Perceptions of race

Leda Cosmides[1], John Tooby[2] and Robert Kurzban[3]

[1]Department of Psychology, University of California, Santa Barbara, CA 93106, USA
[2]Department of Anthropology, University of California, Santa Barbara, CA 93106, USA
[3]Department of Psychology, University of Pennsylvania, 405 Psychology Office Building, 3815 Walnut Street, Philadelphia, PA, USA

Until recently, experiments on person perception had led to two unwelcome conclusions: (1) people encode the race of each individual they encounter, and (2) race encoding is caused by computational mechanisms whose operation is automatic and mandatory. Evolutionary analyses rule out the hypothesis that the brain mechanisms that cause race encoding evolved for that purpose. Consequently, race encoding must be a byproduct of mechanisms that evolved for some alternative function. But which one? Race is not encoded as a byproduct of domain-general perceptual processes. Two families of byproduct hypotheses remain: one invokes inferential machinery designed for tracking coalitional alliances, the other machinery designed for reasoning about natural kinds. Recent experiments show that manipulating coalitional variables can dramatically decrease the extent to which race is noticed and remembered.

Race-based inferences – stereotypes – are easy to activate and inactivate, given the appropriate context [1–2]. But what about race encoding? The race of an individual must be noticed and remembered before a racial stereotype can be activated or racially motivated behavior can occur. Is it possible *not* to notice a person's race?

Race exists in the minds of human beings. But geneticists have failed to discover objective patterns in the world that could easily explain the racial categories that seem so perceptually obvious to adults (see Fig. 1). When, beginning in the mid-1960s, the technology emerged to sequence genes and the proteins they code for, the distribution of alleles could at last be objectively mapped in the human species. What geneticists discovered was an underlying reality that bore no resemblance to existing hereditarian and folk theories of race. The first idea to be falsified was that most genetic variation in the human species served to differentiate races. Within-population genetic variance was found to be ~10 times greater than between-race genetic variance (i.e. two neighbors of the same 'race' differ many times more, genetically speaking, than a mathematically average member of one 'race' differs from an average member of another [3–5]). Second, compared with other similar species, there is little genetic diversity among humans. For example, the diversity of gene sequences among chimpanzees is almost four times higher than for humans [6], despite the fact that their population sizes are far smaller. In fact, the genetic distances for protein loci

between European, Asian and sub-Saharan African populations 'are of the same order of magnitude as those for local populations in other organisms and considerably smaller than those for subspecies.' (Ref. [4], p. 11). That is, by the criteria that biologists typically use to apply the concept of 'subspecies' or 'races', humans do not qualify. Most telling of all, virtually no expressed genes have been identified that are shared by all normal members of one race (and hence could explain a common racial appearance) that are not also present at substantial levels in other races (thereby failing to sort individuals into races) (see Fig. 1).
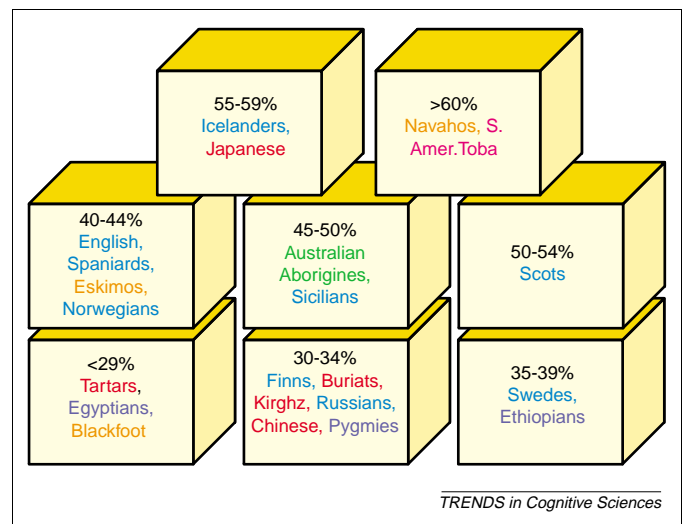

