ABSTRACT—The ability to discriminate visually based on race emerges early in infancy: 3-month-olds can perceptually differentiate and 6-month-olds can perceptually categorize faces by race. Between ages 6 and 8 years, children can sort others into racial groups. But to what extent are these abilities influenced by context? In this article, we review studies on children’s racial categorization and discuss how our conclusions are affected by how we ask the questions (i.e., our methods and stimuli), where we ask them (i.e., the diversity of the child’s surrounding environment), and whom we ask (i.e., the diversity of the children we study). Taken together, we suggest that despite a developmental readiness to categorize others by race, the use of race as a psychologically salient basis for categorization is far from inevitable and is shaped largely by the experimental setting and the larger cultural context.

KEYWORDS—racial categorization; racial stereotyping and prejudice; social development

Racial prejudice is one of the most pressing social issues of our time. Social and developmental psychologists have sought to understand more deeply when racial biases emerge in childhood. Despite the foundational role of racial categorization in stereotyping and prejudice, research with children has focused almost exclusively on the downstream consequences of racial categorization rather than the process of racial categorization itself. In this article, we review what is known about racial categorization from infancy into late childhood, with a focus on recent research. In addition, we argue that researchers need to devote greater attention to the experimental setting and the larger cultural context to advance our theoretical and practical understanding of the development of racial categorization.

WHEN CAN CHILDREN CATEGORIZE BY RACE?

The answer to this question depends largely on how categorization is defined. For example, does noticing differences between racial groups, sorting targets with similar skin color together, identifying physical features as typical of group members, or labeling members of different racial groups provide sufficient evidence of racial categorization? In this article, we define racial categorization as the tendency for race to be perceived as a psychologically salient and meaningful basis for grouping others. This definition builds on the developmental intergroup theory (DIT; 1), in which four main factors contribute to the psychological salience of social categories: (a) perceptual salience (i.e., whether categories are marked by discriminable visual features), (b) proportional group size (i.e., proportionally smaller groups, or minorities, tend to be more distinct), (c) explicit labeling by adults (e.g., “the Black child”), which suggests the dimension merits attention and provides a category label, and (d) implicit use in the environment (e.g., through racial segregation of neighborhoods), which may lead children to independently construct explanations regarding the importance of shared attributes (1). Measuring racial categorization involves administering tasks that map onto these factors, and exploring how and when children consistently and spontaneously use the category to organize information and direct behavior. This definition of racial categorization highlights not only how many inputs (both perceptual and conceptual) integrate to inform children’s categorizations, but also how context directs whether race is salient psychologically and thus used habitually in a psychologically meaningful way. Although outside the scope of this article, one important conceptual input into children’s categorizations is their intuitive theories, including beliefs that social categories are natural
kinds (2). Yet even these intuitive theories may be shaped by cultural context (2–4). Although some factors contributing to the psychological salience of race can emerge quite early in infancy (e.g., perceptual discrimination) and other components depend more on linguistic skills that develop later in childhood (e.g., labeling by race), all are influenced by both the immediate (experimental) and broader (cultural) context.

Infants

Although infants are not attuned to racial differences at birth (5), their ability to differentiate perceptually based on race develops early in homogeneous cultural contexts. By 3 months, White, Black, and Asian infants from countries where their race is the majority (i.e., White infants in the United Kingdom, Black infants in Ethiopia, and Asian infants in China) look longer at same-race faces than at other-race faces (5–7). However, despite this visual preference for same-race faces, young infants do not show impaired recognition of other-race faces that is typically seen in adults (8). Instead, at 3 months, White and Asian infants from countries where their race is the majority can recognize different faces of their race as well as different faces of other races (9, 10). These infants demonstrate a decreasing ability to differentiate other-race faces across many out-groups between 3 and 9 months, and by 9 months, they recognize same-race faces but have difficulty recognizing other-race faces (9, 10), similar to the impaired ability to recognize other-race faces seen in adults (8).

