Degner, J., & Calanchini, J. (2020). Age invariance in implicit bias: Alternative perspectives and their implications for the development of implicit cognition. *Social Cognition, 38*(Suppl.), s135–153. <https://doi.org/10.1521/soco.2020.38.suppl.s135>
- Degner, J., & Wentura, D. (2010). Automatic activation of prejudice in children and adolescents. *Journal of Personality and Social Psychology, 98*(3), 356–374. <https://doi.org/10.1037/a0017993>
- Donges, U. S., Kersting, A., & Suslow, T. (2012). Women's greater ability to perceive happy facial emotion automatically: Gender differences in affective priming. *PLoS ONE, 7*(7), e41745. <https://doi.org/10.1371/journal.pone.0041745>
- Dovidio, J. F., Gaertner, S. L., & Validzic, A. (1998). Intergroup bias: Status, differentiation, and a common in-group identity. *Journal of Personality and Social Psychology, 75*(1), 109–120. <https://doi.org/10.1037/0022-3514.75.1.109>
- 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*(5), 1268–1281. <https://doi.org/10.1111/j.1467-8624.2006.00933.x>
- 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., Chen, E. E., & Banaji, M. R. (2013). Two signatures of implicit intergroup attitudes: Developmental invariance and early enculturation. *Psychological Science, 24*(6), 860–868. <https://doi.org/10.1177/0956797612463081>
- Fiske, S. T., Lin, M. H., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 231–254). Guilford Press. <https://doi.org/10.4324/9781315187280>
- Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology, 117*(3), 522–559. <https://doi.org/10.1037/pspa0000160>
- Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118*(2), 247–279. <https://doi.org/10.1037/a0022327>
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692–731. <https://doi.org/10.1037/0033-2909.132.5.692>
- Gonzalez, A. M., Steele, J. R., & Baron, A. S. (2017). Reducing children's implicit racial bias through exposure to positive outgroup exemplars. *Child Development, 88*(1), 123–130. <https://doi.org/10.1111/cdev.12582>
- Gonzalez, A. M., Steele, J. R., Chan, E. F., Lim, S. A., & Baron, A. S. (2021). Developmental differences in the malleability of implicit racial bias following exposure to counterstereotypical exemplars. *Developmental Psychology, 57*(1), 102–113. <https://doi.org/10.1037/dev0001128>
- Govan, C. L., & Williams, K. D. (2004). Reversing or eliminating IAT effects by changing the affective valence of the stimulus items. *Journal of Experimental Social Psychology, 40*(3), 357–365. <https://doi.org/10.1016/j.jesp.2003.07.002>
- Greenwald, A. G., & Banaji, M. R. (2017). The implicit revolution: Reconciling the relation between conscious and unconscious. *American Psychologist, 72*(9), 861–871. <https://doi.org/10.1037/amp0000238>
- 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>
- Hayward, L. E., Tropp, L. R., Hornsey, M. J., & Barlow, F. K. (2018). Negative (but not positive) contact with Whites directly predicts collective action among racial and ethnic minorities. *British Journal of Social Psychology, 57*(1), 1–20. <https://doi.org/10.1111/bjso.12220>
- Holoien, D. S., & Shelton, J. N. (2012). You deplete me: The cognitive costs of colorblindness on ethnic minorities. *Journal of Experimental Social Psychology, 48*(2), 562–565. <https://doi.org/10.1016/j.jesp.2011.09.010>
- Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14*(6), 640–643. https://doi.org/10.1046/j.0956-7976.2003.psci_1478.x

- Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. *Psychological Science*, *15*(5), 342–345. <https://doi.org/10.1111/j.0956-7976.2004.00680.x>
- Hughes, J. M., Bigler, R. S., & Levy, S. R. (2007). Consequences of learning about historical racism among European American and African American children. *Child Development*, *78*(6), 1689–1705. <https://doi.org/10.1111/j.1467-8624.2007.01096.x>
- Jost, J. T. (2019). The IAT is dead, long live the IAT: Context-sensitive measures of implicit attitudes are indispensable to social and political psychology. *Current Directions in Psychological Science*, *28*(1), 10–19. <https://doi.org/10.1177/0963721418797309>
- Karmali, F., Kawakami, K., Vaccarino, E., Williams, A., Phills, C., & Friesen, J. (2019). I don't see race (or conflict): Strategic descriptions of ambiguously negative intergroup contexts. *Journal of Social Issues: Research on Race and Racism*, *75*(4), 1002–1034. <https://doi.org/10.1111/josi.12353>
- Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. *Advances in Experimental Social Psychology*, *55*, 1–80. <https://doi.org/10.1016/bs.aesp.2016.10.001>
- Kawakami, K., Friesen, J., & Vingilis-Jaremko, L. (2018). Visual attention to members of own and other groups: Preferences, determinants, and consequences. *Social and Personality Psychology Compass*, *12*(312480), 1–16. <https://doi.org/10.1111/spc3.12380>
- Kawakami, K., Hugenberg, K., & Dunham, Y. (2020). Perceiving others as group members: Basic principles of social categorization processes. In P. van Lange, E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (3rd ed., pp. 41–429). Guilford Press.
- Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition and Emotion*, *13*(5), 505–521. <https://doi.org/10.1080/026999399379168>
- Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition*, *27*(4), 623–634. <https://doi.org/10.1521/soco.2009.27.4.623>
- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., . . . 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>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual* (Technical Report A-6). University of Florida.
- Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology*, *69*(3), 397–407. <https://doi.org/10.1037/0022-3514.69.3.397>
- Mitchell, G., & Tetlock, P. (2017). Popularity as a poor proxy for utility: The case of implicit prejudice. In S. O. Lilienfeld & I. D. Waldman (Eds.), *Psychological science under scrutiny: Recent challenges and proposed solutions* (1st ed., pp. 164–195). Wiley. <https://doi.org/10.1002/9781119095910.ch10>
- Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology*, *132*(3), 455–469. <https://doi.org/10.1037/0096-3445.132.3.455>
- Monteiro, M. B., de França, D. X., & Rodrigues, R. (2009). The development of intergroup bias in childhood: How social norms can shape children's racial behaviours. *International Journal of Psychology*, *44*(1), 29–39. <https://doi.org/10.1080/00207590802057910>
- Neto, F., da Conceição Pinto, M., & Mullet, E. (2016). Can music reduce anti-dark-skin prejudice? A test of a cross-cultural musical education programme. *Psychology of Music*, *44*(3), 388–398. <https://doi.org/10.1177/0305735614568882>
- 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>
- Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences*, *15*(4), 152–159. <https://doi.org/10.1016/j.tics.2011.01.005>
- Page-Gould, E., Mendoza-Denton, R., Alegre, J. M., & Siy, J. O. (2010). Understanding the impact of cross-group friendship on interactions with novel outgroup members. *Journal of Personality and Social Psychology*, *98*(5), 775–793. <https://doi.org/10.1037/a0017880>
- Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, *72*(1), 533–560. <https://doi.org/10.1146/annurev-psych-071620-030619>
- Pauker, K., Williams, A., & Steele, J. R. (2016). Children's racial categorization in context. *Child Development Perspectives*, *10*(1), 33–38. <https://doi.org/10.1111/cdep.12155>
- Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*(3), 277–293. <https://doi.org/10.1037/0022-3514.89.3.277>
- Payne, B. K., Vuletic, H. A., & Lundberg, K. B. (2017). Flipping the script on implicit bias research with the bias of crowds. *Psychological Inquiry*, *28*(4), 306–311. <https://doi.org/10.1080/1047840X.2017.1380460>
- Perszyk, D. R., Lei, R. F., Bodenhausen, G. V., Richeson, J. A., & Waxman, S. R. (2019). Bias at the intersection of race and gender: Evidence from preschool-aged children. *Developmental Science*, *22*(3), e12788 <https://doi.org/10.1111/desc.12788>
- Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*, 65–85. <https://doi.org/10.1146/annurev.psych.49.1.65>
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. <https://doi.org/10.1037/0022-3514.90.5.751>
- Plaut, V. C., Thomas, K. M., & Goren, M. J. (2009). Is multiculturalism or color blindness better for minorities? *Psychological Science*, *20*(4), 444–446. <https://doi.org/10.1111/j.1467-9280.2009.02318.x>
- Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Dittmann, R., & Crosby, J. R. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology*, *94*(4), 615–630. <https://doi.org/10.1037/0022-3514.94.4.615>
- Qian, M. K., Quinn, P. C., Heyman, G. D., Pascalis, O., Fu, G., & Lee, K. (2019). A long-lasting effect of perceptual individuation training on reducing implicit racial bias in preschool children. *Child Development*, *90*(3), e290–e305. <https://doi.org/10.1111/cdev.12971>
- Quinn, P. C., Lee, K., Pascalis, O., & Xiao, N. G. (2020). Emotional expressions reinstate recognition of other-race faces in infants following perceptual narrowing. *Developmental Psychology*, *56*(1), 15–27. <https://doi.org/10.1037/dev0000858>
- Raabe, T., & Beelmann, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences. *Child Development*, *82*(6), 1715–1737. <https://doi.org/10.1111/j.1467-8624.2011.01668.x>
- Raissi, A., & Steele, J. R. (in press). Does emotional expression moderate implicit racial bias? Examining bias following smiling and angry primes. *Social Cognition*.
- Rhodes, M., & Baron, A. (2019). The development of social categorization. *Annual Review of Developmental Psychology*, *1*, 359–386. <https://doi.org/10.1146/annurev-devpsych-121318-084824>