Fig. 1. Despite perceptions to the contrary, geneticists have shown that humanity is not divided into distinct racial types (for discussion, see Refs [16,40–42]). Geneticists have investigated the distribution of common protein-building genes – that is, those that can cause detectable phenotypic differences, and so have the potential for differentiating groups by appearance (shared racial appearance cannot be caused by genes that are unexpressed or rare). Most genes influencing appearance are not known, but by using the hundreds of common proteins that *are* known, human population structure could be mapped. The weak dimensions of observable geographical variation are typically clines whose scale, directionality, distribution and slope do not reflect each other (as they should if racial typologies were reflected in the real world) [43]. As an example, sorting human populations by the O allele of the ABO blood group gives a characteristically counterintuitive result, with Icelanders clustering with Japanese, Ethiopians clustered with Swedes, and so forth. Each block contains populations having a similar percentage of the O allele [40,5]. Different colors denote continent of origin. Sorting by other protein-building alleles gives different but equally counterintuitive results. In general, the frequency of a common, protein-building allele will cluster populations into groups that typically violate rather than support traditional racial categories [40,44]. The human species simply cannot be reliably sorted into types based on sets of genes that are shared by most members of the type but different from the genes shared within other types. Significantly, the patterns of genetic and phenotypic variation are sufficiently rich that by choosing alternative criteria, the human mind could be trained to cluster humans into a large number of alternative, mutually contradictory groupings.

Corresponding author: Leda Cosmides (cosmides@psych.ucsb.edu).

Box 1. The memory confusion protocol ('who said what?')

Developed in the 1970s by Taylor *et al*. [7], the memory confusion protocol uses errors in recall to unobtrusively reveal whether subjects are categorizing target individuals along a dimension of interest, such as race or sex. During phase 1, subjects are told that they will see some individuals engaged in a conversation, and that they should try to form an impression of each individual. They are then shown a series of photos of individuals, each of which is paired with a sentence that was supposedly uttered by that individual. The individuals in the photos differ along one or more dimensions of interest – race, sex, age, and so on. In a recent study by Kurzban *et al*. [21], each photo was shown for 8.5 s, and phase 1 lasted ∼4 min. (Fig. I). Phase 2 is a surprise recall test (after a distracter task) in which the subject is shown each sentence (in random order), and asked which individual uttered it (all photos are simultaneously present). This 'who said what?' task is difficult, and subjects make many mistakes. By analyzing the pattern of errors, the experimenter can tell whether the subject had encoded a categorical dimension of interest during phase 1. For example, suppose half the targets were women and half were men. If a subject encodes the sex of targets during phase 1, that subject's errors in phase 2 will not be random: they will be more likely to misattribute a sentence uttered by a man to another man rather than to a woman (and vice versa). Indeed, for any category of interest – race, sex, age, and so on – the experimenter will find that p(within category error) $>$ p(between category error) for any category that the subject did, in fact, encode. By contrast, a subject who did not encode sex during phase 1, will produce errors that are random with respect to that category during phase 2, that is, p(within category error) $=$ p(between category error). Past experiments using the memory confusion protocol showed that when the targets differed in race (black vs. white), subjects made significantly more within-race errors than between-race errors [7–9]. This means that subjects were encoding the race of targets in those experiments.
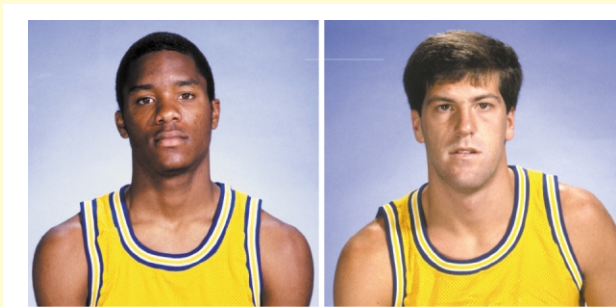

Fig. I. An example of two photos used in a memory confusion protocol experiment [21]. The pattern of errors made in these experiments can show whether subjects have encoded race.

Box 2. Attempts to reduce race encoding

Over the last two decades, considerable effort has been expended on the search to find conditions under which race is not encoded, without success [7–9]. Studies carried out in the US and Britain, using targets that were 'black' and 'white', showed that the size of the race-encoding effect remained roughly the same over the following experimental manipulations:

1. Targets are discussing race relations versus a race-neutral topic [8].
2. Subject is 'black' or 'white' [7–8].
3. Subject is led to believe that s/he would soon be interacting with the targets [8] (a context that is believed to promote the encoding of individuating information over category-based information [10–11]).
4. Subject operates under cognitive load or not [8].
5. Subject is warned that there would be a recall test [7].
6. A competing dimension was included (e.g. targets differed in both sex and race) [9].
7. Attention is drawn to targets' race ('pay attention to people's race'), or away from it ('pay attention to people's sex') [9].

