

The Roots of Racial Categorization

Ben Phillips

bsphilli@asu.edu

Forthcoming in *Review of Philosophy and Psychology*

Abstract. I examine the origins of ordinary racial thinking. In doing so, I argue against the thesis that it is the byproduct of a unique module (e.g. a folk-biology module). Instead, I defend a pluralistic thesis according to which different forms of racial thinking are driven by distinct mechanisms, each with their own etiology. I begin with the belief that visible features are diagnostic of race. I argue that the mechanisms responsible for face recognition have an important, albeit delimited, role to play in sustaining this belief. I then argue that essentialist beliefs about race are driven by some of the mechanisms responsible for “entitativity perception”: the tendency to perceive some aggregates of people as more genuine groups than others. Finally, I argue that coalitional thinking about race is driven by a distinctive form of entitativity perception. However, I suggest that more data is needed to determine the prevalence of this form of racial thinking.

1 Introduction

How and why do we categorize people according to race? Given that racial categorization is automatic (see Ito & Tomelleri, 2017) and relatively widespread, one might be tempted to posit a module that was adapted for encoding race. However, ancestral hunter gatherers moved around on foot, meaning that encounters between the bearers of distinct racial markers were probably rare. It is therefore unlikely that we possess a module that was selected for categorizing others according to race (Cosmides et al., 2003). This leaves us with two main hypotheses.

According to one of these hypotheses, racial concepts are cultural inventions (e.g., see Hochman, 2013, 2019; Jackson, 2017; Smedley & Smedley, 2005, 19). According to the other main hypothesis—popular among evolutionary psychologists—racial categorization is the byproduct of a module that was adapted for other purposes (Atran, 1990, 2001; Cosmides et al., 2003; Gil-White, 2001; Hirschfeld 1996; Pietraszewski et al., 2014, 2015). Are the evolutionary psychologists on the right track, or, is racial thinking a purely cultural invention?

In what follows, I argue that ordinary racial thinking is not the byproduct of a single module, such as a folk-biology module. Instead, I defend a pluralistic thesis according to which different ways of thinking about race are driven by distinct mechanisms, each with their own etiology. Moreover, I argue that by invoking these mechanisms, we can explain some of the key ways in which racial thinking varies both within and across cultures.

In section 2, I examine the folk belief that races are groups of humans distinguished, in part, by visible features. I argue that the mechanisms responsible for face recognition have an important role to play in sustaining this belief. In section 3, I examine the folk belief that race is an inherited, biological, trait. In doing so, I argue against the view—held by a number of evolutionary psychologists—that both essentialist and inheritance beliefs about race are byproducts of the same module (e.g. a folk-biology module). In section 4, I defend the view that essentialist thinking about race is produced by some of the mechanisms responsible for what social psychologists call “entitativity perception”: the tendency to perceive some aggregates of people as more genuine groups than others. More specifically, I distinguish between two varieties of entitativity perception: “categorical” and “dynamical.” I argue that the categorical entitativity perception drives essentialist thinking about race, whereas, dynamical entitativity perception drives the tendency to think of racial groups as coalitions.

I conclude by discussing potential ways in which this pluralistic approach will advance our understanding of racial prejudice, racial stereotyping, and the ethics of racial categorization.

2 Face recognition and the visual markers of race

One belief that is central to folk thinking about race is that the members of different racial groups are distinguished by bodily features, such as skin color and facial physiognomy (see Glasgow, 2009, 69, 119; Hardimon, 2003; Hirschfeld, 1996, 155). The centrality of this belief is reflected in the fact that some individuals are seen as more typical-looking for a given racial group than others (Maddox, 2004, p. 383). For example, Black men are stereotypically thought to have “dark skin tone, broad noses, full lips, and tightly curled hair” (Maddox & Perry, 2018, p. 58). Research has also shown that when a member of a racial group is seen as highly typical for that group, they are subject to more prejudice and stereotyping than other members. Maddox calls this *racial phenotypicality bias* (Maddox, 2004; see also Maddox & Perry, 2018).

The consensus among biologists and race scholars is that there is no non-arbitrary way to sort people according to the visible markers of race (e.g., see Zack, 2002, p. 43). Nonetheless, the belief that visible features are diagnostic of race is still widespread among the folk. Where does this belief come from? In what follows, I argue that the mechanisms responsible for face recognition have a key role to play.

2.1 Face-space and the other-race effect

Face recognition is one of the most widely studied capacities in cognitive science. One especially robust finding is that people tend not to recognize the faces of other-race individuals as well as the faces of own-race individuals. This is widely referred to as “the other-race effect” (Meissner & Brigham, 2001).

There are various theories about what drives the other-race effect. According to one influential approach, people tend not to individuate other-race faces due to a lack of motivation. Evidence for this approach comes from studies in which increasing people's motivation to individuate other-race faces dampens the other-race effect; as well as studies in which decreasing their motivation enhances it (see Young et al., 2012).

Motivational accounts have an important role to play in explaining various instances of the other-race effect; however, they do not explain how and why racial markers are processed as outgroup cues in the first place (I offer some explanations below in section 4).

Similarly, consider Levin's influential view according to which other-race faces are represented as having a "race-specifying feature," whereas, own-race faces are not (Levin, 2000). According to Levin, the encoding of this race-specifying feature means that relative to own-race faces, other-race faces are costly to process: as a result, fewer resources are directed towards individuating them. Like proponents of the motivational account, Levin does not specify how and why other-race faces are processed in terms of "race-specifying features."

How then do face-processing mechanisms become attuned to racial markers? Fortunately, a unifying model has emerged. Based on the pioneering work of Valentine (1991), it is now widely thought that faces are represented as points in a multi-dimensional face-space. The dimensions of face-space are still not fully understood; however, they are thought to represent attributes such as face shape; skin color; and so on. Through experience, face-space is constructed and updated in a way that optimizes recognition (Valentine et al., 2016, p. 1999). Its dimensions thereby come to reflect the facial features that one typically encounters, the origin being occupied by the average face, which is close to the central tendency on all dimensions. If one's diet of faces is largely

restricted to one's own racial group, the dimensions of face-space become optimized for recognizing own-race faces, but not other-race faces.

Developmental studies provide compelling support for the face-space model. By about 9 months of age, infants raised in primarily monoracial environments start to exhibit the other-race effect (see Kelly et al., 2007). Importantly, this developmental trajectory appears to be cross-culturally robust (Kelly et al., 2007; Kobayashi et al., 2016; Suhrke et al., 2014; Tham et al., 2017). There is also strong evidence that the mechanisms driving it are phylogenetically ancient. For example, Sugita (2008) tested infant Japanese macaques on their ability to recognize either human or macaque faces. Those macaques that had only been exposed to human faces had difficulty distinguishing between macaque faces; whereas, those that had only been exposed to macaque faces had difficulty distinguishing between human faces. Similar effects have been found in other primates, including New World monkeys, Old World monkeys, and great apes (see Dufour, Pascalis, & Petit, 2006; Pascalis & Bachevalier, 1998).

There is thus strong evidence for a phylogenetically ancient face-processing mechanism that is sensitive to the visual markers of race. This mechanism causes those raised in primarily monoracial settings to perceptually discriminate between own- and other-race individuals. Importantly, though, distinguishing between own- and other-race individuals is not the function of this mechanism: its function is to optimize face recognition. This goes some way towards explaining why the visible markers of race become salient to children. However, it does not fully explain how and why children acquire representations of race-specific categories, such as *Black*, *White*, and *Asian*. Of course, many children will acquire race-specific concepts from those who already possess them, such as their parents. But given that we are trying to unearth the origins of

racial thinking, this is obviously not a satisfying explanation: it does not tell us how these other individuals acquired the relevant concepts.

2.2 The recognition-byproduct hypothesis

Some theorists have suggested that race-specific categorization is, in fact, a byproduct of face recognition (see Kramer et al., 2017). To see why this might be the case, consider the fact that when it comes to recognizing others, those features that are relatively invariant across viewing conditions serve as the most reliable cues: examples include sex; age; height; and those visible features that are commonly thought to indicate one's race. In contrast, those features that are relatively fleeting do not serve as reliable cues: examples include facial expressions of emotion, the orientation of one's arms, and so on. This is reflected in witness testimony, as well as police descriptions of wanted suspects, which tend to describe individuals in terms of the invariant features listed above.¹

This observation about the role that invariant cues play in recognizing other individuals suggests a simple hypothesis: face-processing mechanisms sort people according to the visual markers of race because this optimizes recognition. Call this “the recognition-byproduct hypothesis.”

The recognition-byproduct hypothesis is supported by computer simulations of face recognition (see Balas, 2012; Dahl et al., 2016; Kramer et al., 2017; O'Toole et al., 1991). For example, Kramer and colleagues (2017) trained a model to recognize faces across a range of viewing conditions. Their model utilized Linear Discriminant Analysis, which is widely used in face classification systems because it is an efficient way of establishing which features best

¹ For example, see <https://home.chicagopolice.org/about/contact-us/how-to-describe-a-suspect/>

describe a class of stimuli (e.g. different images of the same face), thereby discriminating it from members of other classes (e.g. images of other faces).