- Richeson, J. A., & Nussbaum, R. J. (2004). The impact of multiculturalism versus color-blindness on racial bias. *Journal of Experimental Social Psychology, 40*(3), 417–423. <https://doi.org/10.1016/j.jesp.2003.09.002>
- Roberts, S. O., & Gelman, S. A. (2016). Can White children grow up to be Black? Children's reasoning about the stability of emotion and race. *Developmental Psychology, 52*(6), 887–893. <https://doi.org/10.1037/dev0000132>
- Rohr, M., Degner, J., & Wentura, D. (2012). Masked emotional priming beyond global valence activations. *Cognition and Emotion, 26*(2), 224–244. <https://doi.org/10.1080/02699931.2011.576852>
- 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>
- Schultz, D., Ambike, A., Buckingham-Howes, S., & Cheah, C. S. L. (2008). Experimental analysis of preschool playmate preferences as a function of smiles and sex. *Infant and Child Development, 17*(5), 503–507. <https://doi.org/10.1002/icd.566>
- Shutts, K. (2015). Young children's preferences: Gender, race, and social status. *Child Development Perspectives, 9*(4), 262–266. <https://doi.org/10.1111/cdep.12154>
- Shutts, K., Banaji, M. R., & Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Developmental Science, 13*(4), 599–610. <https://doi.org/10.1111/j.1467-7687.2009.00913.x>
- Shutts, K., Pemberton, C. K., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development, 14*(1), 35–62. <https://doi.org/10.1080/15248372.2011.638686>
- Sinclair, L., & Kunda, Z. (1999). Reactions to a black professional: Motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology, 77*(5), 885–904. <https://doi.org/10.1037/0022-3514.77.5.885>
- Steele, J. R., George, M., Cease, M., Fabri, T. L., & Schlosser, J. (2018). Not always Black and White: The effect of race and emotional expression on implicit attitudes. *Social Cognition, 36*(5), 534–558. <https://doi.org/10.1521/soco.2018.36.5.534>
- 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), e12673. <https://doi.org/10.1111/desc.12673>
- Steffens, M. C., & Plewe, I. (2001). Items' cross-category associations as a confounding factor in the Implicit Association Test. *Zeitschrift Für Experimentelle Psychologie, 48*(2), 123–134. <https://doi.org/10.1026//0949-3946.48.2.123>
- Turner, R. N., Hewstone, M., Voci, A., & Vonofakou, C. (2008). A test of the extended intergroup contact hypothesis: The mediating role of intergroup anxiety, perceived ingroup and outgroup norms, and inclusion of the outgroup in the self. *Journal of Personality and Social Psychology, 95*(4), 843–860. <https://doi.org/10.1037/a0011434>
- Vezzali, L., Capozza, D., Giovannini, D., & Stathi, S. (2012). Improving implicit and explicit intergroup attitudes using imagined contact: An experimental intervention with elementary school children. *Group Processes & Intergroup Relations, 15*(2), 203–212. <https://doi.org/10.1177/1368430211424920>
- Vorauer, J. D., & Sasaki, S. J. (2009). Helpful only in the abstract? *Psychological Science, 20*(2), 191–197. <https://doi.org/10.1111/j.1467-9280.2009.02265.x>
- Weisman, K., Johnson, M. V., & Shutts, K. (2015). Young children's automatic encoding of social categories. *Developmental Science, 18*(6), 1036–1043. <https://doi.org/10.1111/desc.12269>
- Williams, A., & Steele, J. R. (2019). Examining children's implicit racial attitudes using exemplar and category-based measures. *Child Development, 90*(3), e322–e338. <https://doi.org/10.1111/cdev.12991>
- Williams, A., Steele, J. R., & Lipman, C. (2016). Assessing children's implicit attitudes using the affect misattribution procedure. *Journal of Cognition and Development, 17*(3), 505–525. <https://doi.org/10.1080/15248372.2015.1061527>
- Wilton, L. S., Apfelbaum, E. P., & Good, J. J. (2019). Valuing differences and reinforcing them: Multiculturalism increases race essentialism. *Social Psychological & Personality Science, 10*(5), 681–689. <https://doi.org/10.1177/1948550618780728>
- Xiao, W. S., Fu, G., Quinn, P. C., Qin, J., Tanaka, J. W., Pascalis, O., & Lee, K. (2015). Individuation training with other-race faces reduces preschoolers' implicit racial bias: A link between perceptual and social representation of faces in children. *Developmental Science, 18*(4), 655–663. <https://doi.org/10.1111/desc.12241>