The evolutionarily derived prediction that race is encoded as a proxy for coalition led to the discovery of the only context known so far to reduce race encoding: a context of conflict in which coalitional membership is uncorrelated with race [21].

The encoding of race was thought to be spontaneous and automatic because the pattern of recall errors that indicates race encoding (Box 1) occurred in the absence of instructions to attend to the race of targets, and across a wide variety of experimental situations. It was thought to be mandatory – encoded with equal strength across all situations – because every attempt to increase or decrease the extent to which subjects encode the race of targets had failed [7–9]. (Box 2) Such results led some to propose that race, sex, and age are 'primary' or 'primitive' dimensions of person perception [12,13], built into our cognitive architecture. Until recently, no context manipulation – whether social, instructional, or attentional – had been able to budge this race effect.

Automatic race encoding is a puzzle

Natural selection would plausibly have favored neuro-computational machinery that automatically encodes an individual's sex and age. For millions of years, our ancestors inhabited a social world in which registering the sex and life-history stage of an individual would have enabled a large variety of useful probabilistic inferences about that individual (e.g. adolescent girl; toddler boy). By contrast, 'race' is a very implausible candidate for a conceptual primitive to have been built into our evolved cognitive machinery. Ancestral hunter-gatherers traveled primarily by foot, making social contact geographically local [14]. Given the breeding structure inherent in such a world, the typical individual would almost never have encountered people drawn from populations genetically distant enough to qualify as belonging to a different 'race' (Fig. 1). If individuals typically would not have encountered individuals of other races, then there could have been no selection for cognitive adaptations designed to preferentially encode such a dimension, much less encode it in an automatic and mandatory fashion. Race encoding may be a robust and reliable phenomenon, but it cannot be

Nevertheless, the claim that humanity is not divided into distinct racial types is often met with incredulity – who am I going to believe, scientists or the clear evidence of my senses?

Incredulity reigns because we do perceive race. Using an unobtrusive measure – a memory-confusion protocol developed by Taylor *et al*. [7] (Box 1) – social psychologists found that when adults encounter a new individual, they encode that individual's race, sex and age [7–9] (see [10–13] for review and discussion). These dimensions can be encoded without other individuating information; for example, one might recall that one's new neighbor is a young, white woman, without remembering anything else about her [7–11] – her name, her hair color, her job. Until recently, it appeared that race – along with sex and age – was encoded in an automatic and mandatory fashion.

caused by computational machinery that was designed by natural selection for that purpose.

This means that race encoding must be a side-effect of machinery that was designed by selection for some alternative function. Three proposals have been advanced:
(1) Race encoding is a byproduct of domain-general perceptual/correlational systems [7].
(2) Race encoding is a byproduct of an essentialist inference system that evolved for reasoning about natural kind categories [15–20].
(3) Race encoding is a byproduct of computational machinery that evolved for tracking coalitions and alliances [21].

Is race encoding a byproduct of perceptual/correlational systems?

Many species, including humans, appear to have computational machinery that is well-designed to pick up correlations between perceived features and events [22,23] – indeed, classical conditioning is possible only because such machinery exists [24]. Given that humans rely on vision more than most animals, and that we have good color vision, could it be that race is automatically encoded merely as a byproduct of the ordinary operation of our visual and correlation-detecting systems?

Hirschfeld [16] reviews evidence from genetics and anthropology that undermines the notion that races are biological kinds, merely 'out there' to be discovered by our perceptual systems. Real human variation is continuous, and substantially uncorrelated (Fig. 1). It is true, however, that if individuals who are descended from populations that inhabited mutually distant parts of the globe for many thousands of years – central Africa, northern Europe, east Asia – are juxtaposed, one will see a few phenotypic features that are found more commonly in one group than another, including some that are intercorrelated in a Roschian, family resemblance fashion. In contrast to the ancestral world, the recent advent of long distance transportation leads to individuals descended from disparate populations sometimes living side by side. Perhaps our perceptual/correlational systems are merely detecting these perceptual clusters. On this view, the automatic, mandatory encoding of race is merely a byproduct of the automatic, mandatory encoding of perceptual attributes such as color and shape, which operate equally on all stimuli, whether people or objects. This perceptual byproduct hypothesis is plausible, and was the first hypothesis proposed to explain race encoding in the memory confusion protocol [7].