Initially, Kramer and colleagues (2017) trained their model on 20 faces. Once the model had optimized its ability to recognize these faces (achieving 98% accuracy on novel images of them), new faces were fed into it. What Kramer and colleagues found is that the model immediately started grouping novel faces according to both race and sex. What's more, it did so very "accurately," correctly categorizing 91% of new faces according to race.²

According to Kramer and colleagues, this shows that recognition is optimized when faces are categorized according to those visual markers that are widely thought to indicate race and sex. Given that the model was not explicitly trained to categorize faces according to race and sex, they go even further, suggesting that an understanding of the social significance of these categories "may be largely irrelevant to acquiring the ability to classify faces along these lines" (2017, 11).³

The recognition-byproduct hypothesis also explains "race-contingent aftereffects." Adaptational aftereffects occur when the firing rate of a sensory neuron begins to drop amidst prolonged exposure to its "preferred" stimulus. This reduction in firing rate has the dual effect of decreasing the neuron's sensitivity to the adapted stimulus, while increasing its sensitivity to

² Kramer and colleagues' (2017) use of the term "accurately" needs to be qualified, for if racial categories are not real, there is nothing for the model to be accurate about. Rather than presupposing realism, we can interpret Kramer and colleagues' statement as follows: the model's categorization of an individual as Black was "accurate" only if that individual would be considered Black by the folk; and the model's categorization of an individual as White was "accurate" only if that individual would be considered White by the folk. This interpretation fits with Kramer and colleagues' methodology, for the faces they fed into the model were found on Google Images, with search terms, such as "Black women," and "White men."

³ Below, in section 2.3, I defend a somewhat weaker claim.

novel stimuli. For example, if you visually adapt to lines that are tilted slightly to the left of the vertical axis, and then view some vertical lines immediately afterwards, the latter will look as though they are tilted slightly to the right. Given that there is scant evidence for adaptational aftereffects in paradigmatically post-perceptual systems, their presence is widely seen as indicating that the process in question is perceptual (see Block, 2014).

To test whether racial categorization produces adaptational aftereffects, Jaquet and colleagues (2007) presented participants with the faces of White European individuals, as well as the faces of individuals who are Chinese by descent. Importantly, the faces were all distorted: some were expanded, while others were contracted. During *the adaptation phase*, participants spent two minutes observing six of these faces, three of which were Chinese and three of which were White. In half of the trials, the Chinese faces were expanded by 60%, while the White faces were contracted by 60%. In the other trials, this pattern was inverted.

Next, in *the postadaptation rating phase*, participants observed two versions of the same face, side by side on the screen: one version was contracted by 10%, while the other was expanded by 10%. Low-level aftereffects are unlikely to transfer across images that differ significantly in size (Zhao & Chubb, 2001). Thus, to ensure that aftereffects were driven by high level face-processing mechanisms, face-size was significantly varied between the adaptation phase and the postadaptation rating phase.

Having observed the test faces for 2 s, participants were asked to indicate which one looked more “normal.” What Jaquet and colleagues found was that if the Chinese faces presented during the adaptation phase were *expanded* then the participant was more likely to rate the *expanded* test face as normal-looking, but only if it was Chinese. The analogous result was found for White faces. By the same token, if the Chinese faces presented during the adaptation phase were

contracted then the participant was more likely to rate the *contracted* test face as normal-looking, but only if it was Chinese. Again, the analogous result was found for White faces (for replications, see Jaquet et al., 2008; Little et al., 2008; and Rhodes et al., 2010).

According to Jaquet and colleagues (2007), race-contingent aftereffects provide evidence for two main hypotheses. First, faces belonging to different-race individuals are processed by distinct (even if overlapping) neural populations. Second, these distinct populations realize distinct face-spaces, the origin of each one containing the average face *for the given racial category* (Jaquet et al., 2008, 748–51; see also Jaquet et al., 2007; Rhodes et al., 2010). When trying to recognize a face, racial markers activate one of these face-spaces. The face in question is then compared to the face at the origin of this particular space. As Jaquet and colleagues point out (2008, p. 739), if we only possessed a single face-space with a single face at the origin, the aftereffects described above would not have occurred. Instead, the contracted and expanded faces that appeared during the adaptation phase would have cancelled one another out.

Given that sex-contingent aftereffects have been uncovered as well (Little et al., 2005), various theorists have claimed that distinct (even if overlapping) populations of neurons realize sex-specific face-spaces (Baudouin & Gallay, 2006; Jaquet et al., 2008, 738; Little et al., 2005).

The recognition-byproduct hypothesis explains why those visible features that are thought to indicate race and sex activate distinct face-spaces, thereby giving rise to the contingent aftereffects described above. These visible features are invariant cues for face recognition. Thus, to optimize recognition, face-spaces are constructed and updated in a way that exploits these cues. The end result is that individuals are categorized according to the visual markers of race

and sex early on within perception, thereby generating the pattern of aftereffects described above.⁴

2.3 *The delimited role of face recognition*

The recognition-byproduct hypothesis faces three main challenges. The first challenge concerns Hirschfeld's view that young children's representations of race are based primarily on verbal, not visual, cues (1993; 1996, pp. 155–157). Hirschfeld bases his claim on a simple experiment with 3- and 4-year-old children (1993). In the first phase of the experiment, the children were told a story about a child who was seeking the help of four different adults. Each adult was described in terms of race, occupation, and a non-racial feature, such as age or body type. Immediately afterwards, the children were given a free recall test, the results of which demonstrated that they had successfully encoded all three categories.

In another phase of the experiment, the children were shown drawings that depicted a similar story. The children were then given two tasks: to describe the events that occurred in the story; and to perform a memory test. Hirschfeld found that in both cases, children rarely mentioned race. This is in contrast to the first phase of the experiment, in which children successfully recalled information about the race of each adult. On the basis of these findings, Hirschfeld concludes that young children's acquisition of racial concepts are guided by linguistic cues in a top-down manner (1993; 1996, pp. 155–157).

While important, Hirschfeld's findings are compatible with the recognition-byproduct hypothesis. First, it is important to recognize that children *do* sort people according to skin color,

⁴ I have been focusing on *face* recognition, but it is worth pointing out that in recognizing others, people utilize non-facial bodily cues as well (e.g., see Rice et al., 2013). This is compatible with the recognition-byproduct hypothesis.

even in the absence of verbal cues. For example, Shutts and colleagues (2013) found that 4-year-olds use racial markers to predict who will be friends with whom. Importantly, these children were not primed with racial terms. In fact, in one of Hirschfeld's own studies—which I discuss below—children as young as 3 years of age used skin color to make stereotype-consistent predictions of behavior (Hirschfeld et al., 2007). Thus, to claim that young children do not notice differences in skin color, and that they do not believe that skin color is diagnostic of race, would be misleading—a point that Hirschfeld acknowledges (1996, p. 155).

Second, face recognition abilities are still developing in early childhood (see Hills & Lewis, 2018; McKone et al., 2009). Dunham and colleagues (2014) found that while adults' racial categorizations are sensitive to both skin color and other physiognomic cues, children rely much more heavily on skin color. They draw the conclusion that “children begin with a skin color heuristic and only gradually develop an appreciation of other category-diagnostic features” (2014, p. 479; see also Balas et al., 2015). As Dunham and colleagues note (2014, p. 480) this is to be expected, for it is widely recognized that as their ability to recognize faces improves, children gradually increase their reliance on configural cues (Freire & Lee, 2001; Mondloch et al., 2002).

Overall, then, the evidence suggests that “race-specific” face-spaces take time to develop. As their recognition abilities improve, children's sensitivity to the physiognomic markers of race improve as well. This enhances their tendency to see some individuals as more typical-looking for a given racial group than others, which, in turn, enhances their tendency towards certain forms of racial stereotyping and prejudice (Maddox, 2004; Maddox & Perry, 2018). Thus, what emerges is the belief that people can be divided up into socially meaningful categories, based on skin color, and other aspects of facial physiognomy.

To be clear, I'm not claiming that the mechanisms responsible for face recognition are *solely* responsible for the emergence of this belief: other factors, including linguistic cues, have an important role to play (see section 3). Rather, the claim is that this belief is produced and sustained, *in part*, by the mechanisms responsible for face recognition.

The second challenge concerns the claim, widely held by race scholars, that racial categorization is a uniquely modern phenomenon (for a useful overview, see Hochman, 2019).⁵ Face recognition is a phylogenetically ancient capacity, but we don't see clear and abundant examples of racial categorization in ancient texts. Does this undermine the recognition-byproduct hypothesis?

Clearly, this depends on how we state the hypothesis. According to the strong version, the mechanisms responsible for face recognition produce representations of race as a byproduct. This seems to be what Kramer and colleagues are committing themselves to when they suggest that understanding the social significance of race “may be largely irrelevant to acquiring the ability to classify faces along these lines” (2017, p. 11).

If bona fide racial categorization did not occur before the modern (or late medieval) era, this suggests that the strong version of the recognition-byproduct hypothesis is false. However, consider a weaker version according to which the mechanisms responsible for face recognition do not produce bona fide representations of racial categories: rather, they sort people according to the visual markers of race, and thereby *predispose* us towards acquiring the belief that visible features are diagnostic of race. On this version of the hypothesis, for bona fide racial

⁵ According to convention, the modern era began with the arrival of Columbus in “The New World,” in 1492. Hochman (2019) argues that the concept of race actually emerged in late medieval Spain, just before the start of the modern era. Obviously, this is still *well* after our capacity for face recognition first emerged.

categorization to occur, additional factors need to be in place; but once they are in place, face recognition plays a significant role in sustaining the belief that visible features are diagnostic of race.