Thus, while 3-month-olds raised in homogeneous cultural contexts show sensitivity to distinctions between racial groups, they can still individuate faces within racial groups. However, the ability to individuate within racial groups apparently changes with development and environmental input—and children become tuned to the faces they encounter most frequently as they age. Consistent with the strong connection in adults between categorical processing of race and impaired recognition of other-race faces (8), this perceptual tuning also apparently coincides with infants’ ability to categorize faces by race (11). Infants can perceptually categorize some faces by race at 6 months (12): Specifically, in one study, when White 6-month-olds with limited exposure to other-race faces were familiarized with many Black or Asian faces (i.e., faces belonging to a single racial category), they distinguished between a new face from the familiarized racial category compared to a new face from a novel racial category (i.e., Asian or Black, respectively; 12). This design tests whether infants categorized a new face from the familiarized category as part of the same category and a face from the novel racial category as part of a different category. However, at 9 months, White infants no longer distinguished between many other-race categories, instead forming a broader distinction between same-race (White = in-group) and other-race faces grouped together (Asian and Black = out-group; 12).

In all the studies with infants we have reviewed, stimuli consisted of color photographs of faces that used both facial features and skin tone as visual markers of race. Thus, we cannot determine whether infants use one or both of these visual cues to process same- and other-race faces. However, in some studies (13), the ability to differentiate same- and other-race faces was not necessarily based solely on low-level perceptual cues such as skin color. When presented with computer-generated faces that depicted prototypical physiognomy and skin tone (i.e., Eurocentric facial features with White skin tone and Afrocentric features with Black skin tone) or faces that isolated these aspects (e.g., Eurocentric features with Black skin tone and Afrocentric features with White skin tone), the neural responses of White majority 9-month-olds in the United States did not differ when viewing prototypical White faces in comparison to faces that isolated Black features (i.e., skin tone or face shape), but did differ in comparison to prototypical Black faces (13). Thus, infants may rely on both facial shape associated with a racial group and skin tone to distinguish from other-race faces.

Do these examples reflect individuals’ ability to perceptually differentiate racial categories or merely to differentiate what is familiar and what is not? As studies often involve comparing familiar and unfamiliar race faces, this effectively assesses whether children can separate their familiar group from a perceptually distinct group (11). To build on this work, researchers should present many groups of unfamiliar other-race faces to further examine infants’ ability to perceptually differentiate and categorize faces based on race (cf. 12).

Although it is unclear whether infants’ abilities to categorize by race reflect more than perceptual differentiation, the central role of cultural context in these effects deserves emphasis. Because biases in visual attention are not present at birth (5), limited exposure to other-race faces may lead to the perceptual narrowing favoring same-race faces. Indeed, in one study, White and Black 3-month-olds in Israel who are exposed frequently to faces from both these racial groups did not look preferentially toward faces of a same-race relative to other-race faces (6). Even minimal exposure to other-race faces in infancy facilitates the ability to recognize other-race faces (14–16). Thus, from a very young age, infants display sensitivity to race that is driven by cultural context, such as the faces they are exposed to in their environment.

Toddlers

Recent studies raise questions about the extent to which young toddlers readily use perceptual cues to categorize new racial group exemplars, even if they appear to do so as 6-month-olds. In one study (17), 19-month-old Jewish Israeli toddlers failed to match new exemplars to a category of exemplars they had just been familiarized with, including those high in perceptual (e.g., gender, race, shirt color) and cultural (e.g., ethnicity) salience, unless the category exemplars were paired with a novel category label (e.g., “Look, a Tiroli”) during familiarization. In contrast, 26-month-olds matched new race and gender exemplars with
the expected category (i.e., selecting a Black target after being familiarized with color photographs of Black people), regardless of whether category exemplars were paired with a novel category label. Thus, younger toddlers’ representation of racial categories apparently relies on cultural input (e.g., category labels) rather than emerging solely based on visual cues.