Several predictions follow directly from the perceptual byproduct hypothesis; evidence from Stangor *et al.* [9] and Hirschfeld [16] weigh strongly against them (see Box 3). In addition to failing these empirical tests, the perceptual byproduct hypothesis has a major limitation: It cannot explain why membership in a racial category – as determined by perceivable, phenotypic surface features – is the basis for making inferences about a person's behavior. Yet one of the most obvious (and disturbing) features of racial thinking is racial stereotypes: inferences about a person's traits, personality, goals, moral dispositions, affiliations, and

---

Box 3. Evidence against the perceptual byproduct hypothesis

Four predictions follow directly from the perceptual byproduct hypothesis. Data from Stangor *et al.* [9] and from Hirschfeld [16] speak against them. Briefly:

Prediction 1. *If race is encoded merely as a byproduct of domain-general categorization processes that cannot help but use color, then subjects will encode the color of objects in an automatic and mandatory manner – even when color differences have no social significance (i.e. they do not indicate race, team membership, personality, status, politics or any other social dimension).*

Using the memory confusion protocol, Stangor *et al.* (Exp. 5) showed that the encoding of color is not automatic and mandatory. They examined whether people automatically encode clothing color (black versus white shirts), when it has no social significance (i.e. does not indicate team, status, coalition, etc.). Targets were black and white women; shirt color was uncorrelated with race. Subjects encoded race, but did not encode shirt color at all (limitations on attention cannot explain this result; when the two dimensions are race and sex, both are strongly encoded [9]). Brewer *et al.* [45: Exp 1] found similar results, even in the absence of competing dimensions (targets all same sex and race).

When judging similarity – especially for human targets – children do not privilege skin color over other perceptual dimensions [16: pp. 93–101].

Prediction 2. *Task demands that decrease or increase the extent to which subjects encode perceptual dimensions in tasks with non-social stimuli should decrease or increase the extent to which subjects encode race.*

Instructions to attend to color (or shape) increase the extent to which it is encoded with non-social stimuli [46,47]. But Stangor *et al.* [Exps 1, 2, note 3] showed that the same task demands have no effect on encoding of race (see Box 2, points 6,7).

Prediction 3. *Perceptual similarity should affect how strongly race is encoded; that is, there should be prototype effects.*

If the encoding of race is merely a byproduct of the low-level, domain-general computational machinery, then one should see prototypicality effects for racial categories, just as one sees for artificial categories [48–50]. E.g., perceived similarity of individuals of a given race should produce powerful effects in categorization experiments, affecting learning, inference, recall, and recognition. Stangor *et al.* (Exp 2) directly tested this prediction and found no prototypicality effects: race was encoded just as strongly in conditions where the same race targets were very different in physical appearance as in those where their physical appearance was very similar.

Prediction 4. *Development of racial categories is not bottom up.*

If racial categories were built inductively from perceptual features in a bottom-up fashion, then recognizing which people are members of each category should be trivial. But it is not: Preschoolers, who know and use racial category terms, are rather poor at telling which people in the world fall into which racial category [16]. Hirschfeld shows their categories are driven by labels, not perceptual features.

---

behavior based on racial category membership. By contrast, the essentialist and coalitional byproduct hypotheses easily explain why category membership – whether racial or not – will support inferences about behavior and traits.

Is race encoding a byproduct of essentialist reasoning?

Although there are important differences between their views, Rothbart and Taylor [15], Hirschfeld [16],

<div style="border:1px solid">

Box 4. Natural kinds and essentialism: three theories

(1) Essentialism and social categories

Rothbart and Taylor [15] propose that the essentialist system can be applied to any social category (including racial ones), and that this occurs whenever cultural beliefs imply that category membership is inalterable and carries inductive potential.

(2) Essentialism and folk sociology

According to Hirschfeld [16], 'humans organize themselves into collectivities and define themselves into social kinds as a function of group membership' (p. 119). In consequence, the human cognitive architecture evolved an intuitive 'theory of society', the function of which is to produce expectations about the skeletal structure of society, including what kinds of people there are. The system uses an essentialist mode of construal. Essentialist reasoning enables (but does not require) the notion that humans are biologically clustered, because this same essentialist mode functions across many domains, and plays a role in generating a folk biology. By pitting race against alternative social categories and perceptual cues, Hirschfeld showed that the inductive inferences of children and adults follow racial category for many traits.