The third challenge concerns ordinary racial distinctions that are not based on visible differences. Hirschfeld points to several potential examples, such as English discourse about the Irish; characterizations of Italians as non-White in the early twentieth century; and the distinction between Northern Europeans of Protestant descent and Irish and Southern Europeans of Catholic descent (1996, pp. 43–44).

Some philosophers have argued that according to the folk concept of race, racial differences are always marked in terms of visible features (see Glasgow, 2009, Chapter 2; Hardimon, 2002). I won't take a stance on this issue, for even if we concede that the folk *sometimes* draw racial distinctions in the absence of distinguishing bodily features, the point remains that the belief that racial distinctions are visually marked is widespread, and the recognition-byproduct hypothesis provides one explanation of why this is the case.

3 Misfiring modules and biological conceptions of race

Up until this point, I have been focusing on the belief that bodily features are diagnostic of race. Another belief that tends to occupy ordinary thinking is the belief that these features are indicative of hidden traits: traits which are shared by all and only members of the given racial group. Various studies have found that the folk tend to construe these hidden traits as inherited (Hirschfeld, 1993, 1996; Gil-White, 2001, Machery & Faucher, 2020). Where does this biological thinking about race come from?

3.1 Atran's account

Some theorists have claimed that biological thinking about race is the byproduct of a module. According to Atran (2001), we possess a “living-kind module”: a psychological adaptation for distinguishing between species. Atran claims that racial markers, such as skin color, “readily meet minimum perceptual triggering conditions for firing the living-kind module” (2001, 538). In misfiring, the module thereby causes us to perceive the bearers of different racial markers as belonging to different species. Atran does not provide any direct evidence for this hypothesis; however, in a series of studies, he has provided evidence for his hypothesized living-kind module (Atran, 1990; Atran & Medin, 2010).

There are various problems with the claim that biological thinking about race is the byproduct of a living-kind module. From as early as 3 months of age, infants do appear to distinguish between various species (Quinn & Eimas, 1997). By 5 years of age, they think that visually similar animals belong to the same natural kind, and that these kinds have strict boundaries (Rhodes, 2012; Rhodes & Gelman, 2009; Rhodes et al., 2014). In contrast, young children appear to see membership in racial groups as considerably more flexible (Rhodes et al., 2014). Moreover, unlike animal categories, children are much more likely to treat racial categories as having inductive value when they are exposed to certain linguistic cues.

For example, Waxman (2010) showed 4-year-old children a picture of a person (e.g. a black woman), before telling them that this person possessed some trait, such as being good at a game called “zaggit.” Waxman found that unless the pictured individual was introduced as a member of a named category (e.g. “This one is a *Wayshan*”), children did not project the trait in question onto other members of the same racialized group. This squares away with other studies demonstrating that children are more likely to form essentialist beliefs about a social category

when attention is drawn to the category through the use of a generic statement, such as “Hispanics are X” or “African-Americans are Y” (Leslie, 2014; Rhodes et al., 2012). The important point is that without these linguistic cues, young children are not as likely to use racial categories as bases for inductive inferences.

This is not the pattern of findings we would expect were racial categorizations the byproduct of a living-kind module. On the contrary, we would expect young children’s species categorizations to develop in lockstep with their racial categorizations.

3.2 Hirschfeld’s account

Hirschfeld agrees that the biological conception of race is the product of a module; however, he denies that this module is dedicated to detecting species. Instead, he argues that we possess a “human-kind module,” which tracks and essentializes those social groups that are salient within a given society (1996, 2001). Given that, in many societies, people with the same racial markers form relatively salient groups, the human-kind module uses these markers to track and essentialize them.

Hirschfeld supports his account with studies of preschool-aged children. In one study, young children were shown a drawing of an adult and two children (Hirschfeld, 1993). Visual similarities between the adult and each child were varied along three dimensions: uniform, body build, and skin color. Some children were asked to say which child is the adult’s: others were asked to say which one was the adult as a child. In both conditions, children chose skin color over uniform and body build (at above chance levels).

In another study, young children were told a story about a child who was born to parents of a given race, but was immediately adopted by parents of a different race (Hirschfeld, 1995). The children were asked to specify what the baby would look like when it was grown up (they did

this by pointing at depicted individuals). Four-year-olds reliably chose the individual with the same skin color as the birth parents. Hirschfeld's findings have been replicated by studies conducted in various countries, such as Spain (Giménez & Harris, 2010) and Madagascar (Astuti et al., 2004). According to Hirschfeld, these studies provide converging evidence that preschoolers are essentialists about race: they see it as an immutable trait that is transmitted via descent (2001, pp. 114–115).

Hirschfeld's account faces a number of challenges. First, the cross-cultural robustness of a psychological phenomenon does not entail that it is generated by a domain-specific adaptation. Everybody believes that cars are a human invention, but this belief is clearly not the product of some domain-specific adaptation: the fact that cars are human-invented is easily acquired via domain-general learning (on this point, see Hochman, 2013, p. 999; Machery & Faucher, 2017, p. 1164).

To show that the biological conception of race is the byproduct of a domain-specific adaptation, we would need to establish that it is probably not the product of domain-general processes. For instance, suppose we were to acquire evidence that when infants perceive people with the same skin color, they expect them to behave in the same way. Furthermore, suppose infants have not had enough exposure to the visible markers of race for this capacity to have been acquired via domain-general processes. In that case, the evidence would be suggesting that a domain-specific adaptation processes racial categories as biological kinds. In other words, a "poverty of the stimulus argument" would be in the offing (see Chomsky, 1965).

In Hirschfeld's study, the youngest children were 3 years of age. We are therefore left with no reason to rule out domain-general processes. The fact that children inherit skin color from their

biological parents is an easily learnable fact: minimal exposure to different-race families would presumably suffice (see Hochman, 2013, p. 999; Machery & Faucher, 2017, p. 1164).

The second main problem concerns Hirschfeld's claim that young children "essentialize" racial categories. While the studies outlined above provide evidence that young children see skin color as inherited and stable, they do not provide evidence that they see skin color as tied to a causally powerful "essence." To assess the hypothesis that young children construe race in this way, Mandalaywala and colleagues (2019) tested children, aged 5 to 7 years, on a version of Hirschfeld's switched-at-birth task. In addition to predicting the skin color of the grown-up adopted child, participants were asked to predict whether the child would possess the same psycho-behavioral dispositions as the birth parent, or, the adoptive parent. The experimenter did this by telling the participants that the baby is now 6 years of age, and has various characteristics, such as *being smart* and *being good at sports*. For each characteristic, the participant was asked whether the child possesses it because of the birth mother, or, because of the adoptive mother.

Participants indicated that the target child would have the same skin color as the birth mother 62% of the time. However, when the birth mother and the adoptive mother belonged to different racial groups, participants only said that the target child would share psycho-behavioral characteristics with the birth mother 31% of the time. Subsequent analyses confirmed that there was no effect of birth mother on predictions about psycho-behavioral characteristics. Mandalaywala and colleagues (2019) conclude that "by focusing on beliefs about the heritability of skin color, previous studies (e.g., Hirschfeld, 1995) overestimated the extent to which young children hold essentialist views of race" (2019, 447). They go on to say that although young children tend to see skin color as inherited, they expect a person's psycho-behavioral dispositions

to be determined by the environment of upbringing, not their inherited race (see also Rhodes, 2013a).

There *are* studies suggesting that young children see racial categories as having inductive value. For example, in one of Hirschfeld's studies, participants saw a drawing of two individuals: one had dark skin, and one had light skin (Hirschfeld et al., 2007). The participants were then asked to make predictions concerning the individuals' characteristics. For example, in one trial, they were told, "One of them is a friendly girl." They were then asked to indicate which one was it was. Hirschfeld and colleagues found that even 3-year-olds tended to give stereotype consistent answers. For example, they tended to predict that the individual with the light skin tone would be the friendly one (see also Williams & Roberson, 1967).

However, despite the evidence of racial stereotyping in young children, there is still no evidence to suggest that a single module is responsible for both the tendency to essentialize race *and* the tendency to see race as inherited. As was just explained, in the only study of its kind, Mandalaywala and colleagues (2019) found that children's belief that race is inherited did not predict essentialist beliefs about race. Moreover, neither children nor adults see membership in other salient groups, such as teachers or nations, as inherited (Gelman, 2003, p. 304; Machery & Faucher, 2017, pp. 1150–1151). However, children and adults *do* harbor a tendency to essentialize various non-racial groups—a point I elaborate on below. This further undermines the claim that a single module is responsible for both (i) the belief that race is inherited; and (ii) the tendency to essentialize racial groups.

3.3 Mallon's account

Mallon defends an account that is similar to Hirschfeld's. He draws a useful distinction between *broad essentialism* and *lineage essentialism*. To be a broad essentialist about a category, C, is to

accept the following two theses: (1) there are certain underlying natural properties the possession of which constitutes membership in C; and (2) possessing these underlying properties causally explains why certain other properties are typical for members of C (Mallon, 2016, pp. 25–26). To be a lineage essentialist about C is to accept (1) and (2), plus the following theses: (3) the offspring of members of C are also members of C; (4) genuine members of C only come from other genuine members of C (Mallon, 2016, p. 29).