Does being able to perceptually differentiate racial categories correspond with viewing race as a meaningful, psychologically salient category that guides behavior (1)? Early in development it does not, because in infancy, looking preferences are unrelated to social behavior. At 10 months, when infants in homogeneous cultural contexts robustly recognize same-race compared to other-race faces, White American infants do not prefer toys offered by video-recorded White women over those offered by video-recorded Black women (18). Even older toddlers fail to demonstrate race-based differences in behavior: White American 2- to 3-year-olds are equally likely to give toys to White or Black women depicted in color photographs (18). Furthermore, when the experimental context places social categories in competition, children may prioritize categories other than race and these may predict behavior (19). When presented simultaneously with color photographs of children or adults that vary systematically by gender and race, White American 3- to 4-year-olds’ friendship selections, inferences about shared preferences, allocation and acceptance of toys, and preference for novel activities and objects are determined more by gender than race (20, 21).

**Children**

Children may perceptually differentiate racial group members based on similar features. But when provided with category labels, by ages 3 or 4, White Canadian children can identify the racial group membership of targets depicted in color photographs (in accordance with adult judgments; 22), and by ages 6–8, both Black and White children can consistently classify others by race (23). However, in studies of target groups other than Blacks and Whites, race is not as psychologically salient. For example, when asked to sort color photographs of children by racial label (White, Black, Asian), only a slim majority (60%) of White, Black, and Asian 3- to 5-year-olds from multiracial schools in the United Kingdom used the terms in a manner consistent with adult categorizations (24). Additionally, when studies included a wider range of stimuli, such as computer-generated faces that varied in their prototypicality (in both skin tone and physiognomy), predominantly White American 4- to 9-year-olds relied more on skin color than physiognomy when categorizing by race (25; see also 20). Children did not use facial features as category-diagnostic information in the same way as adults do, suggesting that children may not have an adult-like conceptualization of race. These results raise the possibility that past findings may depend primarily on children’s directed attention to category labels and skin color.

**LOOKING FORWARD: BRINGING CONTEXT INTO FOCUS**

Although we know much about when children can categorize by race, we do not know a great deal about when they do so spontaneously and what factors affect these categorizations. Furthermore, how much of our conclusion—that race is perceptually discernible by 3 months and explicitly identifiable around 6 years—is based on the stability or homogeneity of the tasks, groups, or environments in studies? In other words, are the conclusions about the development of racial categorization biased by the experimental and cultural contexts in which researchers have asked these questions? We believe they may be.

As an illustration, we used an open-ended measure to capture how 8- to 12-year-olds in the continental United States and Hawaii categorized prototypical White and Black target children, depicted in color photographs, by race (27). While White, Asian, and Latino monoracial and multiracial children in the continental United States typically listed one racial label per target, consistent with adult categorizations (e.g., they labeled the Black target as African American), in Hawaii, White, Asian, and Black monoracial and multiracial children tended to perceive the monoracial targets as multiracial or belonging to many groups. Both White and Black targets were described on average by 3–4 racial/ethnic labels (e.g., labeling the Black target as Black, Chinese, and Native Hawaiian). Perhaps because of their experience with a large multiracial population (23% of Hawaii residents identify as multiracial), children growing up in Hawaii may default to a multiracial prototype and be less likely to rely on perceptual cues to categorize racially because they are less predictive in this environment. This example illustrates how expanding our methods (e.g., moving beyond forced choice or labels provided by the experimenter) and highlighting where research is conducted (e.g., a heterogeneous, highly multiracial environment) can provide new insights into racial categorization.

Although such less structured tasks are not without limits (e.g., reliance on children’s verbal abilities, difficulties in scoring responses), results from these measures can clarify how we interpret responses on more structured tasks that assess children’s racial categorization and ensuing attitudes. Researchers should look carefully at how experimental and cultural contexts affect our understanding of racial categorization across development. Specifically, we need to consider how we ask the questions (i.e., our methods and stimuli), where we ask them (i.e., the diversity of the child’s surrounding environment), and whom we ask (i.e., the diversity of the groups we study).