(3) Essentialism and ethnic groups

Gil-White [17] argues that the human cognitive architecture has a module designed for reasoning about living kinds (species), which includes an essentialist inference system, and is activated by cues of category-based endogamy (within-group breeding) and descent-based membership. Because ethnic groups manifest both cues, they would have activated this module under ancestral conditions, entirely as a byproduct of the module's design. Nevertheless, because different ethnic groups have different cultural norms and they (as opposed to clans and other groups) represent norm boundaries, the essentialist inference that category membership predicts clusters of non-obvious shared properties conferred a net selective advantage. As a result, the living kinds module was co-opted by natural selection, which produced a new module designed for reasoning about ethnic groups. It is especially prone to produce inferences about ethnic differences in status, norms, morals, and customs (rather than phenotypic differences). In studies of Kazakhs and Mongols, Gil-White showed that trait induction is often based on the ethnicity of the birth parents, rather than on the ethnicity of the parents who raised a child.

</div>

Gil-White [17] and several others [18–20] have all proposed that inferences about race are a byproduct of an essentialist inference system that is functionally specialized for reasoning about natural kinds (Box 4).

The way people reason about natural kinds – tigers, oysters, gold, oak trees – is different from the way they reason about arbitrarily defined categories, such as 'white things' (a motley category that includes pearls, doves, some sinks but not others, etc.) [25–27]. If two things are both judged to be members of the same natural kind, we infer that they share many properties in common, including nonobvious or even hidden ones. For example, having seen that one zebra has a heart, eats grass, runs fast, and fears lions, I might infer that other zebras do too. Use of a category label can support the judgment that two entities belong to the same natural kind (e.g. someone tells me that each of those animals is a 'zebra'); so can perceptual similarity (both animals are horsey and striped). But perceptual similarity is not necessary for kind membership – an albino zebra is still a zebra. Children and adults act as although natural kinds have defining *essences*, 'underlying natures that make them the thing that they are' (Ref. [26], pp. 1476–1477).

Many features of racial categorization and reasoning become comprehensible if one assumes that they are governed by this inference system. Cross-culturally recurrent features of racial thinking include the following notions (see Hirschfeld [16]): (1) there are different *kinds* of humans; (2) people of different races are different in kind; (3) being of a certain race causes many properties, both physical and non-physical, including nonobvious ones such as inner traits and dispositions (temperament, character, shared 'blood'); (4) possessing the correct 'underlying nature' is what makes one of given race, regardless of perceptual properties; and (5) how you look may be a good clue to your race, but a person may look like one race but 'really' be of another. If the human mind were (mis)interpreting race as indicating membership in a natural kind category, these are precisely the assumptions and inferences that the essentialist system would produce.

The essentialist system doesn't care whether human 'races' *in fact* form natural kinds. The system is activated by particular input conditions, and any stimuli that fit those conditions will be treated as a natural kind. On this view, some of the system's input conditions invite the inference that people are divided into different races, constituting different 'kinds' of people. Just as being told that there are 'dolphins' and 'fish' leads children to encode otherwise similar-looking fishy creatures as members of two different natural kinds [25], being told that there are 'black people' and 'white people' leads children to encode otherwise similar-looking people as members of two different natural kinds [16–18,20]. Perceptual similarity might also feed the system: The clusters of correlated phenotypic traits that one sees in modern societies could be interpreted as distinguishing several different, phenoty-pically defined 'kinds' of people [19]. Indeed, the presence of kind labels may organize and drive perception, grace-fully explaining why children who know racial terms do not sort individuals into racial categories on the basis of the perceptual features used by adults [16] (Box 3).

Empirical work on this topic has explored inference, not encoding (e.g. [16,17]). But the application to race encoding is straightforward. If subsets of people fit the input conditions for the essentialist system, it will swing into action. Once it infers (wrongly, in this case) that people are divided into intrinsically different racial kinds and has acquired some criteria for sorting individuals into racial categories [28], it should cause a person's race to be automatically encoded. A person's race would be auto-matically encoded for the same reason that their sex is – and, indeed, for the same reason that one automatically encodes whether the organism one is looking at is a lion, a bird, or a human. Knowing which natural kind an entity belongs to supports many inferences, and the system has mistakenly mapped arbitrary racial categories onto the conceptual apparatus for reasoning about natural kinds.