Mallon argues that lineage essentialism about race emerges from a system that is innate, domain-specific, and species-typical. In gathering evidence for this claim, Mallon goes well beyond Hirschfeld’s findings. For example, he provides qualitative historical evidence for lineage essentialism outside the modern European-American context (2016, pp. 33–34). In examining the unique patterns of racial classification among both Brazilians and the Vezo of Madagascar, Mallon also provides a novel account of how inheritance thinking about race is moderated by local cultural conditions (2016, pp. 35–42).

Mallon thus clarifies and extends Hirschfeld’s thesis in important ways. However, in claiming that there is a unique system that is responsible for both broad essentialism about race, and inheritance thinking about race, he faces the problem that was raised above: namely, that even when young children exhibit inheritance thinking about race, this does not predict broad essentialist beliefs about race. In other words, the developmental evidence does not support the claim that the two components of lineage essentialism are produced by one and the same mechanism.

3.4 Gil-White’s account

Gil-White’s account faces the same problem. He argues that folk-biological mechanisms were exapted during the Pleistocene for the purposes of detecting members of other “ethnies”: cultural

groups made up of smaller units called “bands” (see Richerson & Boyd, 2005). As Gil-White (2001, pp. 518–519) points out, ethnies would have displayed some of the key markers of species. First, due to culturally transmitted norms, members of the same ethnie would have exhibited similar behaviors. Second, ethnies would have been largely endogamous. Finally, because members of the same ethnie would have shared the same type of clothing, bodily markings, and language, they would have looked and sounded similar.

Given the similarities between species and ethnies, Gil-White claims that individuals who began to see members of other ethnies as members of other species would have enjoyed a boost in reproductive success. Why? Because differences in norms would have made attempts to coordinate with members of other ethnies costly (in terms of energy and time wasted). This plausibly led to the exaptation of Atran’s living-kind module, the result being that the detection of other ethnies became automatic and stimulus-driven (Gil-White, 2001, p. 519).

Gil-White (2001, p. 534) goes on to suggest that in contemporary contexts, we automatically see members of other races as members of other ethnies, because racial markers, such as skin color, trigger the ethnie-detector (for similar approaches, see Machery & Faucher, 2005; Moya, 2013; Moya & Boyd, 2015; Moya & Henrich, 2016).

In marshalling support for his hypothesis, Gil-White (2001) appeals to a study of two seminomadic pastoralist groups in Mongolia: the Torguuds and the Kazakhs. According to Gil-White’s findings, members of both groups believe that children inherit psycho-behavioral traits from their biological parents: traits that are seen as being essential to membership in the given ethnic group (2001, pp. 523–526).

Gil-White’s account faces two main problems. First, infants do appear to use ethnic markers, such as language, as bases for inductive inferences about group membership. For example,

infants expect speakers of the same language to form affiliations, and to share food preferences (see Liberman et al., 2017). However, infants do not expect individuals with the same racial markers to form affiliations. It is not until about 4 years of age that children draw these sorts of inferences (Shutts et al., 2013). In fact, there is no evidence to suggest that infants engage in racial categorization, let alone see racial groups as ethnically-like.

The second problem concerns Gil-White's claim that the ethnicity-detector produces both the belief that race is inherited, and the belief that members of the same racial group share psychobehavioral traits. Once again, the developmental evidence suggest that these beliefs are not produced by one and the same mechanism. I have already pointed out that, for all we know, children acquire the former belief via domain-general processes. In what follows, I focus on the mechanisms responsible for the latter belief.

4 Entitativity perception and essentialist beliefs about race

One of the core findings in social psychology is that we tend to essentialize and dehumanize the members of outgroups. *Racial* outgroups are no exception. In what follows, I elaborate on this line of reasoning by focusing on the mechanisms responsible for "entitativity perception."

4.1 Two kinds of entitativity perception

Social psychologists use the term "entitative" to denote the quality of being a group. For example, a street gang seems more entitative (or "groupy") than an aggregate of people waiting in line at a coffee shop. Research suggests that there are two kinds of entitativity perception: *categorical* entitativity perception and *dynamical* entitativity perception (Ip et al., 2006; Rutchick et al., 2008). Groups whose members are perceived as highly similar are perceived as high in categorical entitativity; whereas, groups whose members are perceived as engaging in

acts of affiliation and coordinated action (i.e. coalitional groups) are perceived as high in dynamical entitativity.

Empirical support for the distinction comes from studies in which participants are primed to think about the unity of a group in different ways. For instance, Rutchick and colleagues (2008) asked participants to think about a group of bees. In the *categorical* condition, they were shown a taxonomic chart, and were asked to think about each bee as a member of the same biological category. In the *dynamic* condition, they were shown diagrams outlining the different roles that drones, workers, and queens play in a functioning beehive. Participants were then asked to rate the entitativity of various groups, including a Bhutanese social caste; a political party; and a group of students. They were also asked to list the three most important things about each group.

The categorical prime caused participants to focus on the degree of similarity between group members, which drove their impression of entitativity. The dynamic prime caused them to focus on the degree of interaction among group members, which drove their impression of entitativity (see also Haslam et al., 2000; Ip et al., 2006; Spencer-Rodgers et al., 2007).

Other studies have found that entitativity perception—in both its forms—is a key precursor to stereotyping (e.g., see Crawford et al., 2002; Spencer-Rodgers et al., 2007). Given that dynamical entitativity perception involves the detection of coalitional groups, the stereotypes that it fuels tend to be about group goals and achievements (Spencer-Rodgers et al., 2007). In contrast, categorical entitativity perception involves the detection of groups whose members are highly similar (they need not be affiliated with one another). For this reason, it tends to drive

broad essentialism (Haslam et al., 2000; Rothbart & Taylor, 1992; Spencer-Rodgers et al., 2007; Yzerbyt et al., 1997).⁶

4.2 The role of categorical entitativity perception

There are good reasons for thinking that racial essentialism is driven, in large part, by categorical entitativity perception. As was explained above, for people raised in primarily monoracial settings, the members of racial outgroups tend to look more similar to one another than the members of their racial ingroup. We also examined the evidence for “race-specific” face-spaces. These face-spaces enhance the impression that members of the same racial group are physically similar, for same-race faces tend to cluster around the typical face for that group.

Given that the members of racial outgroups tend to be seen as physically similar, it is a short step to categorical entitativity perception, and thus essentialist thinking about race. To test this hypothesis, Dasgupta and colleagues (1999) told participants about two groups of humanoid creatures: “Gs” and “Hs.” Participants were then presented with a series of pictures in which Gs, but not Hs, were present. In some trials, all Gs had the same skin color, while in other trials, their skin color varied. After each trial, participants answered questions about the psychological traits of Gs. Dasgupta and colleagues found that participants were more likely to attribute common psychological traits to Gs when they had the same skin color (for a similar finding, see Ip et al., 2009).

This, and other studies like it, focus on bodily indicators of race, but entitativity perception can be driven by perceived similarities in behavior too (e.g., see Crawford et al., 2002; Fessler & Holbrook, 2016). For example, in Crawford and colleagues’ seminal study, participants read

⁶ In what follows, I use the term “essentialism” to denote broad essentialism (as defined above in section 3.3).

about high and low entitativity groups. Members of the high entitativity group were described as having similar backgrounds, beliefs, and personalities: they were also described as acting in a similar manner “across a variety of situations” (2002, p. 1080). In fact, physical features were not mentioned at all. Nonetheless, Crawford and colleagues found that describing a group as such caused participants to project the observed traits of one member onto others in the group.

This provides us with one explanation of why racial distinctions are sometimes drawn in the absence of distinguishing bodily features: it is because the groups in question are perceived as psychologically and behaviorally homogenous. For example, given the cultural norms and practices shared by Italian-Americans in the early twentieth century, those who racialized them may have done so, in part, because they perceived them as psychologically homogenous, and thus, categorically entitative.⁷

4.2.1 Explaining variation in essentialist thinking

In developing the view that essentialist thinking about race is driven by categorical entitativity perception, an immediate challenge is worth considering: namely, why do a significant proportion of people disavow essentialist claims about race?

Some studies *have* found that people tend to see racial categories as high in categorical entitativity (e.g. see Haslam et al., 2000). However, other studies suggest that folk beliefs vary to a significant degree. For instance, Shulman and Glasgow (2010) found that 25% of US adults do not believe that race is real. And among the 75% who said that race is real, a significant

⁷ No doubt, various other factors have a role to play in explaining this sort of case. For example, knowing that Italian-Americans immigrated from the same geographical region may have caused some to construe them as sharing an *inherited* essence. Moreover, there is evidence that people with a “social dominance orientation”—i.e. a general preference for social inequality—tend to biologize and dehumanize social outgroups (Costello & Hodson, 2010).

proportion rejected the view that races are natural kinds. For example, 21% of realists regarded races as purely social categories, while 25% regarded races as hybrid biosocial categories (see also Condit et al., 2004; Dubriwny et al., 2004). How do we explain this sort of variation?

The first thing to note about the studies mentioned above is that they surveyed participants on their explicit, self-reported, beliefs about race. Various studies have found that when it comes to *implicit* attitudes about race (and other social categories), people tend to be more essentialist in their thinking (e.g., see Chas et al., 2018; Goff et al., 2008; Karasawa et al., 2019).