(Appendix follows)

Appendix
Sample Headers for the Child IATs

Figure A1

Race-Categorization Condition (Studies 1 and 2)



Emotion-Categorization Condition (Studies 1 and 2)



Ambiguous-Categorization Condition (Studies 2 and 3)



Note. The side and pairing order for each header in the critical blocks were counterbalanced between participants. Children were trained to use the headers to sort the images. The sun and cloud were used as a pictorial reminder for pleasant and unpleasant in each condition. See the online article for the color version of this figure.

Received May 22, 2020
Revision received March 20, 2021
Accepted March 26, 2021 ■

Supplemental Materials

Additional Analyses. As noted in Footnote 3, there has been debate as to whether a comparison to 0 on the IAT provides a meaningful, or arbitrary, metric (e.g., Blanton & Jaccard, 2006; Mitchell & Tetlock, 2017, cf. Jost, 2019). We outlined these scores for each study below, however, due to the potential for 0 on the IAT to provide an arbitrary metric, caution should be used in the interpretation of these scores.

In Study 1, to further examine these associations, one-sample *t*-tests comparing mean *D* scores to 0 were conducted separately for participants in each condition. Children in the Race-Categorization condition were faster to pair neutral White faces with pleasant images and smiling Black faces with unpleasant images relative to the reverse pairing, $t(62) = 9.68, p < .001, d = 1.22$; and this was true for both younger ($D = .46, SD = .46, t(27) = 5.35, p < .001, d = 1.01$), and older ($D = .59, SD = .42, t(34) = 8.34, p < .001, d = 1.41$), children. By contrast, children in the Emotion-Categorization condition showed the reverse pattern. They were faster to pair smiling Black faces with pleasant images and neutral White faces with unpleasant images relative to the reverse pairing, $t(54) = -12.91, p < .001, d = -1.74$; and this was true for both younger, ($D = -.62, SD = .42, t(24) = -7.39, p < .001, d = -1.48$), and older ($D = -.75, SD = .37, t(29) = -10.99, p < .001, d = -2.03$), children.

Consistent with the results from Study 1, in Study 2, one-sample *t*-tests comparing mean *D* scores to 0 revealed that children in the Race-Categorization condition were faster to pair neutral White faces with pleasant images and smiling Black faces with unpleasant images relative to the reverse pairing, $t(32) = 6.86, p < .001, d = 1.19$. By contrast, and also consistent with Study 1, children in the Emotion-Categorization condition were faster to pair smiling Black faces with pleasant images and neutral White faces with unpleasant images relative to the reverse pairing, $t(27) = -5.26, p < .001, d = -.98$. Finally, participants in the

Ambiguous-Categorization condition, who were not given explicit instructions about how to categorize these targets, demonstrated no reliable bias, $t(36) = .82, p = .42, d = .14$.