Is race encoding a byproduct of coalitional psychology?

According to Kurzban, Tooby, and Cosmides [21], no part of the human cognitive architecture is designed specifically to encode race. The (apparently) automatic and mandatory encoding of race is instead a byproduct of adaptations that evolved for an alternative function that was a regular part of the lives of our foraging ancestors: detecting coalitions

and alliances. Hunter-gatherers lived in bands, and neighboring bands often came into conflict with one another [29–31]. Moreover, there were also coalitions and alliances within bands [32] (a pattern also found in related primate species [33]). To foresee the likely social consequences of alternative courses of action, and to navigate their social world successfully, our ancestors would have benefitted by being equipped with neurocognitive machinery that tracked these shifting alliances.

Computational machinery that is well-designed for detecting coalitions and alliances in the ancestral world should be sensitive to two factors: (1) patterns of coordinated action, cooperation, and competition, and (2) cues that predict – whether purposefully or incidentally – each individual's political allegiances [34–36].

Alliance cues
Like other behaviors, actions that reveal coalitional dispositions are usually transitory. Alliance tracking machinery should therefore be designed to note these rare revelatory behaviors when they occur, and then use them to isolate further cues that happen to correlate with coalition but that are more continuously present and perceptually easier to assay. Because this circuitry detects correspondences between allegiance and appearance, stable dimensions of shared appearance – which might be otherwise meaningless (e.g. dress, dialect, etc.) – emerge in the cognitive system as markers of social categories. Coalitional computation increases their subsequent perceptual salience, and encodes them at higher rates.

Dynamic revision
Patterns of alliance usually change whenever fresh issues arise whose possible resolutions differentially affect new subsets of the local social world. Consequently, coalitions shifted over time, varying in composition, surface cues, duration and internal cohesion. To track these changes, cue validities would need to be computed and revised dynamically: No single coalitional cue (including cues to race) should be uniformly encoded across all contexts. Furthermore, arbitrary cues – such as skin color – should pick up significance only insofar as they acquire predictive validity for coalitional membership [36].

In societies that are not completely racially integrated, shared appearance – a highly visible and always present cue – can be correlated with patterns of association, cooperation, and competition [36]. Under these conditions, coalition detectors may perceive (or misperceive) race-based social alliances, and the mind will map 'race' onto the cognitive variable *coalition*. According to this hypothesis, race encoding is not automatic and mandatory. It appeared that way only because the relevant research was conducted in certain social environments where the construct of 'race' happened, for historical reasons [16], to be one valid probabilistic cue to a different underlying variable, one that the mind *was* designed to automatically seek out: coalitional affiliation [34–36].

Is coalition encoded?
Using the memory confusion protocol, Kurzban *et al.* [21] first showed that people do automatically encode the

coalitional alliances of targets. The targets were males, some black, some white; each made statements suggesting allegiance with one of two antagonistic coalitions. Crucially, race was not correlated with coalitional affiliation.

Subjects encoded coalitional alliance even in the absence of shared appearance cues – merely from patterns of agreement and disagreement. When a shared appearance cue – jersey color – was added, coalition encoding was boosted dramatically, to levels higher than any found for race. [N.B. Jersey color is not encoded at all when it lacks social meaning (Box 3)].

Race as a proxy for coalition?
The results further showed that, as predicted, race encoding is *not* mandatory. When coalition encoding was boosted by a shared appearance cue, there was an accompanying *decrease* in race encoding, which was diminished in one experiment and eliminated in another. Other tests showed that the decrease in race encoding could not be attributed to domain-general constraints on attention.

Subjects had a lifetime's experience of race predicting patterns of cooperation and conflict. The decreases in these experiments occurred in response to only 4 min of exposure to an alternative world where race did not predict coalitional alliance. This is expected if (1) race is encoded (in real life) because it serves as a rough-and-ready coalition cue, and (2) coalition cues are revised dynamically, to reflect newly emerging coalitions. There are many contexts that decrease racial *stereotyping* (inferences); creating alliances uncorrelated with race is the first social context found that decreases race *encoding*.