The second thing to note is that various factors modulate the way in which similarity cues are consumed by the mechanisms responsible for categorical entitativity perception. For instance, in a study by Ip and colleagues (2009), participants watched animations depicting humanoid creatures called “Volotos.” In some trials, each creature had the same skin color, while in the others, skin color varied. Moreover, in some trials, the creatures moved in synchrony, while in others, they stood still. Finally, in some trials, a second group of creatures was located nearby. Immediately after observing each animation, participants were asked to rate the target group’s degree of entitativity. They were also asked to specify the degree to which group members shared common psychological traits, as well as the degree to which they shared common goals.

What Ip and colleagues found is that when skin color varied, but group members moved in synchrony, this led to the attribution of common goals. In contrast, when skin color was uniform, this led to the attribution of common psychological traits, but only if a second group of creatures, with different skin colors, was present. Ip and colleagues conclude that inferring common psychological traits from common physical traits is constrained by what Turner (1987) calls “the principle of metacontrast.” According to this principle, perceivers are more likely to process an aggregate of people as entitative when they detect low intragroup variability and high intergroup

variability with respect to some salient feature. Thus, for skin color to be seen as diagnostic of group membership, it must be seen as varying to a significant degree between, but not within, groups.

This provides us with another way to explain variation in essentialist thinking about race. Consider an individual who lives in a city, the neighborhoods of which are segregated according to race, versus an individual who lives in a more racially integrated setting. The first individual is more likely to see racial markers as diagnostic of group membership than the second one. The first individual is therefore more likely to perceive racial groups as high in categorical entitativity.

This hypothesis is supported by a number of studies, which have found that living in a racially homogenous setting predicts essentialist thinking about race. For instance, Pauker and colleagues (2018) performed a longitudinal study of people who moved from continental United States to Hawaii: the most racially diverse area in the United States. Utilizing the *Race Conceptions Scale* (Williams & Eberhardt, 2008), they found that essentialist beliefs about race decreased over time. More specifically, they found that those who showed the greatest decrease in essentialist thinking had the greatest increase in diversity of acquaintances (for similar findings, see Deeb et al., 2011; Mandalaywala et al., 2019; Pauker et al., 2016).

Other lines of research suggest that an individual's immediate social environment affects the degree to which *perceptual* mechanisms process racial markers as entitativity cues. For example, Phillips and colleagues (2018) tested the hypothesis that group impressions are extracted via ensemble-coding mechanisms: perceptual mechanisms which extract the central tendency of a group of objects in a rapid, and parallel, manner. In each trial, participants were shown two successive sets of faces. In one study, they were asked to specify which group was more racially

diverse. In another study, they were asked to specify which group was more diverse in terms of gender. In a third study, they were tasked with specifying which group was more hierarchical. Importantly, each image was presented for 200 ms, before being backwards masked. The brief presentation time was designed to prevent serial processing (saccades take longer than 200 ms). Participants were above chance at detecting the more racially diverse group, as well as the more hierarchical group.

For our purposes, this sort of finding is important because it suggests that a crowds' degree of physical homogeneity is calculated by a stimulus-driven mechanism. This therefore provides us with another explanation of why racial essentialism is enhanced in racially segregated settings. It is because this mechanism is being exposed to crowds that exhibit relatively low variation in racial markers, the result of which is categorical entitativity perception and thus a tendency towards essentialist thinking.

Other studies provide additional support for the hypothesis that categorical entitativity perception is driven, in part, by ensemble-coding mechanisms. Lamer and colleagues (2018) showed images of racially diverse crowds to participants (half of the individuals depicted were Black; the other half were White). Each image was presented for 300 ms. For participants in the emotional segregation condition, there was a large difference in the average emotional expression across Black and White subgroups; whereas, for participants in the control condition, the difference was small. Participants were then given two tasks. In the race categorization task, they had to identify a racially ambiguous face as "Black," "White," or "Biracial." The purpose of this task was to assess the extent to which participants draw strict boundaries between racial groups. In the race essentialism task, the Race Conceptions Scale was used to assess participants' degree of racial essentialism.

Lamer and colleagues found that the participants in the emotional segregation condition saw the boundaries between racial categories as stricter than those in the control condition did. Participants in the emotional segregation condition were also more essentialist in their thinking about race. By including a condition in which participants only saw a single face, rather than a crowd, Lamer and colleagues were also able to show that these effects were mediated by ensemble coding. This supports the hypothesis that when racially homogenous crowds are visually salient, ensemble coding is triggered. This results in categorical entitativity perception, and thus, essentialist thinking about race (see also Enos & Celaya, 2018).

4.2.2 The etiology of categorical entitativity perception

While the studies outlined above suggest that racial essentialism is driven by the mechanisms responsible for categorical entitativity perception, they do not directly address the origins of these mechanisms. It is possible that categorical entitativity perception arises from a domain-general tendency to essentialize categories: one that manifests itself differently across social and non-social domains (for a relevant discussion, see Gelman, 2003, Chapter 11). Alternatively, it is possible that there is a domain-specific mechanism that was adapted for essentializing human groups on the basis of perceived similarities between members.

The fact that categorical entitativity perception is driven by low-level ensemble-coding mechanisms provides initial support for the latter view. In fact, in interpreting their findings, Lamer and colleagues (2018) suggest that although ensemble-coding mechanisms evolved for the purposes of processing arrays of non-social features, such as shapes and colors, they may have been exapted for the purposes of detecting social groups as well.

While plausible enough, additional data is needed to assess this sort of hypothesis. At present, research into “people perception”—the tendency to form rapid, stimulus-driven, impressions of

crowds—is still in its infancy (for an overview, see Phillips et al., 2014). By further examining the developmental and cross-cultural profile of ensemble-coding mechanisms, we stand to gain a much better understanding of the origins of categorical entitativity perception, and thereby racial essentialism.⁸

4.3 *The role of dynamical entitativity perception*

I have just argued that categorical entitativity perception drives essentialist thinking about race. What role, if any, does *dynamical* entitativity perception play in driving ordinary thinking about race?

Some evolutionary psychologists argue that we possess an “alliance-detection system,” which was adapted for tracking alliances (Cosmides et al., 2003; Kurzban et al., 2001; Pietraszewski et al., 2014, 2015; see also Hirschfeld, 2001, p. 113). Given that alliances are dynamically entitative groups, another way to state this view is that the alliance-detection-system tracks groups that are high in dynamical entitativity. If societal conditions are such that racial markers

⁸ The account I have just defended is compatible with the view that racial categorization is a modern phenomenon. Plausibly, before the modern concept of race emerged, the most salient social groups cut across racial lines. For instance, Ivan Hannaford argues persuasively that before the Renaissance and the Reformation, “one’s identity and social status were determined by one’s political and religious commitments, not the color of one’s skin” (1996, pp. 147–148). Once racial markers became salient cues for group membership, this would have engaged the mechanisms responsible for categorical entitativity perception, thereby fueling the acquisition of distinctively modern, essentialist conceptions, of race. What might have caused racial markers to become salient, though? A large body of research has found that people tend to form social networks with those who are similar to them. Psychologists refer to this as *homophily* (for an overview, see McPherson et al., 2001). There has been little research into how people weigh the different dimensions of similarity (e.g. race versus age) when forming social networks. Plausibly, if one’s social standing is largely determined by one’s religious and political beliefs, the visible markers of race will not be especially salient. However, in a society in which these beliefs are less powerful drivers of social standing, racial markers may be salient enough to shape the emergence of homophilous social networks. The emergence of these racially skewed networks will, in turn, increase the salience of racial markers. Of course, there is a lot more to be said about the historical and psychological factors that drive racial segregation. Suffice it to say that additional research into the multidimensional nature of homophily may provide important insights.

are a reliable indicator of who is allied with whom then, according to these theorists, the alliance-detection-system picks up on this, and represents bearers of the same markers as allies.

Evidence for the theory comes from a series of studies based on the “Who Said What?” paradigm. Subjects are shown a series of pictures, each one depicting a person issuing a different statement. Then, in the second phase of the study, subjects are given a surprise memory test, in which they are asked to recall who said what. The pattern of mistakes that subjects make is interpreted as revealing the categories they automatically represent. The assumption here is that if *S* automatically categorizes *A* and *B* in the same way (e.g. *as women*) then that raises the likelihood that *S* will attribute a statement to *A*, when it was actually *B* who made the statement. This is referred to as a “within-category error.” “Between-category errors” occur when the subject confuses members of distinct categories. As it turns out, subjects tend to make significantly more within-category errors than between-category ones.

Initial studies found that the most persistent and widespread within-category errors concern sex, race, and age. However, in a series of more recent studies, proponents of the alliance-detection theory have shown that when alliances cut across racial lines, within-category errors for race decrease significantly, and are replaced by within-category errors for the relevant alliance-category (Kurzban et al., 2001; Pietraszewski et al., 2014, 2015). Pietraszewski and colleagues (2014) conclude that “categorizing people by their race is a reversible product of a cognitive system specialized for detecting alliance categories and regulating their use” (2014, p. 1).

4.3.1 Do the folk really see racialized groups as alliances?

In responding to the alliance-detection account, some theorists have suggested that racialized groups do not have the hallmarks of alliances (see Glasgow, 2009, pp. 68–69; Machery & Faucher, 2017, pp. 1157–1158). Unfortunately, this is an understudied issue.