In Study 3, one-sample t -tests comparing mean D scores to 0 revealed that children in the Blatant Intervention condition showed more positive associations with smiling Black targets relative to neutral White targets, $t(52) = -11.22, p < .001, d = -1.64$. By contrast, children in the Subtle Intervention condition did not demonstrate a preference on this measure, $t(54) = -.58, p = .57, d = -.08$. In contrast to the results of Study 2, where children demonstrated no bias on an Ambiguous-Categorization IAT, children in the Control condition showed more positive associations with neutral White targets relative to smiling Black targets, $t(59) = 2.59, p = .01, d = -.35$.

Additional Control Conditions. As noted in Footnote 4, Studies 1 and 2 included an additional control condition. As a secondary goal in Study 1, we examined whether the magnitude of children's biases would be influenced by the emotional expression of target faces. To address this question, in Study 1, an additional 55 children were recruited and were randomly assigned to a third condition, the Race-Control condition. Children in this control condition saw the same headers provided in the Race-Categorization condition but were instead asked to categorize Black and White targets with identical emotional expressions (smiling) whose pictures had been cropped at the mouths (showing only the top half of the face). One of these child participants chose not to complete the entire IAT and the data of five children were removed from analyses due to a response error rate of 35% or greater (Cvencek et al., 2011), resulting in 49 children in this condition. A 3(Condition: Race-Categorization, Emotion-Categorization, Race-Control) X 2(Age: Younger or Older) ANOVA revealed only a main effect of condition, $F(2, 161) = 152.72, p < .001, \eta_p^2 = .66$. Simple effects tests revealed that the Race-Control condition ($D = .50, SD = .38$) did not differ from the Race-Categorization ($D = .54, SD = .44$), $p = .89$, suggesting that when

children were categorizing by race, the emotional expression of the faces did not affect the speed of their associations. As expected, both the Race-Control and Race-Categorization conditions differed from the Emotion-Categorization ($D = -.69$, $SD = .39$) condition, $ps < .001$. No main effect of age $F(1, 161) = .68$, $p = .41$, $\eta_p^2 = .004$ or interaction $F(2, 161) = 1.98$, $p = .14$, $\eta_p^2 = .14$, emerged.

In Study 2, we similarly sought to determine whether children's associations with smiling targets would be influenced by the race of the target faces. To address this question, an additional 34 children were recruited and were randomly assigned to a fourth condition, an Emotion-Control condition. Children in this control condition saw similar headers to the Emotion-Categorization condition, however in this condition smiling Black faces were replaced with smiling White faces. Thus, children in the Emotion-Control condition were asked to categorize smiling White faces and neutral White faces as well as pleasant and unpleasant images; whereas children in the Emotion-Categorization condition categorized faces that also differed systematically by race, specifically, smiling Black and neutral White faces. In a seemingly counterintuitive way, designed to match the scoring for the other conditions, this implicit measure was scored such that higher scores indicate faster responding to neutral faces paired with pleasant images and smiling faces paired with unpleasant images relative to the reverse pairing. The data were removed for two children: One had a response error rate of 35% or greater and the other was a numerical outlier (Cvencek et al., 2011). A 4(Condition: Race-Categorization, Emotion-Categorization, Ambiguous-Categorization, Emotion-Control) X 2(Age: Younger or Older) ANOVA revealed only a main effect of condition, $F(3, 121) = 25.52$, $p < .001$, $\eta_p^2 = .39$. Simple effects tests revealed that the magnitude of associations in the Emotion-Control ($D = -.54$, $SD = .53$) condition did not differ from the Emotion-Categorization ($D = -.52$, $SD = .53$) condition, $p > .99$. When children were categorizing by emotional expression the race of the

target did not affect the speed of their associations. As expected, the Emotion-Control also differed from both the Race-Categorization condition ($D = .44, SD = .37$) and Ambiguous-Categorization conditions ($D = .09, SD = .63$), $ps < .001$. No main effect of age $F(1, 121) = .009, p = .92, \eta_p^2 < .001$, or interaction $F(3, 121) = .71, p = .55, \eta_p^2 = .02$ emerged.