**Methods and Stimuli**

Many of the tasks used to examine racial categorization inadvertently increase the salience of race in the experiment by, for example, explicitly using racial labels, using racially prototypical targets, or making comparisons that differ only by race and not by other competing social categories (e.g., gender, age). In
open-ended spontaneous description tasks (e.g., a child sees a target and is prompted, “Tell me about this person; what do you see?”), White, Black, and Asian preschool and elementary school children in monoracial and multiracial cultures mention race rarely (24, 28, 29). However, when children are asked to sort photos that vary by dimensions (e.g., race, gender, facial expression, age, clothing) into piles that “go together,” children’s use of race as a spontaneous sorting dimension increases with age (24, 30), becoming more reliable around 6 years (30). How racial categorization is assessed can therefore lead to differing conclusions about the extent to which children spontaneously categorize others by race.

Attending to whether the experimental context makes race psychologically salient does not inherently value unstructured over structured tasks. Rather, it should help us expand our repertoire of experimental tasks, interpret more effectively results that vary across experimental context, and provide further insight into the conditions under which others will be spontaneously or deliberately categorized by race. For example, attention to experimental context may affect the interpretation of valuable, highly structured measures, such as those that assess children’s implicit racial biases. In tasks where targets are categorized by race (i.e., the Implicit Association Test), White American participants display an implicit pro-White (relative to Black) bias at 6 years that remains stable into adulthood (31). But measures that do not require overt racial categorization (i.e., the Affective Priming Task) yield a different developmental trajectory: Among White German 9- to 15-year-olds, implicit bias (in the form of out-group negativity) emerged only in early adolescence (32; see also 33). Thus, even among implicit measures, racial salience in the experimental context may affect researchers’ conclusions. Experimental contexts that increase the salience of racial categories may overestimate the extent to which children use race spontaneously when perceiving other people.

Similarly, the focus on prototypical exemplars of various racial groups may artificially heighten children’s attention to race. Not only does this drastically oversimplify the task children face when they meet a new person, but also the representation of stimuli in most experiments reduces within-race variation and underestimates the dynamic nature of how we perceive other people (34). We must broaden the range of stimuli used to include racially ambiguous and multiracial targets to deepen our understanding of the categorization process (35–37). Similar to adults, primarily majority (i.e., White American) children are flexible in how they categorize racially ambiguous faces, integrating both visual and top-down category cues (38), or using their intuitive understanding of race as distinct and immutable (i.e., essentialist reasoning) to guide how they process and remember racially ambiguous faces (39). Examining racially ambiguous and multiracial targets can facilitate our understanding of how conceptual knowledge may bias the category judgments of perceptually identical stimuli. Researchers should also examine the extent to which different social categories (e.g., race and gender) intersect to inform perception and social categorization (40). Finally, studies have begun to rely on more implicit measures of spontaneous categorization (33, 41, 42), which is an important area to develop.

Diversity of Cultural Contexts and Populations
As a whole, most research on racial categorization has been conducted in relatively homogenous cultural contexts (often in the United States), primarily with White children. Although we have cited research from several countries (e.g., Canada, China, Ethiopia, Israel, the United Kingdom, the United States), researchers must examine both racially homogeneous and heterogeneous cultural contexts and groups. We need to include more racial-minority children in this work, including multiracial children who have been almost entirely excluded (cf. 4, 43). In studies that explicitly examined more heterogeneous cultural contexts, where children have exposure to people from a variety of racial groups, diversity can allow children to maintain greater flexibility in components of racial categorization. For example, in one study, infants with intensive cross-race experience did not look preferentially toward same-race faces (6), and in another study, older children in a more diverse city were less likely than children in a rural community to view race as a natural kind (44). In addition, even within the same cultural context, children from a minority group (e.g., Black) may categorize others by race more readily (24, 45), and integrate perceptual and conceptual knowledge about race earlier to inform category judgments (36).

CONCLUSION
In this article, we reviewed studies on racial categorization in childhood and put their findings in context by highlighting that how, where, and to whom we ask our research questions can influence our conclusions. While race is perceptually discriminable early in infancy and used spontaneously by children as young as 6 years to sort others, racial categorization depends on the immediate (experimental) and broader (cultural) context. To deepen our knowledge of the conditions under which children consistently and spontaneously categorize others by race, we must deepen our understanding of how context can influence the cues that children attend to when categorizing others.

REFERENCES