Lickel and colleagues (2000) asked participants to rate various groups in terms of their degree of entitativity; degree of interaction among group members; and degree to which group members have common goals and outcomes. The only racial group that participants were asked about were “Blacks.” This group was rated as relatively low in entitativity (2000, p. 227). Moreover, unlike families and other paradigmatically coalitional groups, this racial group was not seen as possessing the hallmarks of a dynamically entitative group. For example, its members were only rated as sharing common goals and outcomes to a relatively small degree (2000, p. 231). It might be tempting to conclude that the folk do not see racial groups as alliances; however, to draw that conclusion, different sorts of studies would be required.

First, it may be that the folk *implicitly* process racial groups as high in dynamical entitativity: in Lickel and colleagues’ study, only explicit measures were utilized. A number of studies have found that groups are automatically represented as dynamically entitative when certain cues are visible, such as synchronous movement (Bloom & Veres, 1999; Fawcett & Tunçgenç, 2017; Fessler & Holbrook, 2016). In racially segregated settings, these sorts of cues might cause people to implicitly process racial groups as dynamically entitative. For example, frequently seeing members of the same racial group walk together, talk together, and engage in joint attention, might trigger the mechanisms in question.

Second, as the study by Rutchick and colleagues (2008) shows, individuals can be primed to think of the same group in different ways: namely, as categorically entitative or as dynamically

entitative. It may be that in some contexts, racial groups are thought of as dynamically entitative, whereas, in others they are not (on this point, see Phillips, 2020).

In one of the only studies of its kind, Pietraszewski (2016) utilized a version of the *Who said what?* paradigm to assess how racially charged primes would affect participants' performance. Participants saw a series of photographs: half depicted White individuals, while the other half depicted Black individuals. Some of the individuals made statements indicating that they work for the same charity (i.e. a dynamically entitative group). Importantly, some of the individuals in the charity group were Black, while others were White. Before being asked to recall who said what, participants completed a distractor task. In the prime condition, participants were asked to think about racial conflict and inequality: they also viewed photographs depicting racial conflict. The purpose of this prime was to make participants think about the "coalitional relevance" of race (2016, 87). In the nonprime condition, participants were asked to name US state capitals.

Pietraszewski (2016) found that in the nonprime condition, participants tended to confuse members of the charity group. However, in the prime condition, they tended to confuse members of the same racial group. According to Pietraszewski, participants confused members of the same racial group in the prime condition because they were construing them as members of the same alliance; whereas, in the nonprime condition, participants only confused members of the same charity group, because this was the only salient alliance.

Pietraszewski's study provides some evidence that people can be primed to think of racial groups as alliances (i.e. as high in dynamically entitativity). The study by Rutchick and colleagues (2008) also provides some indirect evidence. Recall that one of the target groups was a Bhutanese social caste. Exposure to the dynamical prime—a description of a functioning beehive—enhanced participants' impression that this caste is a dynamically entitative group.

Given the close resemblance between social castes and racial groups, it would not be surprising to find that certain primes can cause people to think of racial groups as dynamically entitative as well.⁹

4.3.2 The etiology of dynamical entitativity perception

What is the origin of dynamical entitativity perception? As was explained above, Cosmides and colleagues argue that we possess a mechanism that was adapted for tracking alliances. In addition to the evidence they adduce, a number of studies suggest that children regard coalitional groups as an important type of social category (for an overview, see Rhodes, 2013a). For instance, Kalish and Lawson (2008) found that young children consistently regard the members of novel social categories (introduced via novel labels) as having obligations to one another (for similar findings, see Rhodes, 2012).

In fact, there is evidence that infants engage in dynamical entitativity perception. Fawcett and Tunçgenç (2017) found that 15-month-olds use synchronous movement to predict who is affiliated with whom. By measuring looking times, Powell and Spelke (2018) found that infants as young as 4 months of age expect agents to form affiliations with those whose movements they imitate.

Thus, the developmental evidence supports the hypothesis that we possess mechanisms that were adapted for tracking alliances. However, to assess the hypothesis that ordinary racial thinking is driven, to a significant extent, by these mechanisms, more cross-cultural and developmental studies are needed. One developmental study found that preschool-age children use the visible markers of race to predict allegiances (Shutts et al., 2013). In extending this finding, future studies should present both children and adults with a wide variety of dynamical

⁹ See Wilkerson (2020) for a discussion of the resemblances between social castes and racial groups.

entitativity cues, such as video footage of same-race individuals walking together, and engaging in other forms of affiliative activity. If, on the basis of these cues, participants tend to represent members of the given racial group as allies, this will constitute evidence that dynamical entitativity perception has a significant role to play in explaining the origins of racial thinking.

5 Conclusion

I have defended a pluralistic thesis according to which different ways of thinking about race are driven by distinct mechanisms. I began by arguing that the mechanisms responsible for face recognition have an important, albeit delimited, role to play in sustaining the belief that visible features are diagnostic of race. I then argued that essentialist beliefs about race are driven, in large part, by some of the mechanisms responsible for entitativity perception. In elaborating on the role that these mechanisms play, I distinguished between two varieties of entitativity perception: categorical and dynamical.

Categorical entitativity perception drives broad essentialist thinking about race. Moreover, the evidence reviewed above suggests that this form of entitativity perception is subserved by low-level ensemble-coding mechanisms. It is possible that these mechanisms were exapted for the purposes of detecting social groups. Assessing this sort of hypothesis will require more research into the developmental, and cross-cultural, profile of these mechanisms.

In examining the role of dynamical entitativity perception, I argued that there are good reasons for thinking that it drives the tendency to think of racial groups as coalitions. The evidence reviewed above suggests that this variety of entitativity perception is subserved by early-developing mechanisms, which are tuned to low-level cues, such as synchronous movement. This supports the hypothesis that coalitional thinking about race is produced by mechanisms that were adapted for tracking coalitions. However, if we are to determine the

prevalence of coalitional thinking about race, more developmental and cross-cultural studies are needed.

Interestingly, two studies found that dynamical entitativity perception is the predominant mode of group perception among Japanese individuals; whereas, categorical entitativity perception is the predominant mode among American individuals (Yuki, 2003; Kurebayashi et al., 2012). Additional studies of this sort may generate novel explanations of cross-cultural variation in racial categorization. For example, it is possible that Japanese individuals are more likely than American individuals to see racial groups as coalitional; whereas, American individuals are more likely than Japanese individuals to see members of the same racial group as sharing a fixed essence.

In addition to explaining how and why we categorize people according to race, the pluralistic approach I have defended has implications for other issues surrounding the psychology of race. For example, there is evidence that the different forms of entitativity perception give rise to stereotypes with distinctive contents (Spencer-Rodgers et al., 2007). Thus, studying the relation between entitativity perception and racial categorization may provide new insights into how racial stereotyping works (for further discussion, see Phillips, 2020).

Other studies have found that when people perceive an outgroup as dynamically entitative, they tend to hold each member responsible for the actions of the group. This can lead to acts of vicarious retribution in which a member of the outgroup is punished, even if they had nothing to do with the incident that provoked the punitive action (Lickel et al., 2006). Thus, by examining the relation between dynamical entitativity perception and racial categorization, researchers may gain new insights into those forms of racial prejudice that involve vicarious retribution (for a number of potential examples, see Smith, 2011).

Finally, the pluralistic account I have defended has consequences for debates about the ethics of racial categorization. According to eliminativists, we should eliminate racial categorization altogether. Conservationists, however, argue that instead of eliminating racial categorization, the best way to deal with racial inequality is to retain racial categorization, but to remove its most pernicious aspects, such as the belief that racial groups are biologically distinct (for an overview, see Kelly, Machery, & Mallon, 2010).

I have argued that cues for categorical entitativity perception are highly salient when racial groups are segregated. Some of the cues for dynamical entitativity perception, such as synchronicity, are also highly salient when racial groups are segregated. This suggests that the prospects for eliminating racial categorization hang on the feasibility of eliminating racial segregation.

Suppose, however, that removing *biological* thinking about race is the goal. According to the cognitive-evolutionary theories surveyed in section 3, when racial categorization occurs, this brings inheritance thinking in its train. If one of these theories is right then preserving racial categorization, while jettisoning biological thinking about race, may be especially difficult (on this point, see Kelly, Machery, & Mallon, 2010, pp. 449–450). In contrast, according to the pluralistic account I have defended, racial categorization does not necessarily bring inheritance thinking in its train. Most obviously, the mechanisms responsible for coalitional thinking about race do not produce inheritance thinking about race. As was mentioned above, though, there are reasons for thinking that seeing racial groups as coalitional engenders distinctive forms of racial prejudice and stereotyping. In pursuing a conservationist approach to racial categorization, these pernicious effects will require serious consideration.

References

- Astuti, R., G. E. A. Solomon, & S. Carey (2004). Constraints on cognitive development: A case study of the acquisition of folkbiological and folksociological knowledge in Madagascar. *Monographs of the society for research in child development*, 69(3): i, v, vii–viii, 1–161.
- Atran, S. (1990). *Cognitive foundations of natural history*. Cambridge: Cambridge University Press.
- Atran, S. (2001). Comment on F. J. Gil-White’s article ‘Are ethnic groups biological “species” to the human brain? Essentialism in our cognition of some social categories. *Current Anthropology*, 42(4), 537–538.
- Atran, S., & Medin, D. (2010). *The native mind and the cultural construction of nature*. Cambridge, MA: MIT Press.
- Balas, B. (2012). Bayesian face recognition and perceptual narrowing in face-space. *Developmental Science*, 15(4), 579–588.
- Balas, B., Peissig, J., & Moulson, M. (2015). Children (but not adults) judge similarity in own- and other-race faces by the color of their skin. *Journal of Experimental Child Psychology*, 130, 56–66.
- Baudouin, J., & Gallay, M. (2006). Is face distinctiveness gender based? *Journal of Experimental Psychology: Human Perception and Performance*, 32, 789–798.
- Block, N. (2014). Seeing-as in the light of vision science. *Philosophy and Phenomenological Research*, 89(1), 560–572.
- Bloom, P., & Veres, C. (1999). The perceived intentionality of groups. *Cognition*, 71(1), B1–B9.
- Chas, A., Betancor, V., Delgado, N., & Rodríguez-Pérez, A. (2018). Children consider their own group to be more human than other social groups: Evidence from indirect and direct measures. *Social Psychology*, 49(3), 125–134.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Condit, C. M., Parrott, R. L., Harris, T. M., Lynch, J., & Dubriwny, T. (2004). The role of “genetics” in popular understandings of race in the United States. *Public Understanding of Science*, 13, 249–272.

- Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Science*, 7(4), 173–179.
- Costello, K., & Hodson, G. (2010). Exploring the roots of dehumanization: The role of animal-human similarity in promoting immigrant humanization. *Group Processes and Intergroup Relations*, 13, 3–22.
- Crawford, M. T., Sherman, S. J., & Hamilton, D. L. (2002). Perceived entitativity, stereotype formation, and the interchangeability of group members. *Journal of Personality and Social Psychology*, 83, 1076–1094.
- Dahl, C. D., Rasch, M. J., Bülthoff, I., & Chen, C. (2016). Integration or separation in the processing of facial properties—A computational view. *Scientific Reports*, 6, 20247.
- Dasgupta, N., Banaji, M. R., & Abelson, R. P. (1999). Group entitativity and group perception: Associations between physical features and psychological judgment. *Journal of Personality and Social Psychology*, 77, 991–1003.
- Deeb, I., Segall, G., Birnbaum, D., Ben-Eliyahu, A., & Diesendruck, G. (2011). Seeing isn't believing: The effect of intergroup exposure on children's essentialist beliefs about ethnic categories. *Journal of Personality and Social Psychology*, 101, 1139–1156.
- Dubriwny, T. N., Bates, B. R., & Bevan, J. L. (2004). Lay understandings of race: Cultural and genetic definitions. *Community Genetics*, 7, 185–195.
- Dufour, V., Pascalis, O., & Petit, O. (2006). Face processing limitation to own species in primates: a comparative study in brown capuchins, Tonkean macaques and humans. *Behavioural Processes*, 73, 107–113.
- Dunham, Y., Stepanova, E. V., Dotsch, R., & Todorov, A. (2014). The development of race-based perceptual categorization: skin color dominates early category judgments. *Developmental Science*, 18(3), 469–483.
- Enos, D. E., & Celaya, C. (2018). The effect of segregation on intergroup relations. *Journal of Experimental Political Science*, 5(1), 26–38.

- Fawcett, C., & Tunçgenç, B. (2017). Infants' use of movement synchrony to infer social affiliation in others. *Journal of Experimental Child Psychology*, 160, 127–136.
- Fessler, D. M., & Holbrook, C. (2016). Synchronized behaviour increases assessments of the formidability and cohesion of coalitions. *Evolution and Human Behavior*, 37(6), 502–509.
- Freire, A., & Lee, K. (2001). Face recognition in 4- to 7-year-olds: processing of configural, featural, and paraphernalia information. *Journal of Experimental Child Psychology*, 80(4), 347–371.
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Press.
- Gil-White, F. (2001). Are ethnic groups biological 'species' to the human brain. *Current Anthropology*, 42(4), 515–554.
- Giménez, M., & Harris, P. L. (2010). Understanding constraints on inheritance: Evidence for biological thinking in early childhood. *British Journal of Developmental Psychology*, 20(3), 307–324.
- Glasgow, J. (2009). *A theory of race*. New York: Routledge.
- Goff, P. A., Eberhardt, J. L., Williams, M. J., & Jackson, M. C. (2008). Not yet human: Implicit knowledge, historical dehumanization, and contemporary consequences. *Journal of Personality and Social Psychology*, 94(2), 292–306.
- Hannaford, I. (1996). *Race: The history of an idea in the West*. Washington: Woodrow Wilson Center Press.
- Hardimon, M. (2003). The ordinary concept of race. *The Journal of Philosophy*, 100(9), 437–455.
- Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, 39, 113–127.
- Hills, P. J., & Lewis, M. B. (2018). The development of face expertise: Evidence for a qualitative change in processing. *Cognitive Development*, 48, 1–18.
- Hirschfeld, L. A. (1993). Discovering social difference: The role of appearance in the development of racial awareness. *Cognitive Psychology*, 25, 317–350.
- Hirschfeld, L. A. (1995). Do children have a theory of race? *Cognition*, 54, 209–252.

- Hirschfeld, L. A. (1996). *Race in making: Cognition, culture, and the child's construction of human kinds*. Cambridge, MA: MIT Press.
- Hirschfeld, L. A. (2001). On a folk theory of society: Children, evolution, and mental representations of social groups. *Personality and Social Psychology Review*, 5(2), 107–117.
- Hirschfeld, L. A., Bartmess, E., White, S., & Frith, U. (2007). Can autistic children predictive behavior by social stereotypes? *Current Biology*, 17(12), 451–452.
- Hochman, A. (2013). Do we need a device to acquire ethnic concepts? *Philosophy of Science*, 80(5), 994–1005.
- Hochman, A. (2019) Is race modern? Disambiguating the question. *Du Bois Review: Social Science Research on Race*, 16(2), 647–665.
- Ip, G. W. M., Chiu, C. Y., & Wan, C. (2006). Birds of a feather and birds flocking together: Physical versus behavioral cues may lead to trait- versus goal-based group perception. *Journal of Personality and Social Psychology*, 90, 368–381.
- Ito, T. A., & Tomelleri, S. (2017). Seeing is not stereotyping: the functional independence of categorization and stereotype activation. *Social Cognitive and Affective Neuroscience*, 758–764.
- Jackson, J. P. (2017). Cognitive/evolutionary psychology and the history of racism. *Philosophy of Science*, 84(2), 296–314.
- Jaquet, E., Rhodes, G., & Hayward, W. G. (2007). Opposite aftereffects for Chinese and Caucasian faces are selective for social category information and not just physical face differences. *Quarterly Journal of Experimental Psychology*, 60, 1457–1467.
- Jaquet, E., Rhodes, G., & Hayward, W. G. (2008). Race-contingent aftereffects suggest distinct perceptual norms for different race faces. *Visual Cognition*, 16, 734–753.
- Kalish, C., & Lawson, C. (2008). Development of social category representations: Early appreciation of roles and deontic relations. *Child Development*, 79, 577–593.
- Karasawa, M., Asai, N., & Hioki, K. (2019). Psychological essentialism at the explicit and implicit levels: The unique status of social categories. *Japanese Psychological Research*, 61(2), 107–122.

- Kelly, D., Machery, E., & Mallon, R. (2010). Race and racial cognition. In J. Doris & the Moral Psychology Reading Group (Eds.), *The Moral Psychology Handbook* (pp. 433–472). Oxford: Oxford University Press.
- Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science*, 18, 1084–1089.
- Kobayashi, M., Macchi Cassia, V., Kanazawa, S., Yamaguchi, M. K., & Kakigi, R. (2018). Perceptual narrowing towards adult faces is a cross-cultural phenomenon in infancy: A behavioral and near-infrared spectroscopy study with Japanese infants. *Developmental Science*, 21(1), e12498.
- Kramer, R. S., Young, A. W., Day, M. G., & Burton, A. M. (2017). Robust social categorization emerges from learning the identities of very few faces. *Psychological Review*, 124(2), 115–129.
- Kurebayashi, M. A., Hoffman, L. R., Ryan, C. S., & Murayama, A. (2012). Japanese and American perceptions of group entitativity and autonomy: A multilevel analysis. *Journal of Cross-Cultural Psychology*, 43, 349–365.
- Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, 98, 15387–15392.
- Lamer, S. A., Sweeny, T. D., Dyer, M. L., & Weisbuch, M. (2018). Rapid visual perception of interracial crowds: Racial category learning from emotional segregation. *Journal of Experimental Psychology: General*, 147(5), 683–701.
- Leslie, S.-J. (2014). Carving up the social world with generics. *Oxford Studies in Experimental Philosophy*, 1, 208–232.
- Levin, D.T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General*, 129, 559–574.
- Lieberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). The origins of social categorization. *Trends in Cognitive Sciences*, 21(7), 556–568.

- Lickel, B., Hamilton, D. L., Wiczokowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, 78, 223–246.
- Lickel, B., Miller, N., Stenstrom, D. M., Denson, T. F., & Schmader, T. (2006). Vicarious retribution: The role of collective blame in intergroup aggression. *Personality and Social Psychology Review*, 10(4), 372–390.
- Little, A. C., DeBruine, L. M., & Jones, B. C. (2005). Sex-contingent face aftereffects suggest distinct neural populations code male and female faces. *Proceedings of the Royal Society of London, Series B*, 272, 2283–2287.
- Little, A. C., DeBruine, L. M., Jones, B. C., & Watt, C. (2008). Category contingent aftereffects for faces of different races, ages and species. *Cognition*, 106, 1537–1547.
- Machery, E., & Faucher, L. (2005). Social construction and the concept of race. *Philosophy of Science*, 72(5), 1208–1219.
- Machery, E., & Faucher, L. (2017). Why do we think racially? Culture, evolution, and cognition. In H. Cohen and C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science*, 2nd edition (pp. 1135–1176). Amsterdam: Elsevier.
- Machery, E., & Faucher, L. (2020). The folk concept of race. Forthcoming in T. Marques and Å. Wikforss (Eds.), *Shifting concepts: The philosophy and psychology of conceptual variability*. Oxford University Press.
- Maddox, K. B. (2004). Perspectives on racial phenotypicality bias. *Personality and Social Psychology Review*, 8(4), 383–401.
- Maddox, K. B., & Perry, J. M. (2018). Racial appearance bias: Improving evidence-based policies to address racial disparities. *Policy Insights from the Behavioral and Brain Sciences*, 5(1), 57–65.
- Mallon, R. (2016). *The construction of human kinds*. Oxford: Oxford University Press.

- Mandalaywala, T., Ranger-Murdock, G., Amodio, D. M., & Rhodes, M. (2019). The nature and consequences of essentialist beliefs about race in early childhood. *Child Development*, 90(4), e427–e453.
- McKone, E., Crookes, K., & Kanwisher, N. (2009). The cognitive and neural development of face recognition in humans. In M. S. Gazzaniga, E. Bizzi, L. M. Chalupa, S. T. Grafton, T. F. Heatherton, C. Koch, J. E. LeDoux, S. J. Luck, G. R. Mangan, J. A. Movshon, H. Neville, E. A. Phelps, P. Rakic, D. L. Schacter, M. Sur, & B. A. Wandell (Eds.), *The cognitive neurosciences* (pp. 467–482). Massachusetts Institute of Technology.
- McPherson, M., Smith-Lovin, L., & Cook, J.M. (2001). Birds of a feather: Homophily in social networks. *Review of Sociology*, 27, 415–444.
- Meissner, C. A. & Brigham, J. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7(1), 3–35.
- Mondloch, C.J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, 31(5), 553–566.
- Moya, C. (2013). Evolved priors for ethnolinguistic categorization: A case study from the Quechua-Aymara boundary in the Peruvian Altiplano. *Evolution and Behavior*, 34, 265–272.
- Moya, C., & Boyd, R. (2015). Different selection pressures give rise to distinct ethnic phenomena: A functionalist framework with illustrations from the Peruvian Altiplano. *Human Nature*: <http://dx.doi.org/10.1007/s12110-015-9224-9>.
- Moya, C., & Henrich, J. (2016). Culture-gene coevolutionary psychology: Cultural learning, language, and ethnic psychology. *Current Opinion in Psychology*, 8, 112–118.
- O’Toole, A. J., Abdi, H., Deffenbacher, K. A., & Bartlett, J. C. (1991). Classifying faces by race and sex using an autoassociative memory trained for recognition. In K. J. Hammomd & D. Gentner (Ed.), *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society* (pp. 847–851). Hillsdale, NJ: Erlbaum.

- Pascalis, O., & Bachevalier, J. (1998). Face recognition in primates: A cross-species study. *Behavioural Processes*, 43, 87–96.
- Pauker, K., Carpinella, C., Meyers, C., Young, D. M., & Sanchez, D. T. (2018). The role of diversity exposure in Whites' reduction in race essentialism over time. *Social Psychological and Personality Science*, 9(8), 944–952.
- Pauker, K., Xu, Y., Williams, A., & Biddle, A. M. (2016). Race essentialism and social contextual differences in children's racial stereotyping. *Child Development*, 87, 1409–1422.
- Phillips, B. 2019. The shifting border between perception and cognition. *Noûs*, 53 (2): 316–346.
- Phillips, B. 2020. Entitativity and Implicit Measures of Social Cognition. *Mind & Language*.
<https://doi.org/10.1111/mila.12350>.
- Phillips, B. 2021. Seeing seeing. *Philosophy and Phenomenological Research*, 102 (1): 24–43
- Phillips, L. T., Slepian, M. L., & Hughes, B. L. (2018). Perceiving groups: The people perception of diversity and hierarchy. *Journal of Personality and Social Psychology*, 114(5), 766–785.
- Phillips, L. T., Weisbuch, M., & Ambady, N. (2014). People perception: Social vision of groups and consequences for organizing and interacting. *Research in Organizational Behavior*, 34, 101–127.
- Pietraszewski, D. (2016). Priming race: Does the mind inhibit categorization by race at encoding or recall? *Social Psychological and Personality Science*, 7(1), 85–91.
- Pietraszewski, D., Cosmides, L., & Tooby, J. (2014). The content of our cooperation, not the color of our skin: An alliance detection system regulates categorization by coalition and race, but not sex. *PLoS ONE*, 9, e88534. <https://doi.org/10.1371/journal.pone.0088534>
- Pietraszewski, D., Curry, O., Petersen, M. B., Cosmides, L., & Tooby, J. (2015). Constituents of political cognition: Race, party politics, and the alliance detection system. *Cognition*, 140, 24–39.
- Powell, L. J., & Spelke, E. S. (2018). Human infants' understanding of social imitation: Inferences of affiliation from third-party observations. *Cognition*, 170, 31–48.
- Quinn, P. C., & Eimas, P. D. (1997). A reexamination of the perceptual-to-conceptual shift in mental representations. *Review of General Psychology*, 1, 271–287.

- Rhodes, M. (2012). Naïve theories of social groups. *Child Development*, 83(6), 1900–1916.
- Rhodes, M. (2013a). The conceptual structure of social categories: The social allegiance hypothesis. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us* (pp. 258–262). Oxford University Press.
- Rhodes, M. (2013b). How two intuitive theories shape the development of social categorization. *Child Development Perspectives*, 7(1), 12–16.
- Rhodes, M., & Gelman, S. A. (2009). Five-year-olds' beliefs about the discreteness of category boundaries for animals and artifacts. *Psychonomic Bulletin & Review*, 16, 920–924.
- Rhodes, M., Gelman, S. A., & Karuza, J. C. (2014). Preschool ontology: The role of beliefs about category boundaries in early categorization. *Journal of Cognition and Development*, 15(1), 78–93.
- Rhodes, G., Watson, T. L., Jeffrey, L., & Clifford, C. W. G. (2010). Perceptual adaptation helps us identify faces. *Vision Research*, 50, 963–968.
- Rice, A., Phillips, P. J., Natu, V., An, X., & O'Toole, A. J. (2013). Unaware person recognition from the body when face identification fails. *Psychological Science*, 24(11), 2235–2243.
- Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: The University of Chicago Press.
- Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G. R. Semin & K. Fiedler (Eds.), *Language, interaction and social cognition* (pp. 11–36). London: Sage.
- Rutchick, A. M., Hamilton, D. L., & Sack, J. D. (2008). Antecedents of entitativity in categorically and dynamically construed groups. *European Journal of Social Psychology*, 38, 905–921.
- Shutts, K., Roben, C., & Spelke, E. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development*, 14(1), 35–62.
- Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American Psychologist*, 60(1), 16–26.

- Smith, D. L. (2011). *Less than human: Why we demean, enslave, and exterminate others*. New York: St. Martin's Press.
- Spencer-Rodgers, J., Hamilton, D. L., & Sherman, S. J. (2007). The central role of entitativity in stereotypes of social categories and task groups. *Journal of Personality and Social Psychology*, 92, 369–388.
- Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences*, 105, 394–398.
- Suhrke, J., Freitag, C., Lamm, B., Teiser, J., Fassbender, I., Poloczek, S., et al. (2014). The other-race effect in 3-year-old German and Cameroonian children. *Frontiers in Psychology*, 5, 198.
- Tham, D. S. Y., & Woo, P. J. (2019). Development of the other-race effect in Malaysian-Chinese infants. *Developmental Psychobiology*, 61(1), 107–115.
- Turner, J. C. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, UK: Basil Blackwell.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology Human Experimental Psychology*, 43a, 161–204.
- Valentine, T., Lewis, M. B., & Hills, P. J. (2016). Face-space: A unifying concept in face recognition research. *The Quarterly Journal of Experimental Psychology*, 69(10), 1996–2019.
- Waxman, S. (2010). Names will never hurt me? Naming and the development of racial and gender categories in preschool-aged children. *European Journal of Social Psychology*, 40, 593–610.
- Wilkerson, I. (2020). *Caste: The origins of our discontents*. New York: Random House.
- Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology*, 94, 1033–1047
- Williams, J. E., & Roberson, J. K. (1967). A method for assessing racial attitudes in preschool children. *Educational and Psychological Measurement*, 27(3), 671–689.

- Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: a critical review of theories of the cross-race effect. *Personality and Social Psychology Review*, 16(2), 116–142.
- Yuki, M. (2003). Intergroup comparison versus intragroup relationship: A cross-cultural examination of social identity theory in North American and East Asian cultural contexts. *Social Psychology Quarterly*, 66, 166–183.
- Yzerbyt, V. Y., Rocher, S., & Schadron, G. (1997). Stereotypes as explanations: A subjective essentialistic view of group perception. In R. Spears, P. Oakes, N. Ellemers, & A. Haslam (Eds.), *The psychology of stereotyping and group life* (pp. 20–50). London: Basil Blackwell.
- Zack, N. (2002). *Philosophy of science and race*. New York: Routledge.
- Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision Research*, 41, 2979–2994.