The coalitional byproduct hypothesis fits neatly with the literature on in group favoritism and outgroup derogation, readily explaining why stereotypes of racial outgroups often include derogatory elements. It dovetails with Sidanius and Pratto's [36] findings about social dominance relations between racial groups. It can also explain Levin's [28] recent discoveries that a deficit in cross-race face recognition: (1) is not universal; (2) does not disappear after exposure *per se* to faces of another race; (3) is rare in coalitional sports fans (where race does not predict team membership); and (4) is caused by excessive focus on facial features that discriminate races – a case of racial (coalitional?) categorization driving the perception of similarity, rather than vice versa.

Is race encoding a byproduct of coalitional psychology or an essentialist system?
The human cognitive architecture contains many functionally specialized computational systems. It would not be surprising to find that it contains a system designed for tracking coalitional alliances coexisting alongside several different essentialist inference systems: one designed for 'living kinds' (species) [18–20], another for social groupings [15–16], and a third for ethnic groups [17]. (see Box 4). But which of these is responsible for our perceptions of race? It is too early to answer that question, but a few observations can be made.

Encoding?

That coalitional context could decrease race encoding was specifically predicted by the coalitional byproduct view; it might also be consistent with Hirschfeld's folk sociology, although not required by it. We note, however, that it is not a natural prediction of the other essentialist views. For example, one's ethnicity (and one's species) remains the same, regardless of which coalition one temporarily joins. If race encoding were a byproduct of a 'living kinds' template [18–20] or an ethnicity module [17], then four minutes of exposure to an orthogonal coalitional conflict should not have decreased race encoding to such low levels. If appearance [19] or cultural beliefs about the immutability and inductive potential of race cause racial groups to be interpreted as natural kinds [15], then race encoding should decrease only as these change. Four minutes of exposure to temporary coalitions (and no change in appearance) should have no effect on such beliefs, and therefore no effect on race encoding.

Folk concepts and cognitive architecture

The use of cognitive science to understand real world cultural phenomena, such as folk concepts of race, is a promising and exciting development [16–20,37]. But it is not without difficulties.

Racial thinking, for example, includes inferences about coalitional identity *and* about phenotypic traits; Hirschfeld's experiments suggest that inferences about these two types of properties do not follow the same logic ([16], pp. 166–180). Inferences about bodily traits reflect those of a 'living kinds' template, whereas inferences about identity appear to follow a more coalitional logic. Thus, different folk beliefs about race can be generated by different inferential machinery. Moreover, folk beliefs about race (or religion or other matters) can be a byproduct of different evolved inference mechanisms in different subpopulations within a culture, depending on the history and local distribution of beliefs [18–20,37]. A facet of 'race' might be understood using a 'living kinds' concept in some places and times; in others, coalitional concepts or ethnicity templates might guide inference. By contrast, inferences about the solidity and motion of rocks should be similar across time and place. This is because the mind contains inference systems *designed* for reasoning about inanimate objects [38,39], whereas reasoning about race is necessarily a byproduct of systems designed for other functions.

Highly articulated models of the underlying inference mechanisms and careful tests will be needed to untangle these issues, but the payoff will be cultural theories that are truly explanatory (and perhaps predictive) because they grow out of an understanding of our evolved human nature.

References

1 Macrae, C.N. and Bodenhausen, G.V. (2000) Social cognition: Thinking categorically about others. *Annu. Rev. Psychol.* 51, 93–120
2 Wheeler, M.E. *et al.* (2001) fMRI study of three cognitive tasks that differentially modulate stereotype accessibility and human amygdala response to racial outgroup faces. *Soc. Neurosci. Abstr.* 27, 317.13
3 Lewontin, R. (1972) The apportionment of human diversity. *Evol. Biol.* 6, 381–398
4 Nei, M. and Roychoudhury, A. (1982) Genetic relationship and evolution of human races. *Evol. Biol.* 14, 1–59
5 Nei, M. and Roychoudhury, A. (1993) Evolutionary relationships of human populations on a global scale. *Mol. Biol. Evol.* 10, 927–943
6 Kaessmann, H. *et al.* (1999) Extensive nuclear DNA sequence diversity among chimpanzees. *Science* 286, 1159–1162
7 Taylor, S. *et al.* (1978) Categorical bases of person memory and stereotyping. *J. Pers. Soc. Psychol.* 36, 778–793
8 Hewstone, M. *et al.* (1991) Social categorization and person memory: The pervasiveness of race as an organizing principle. *Eur. J. Soc. Psychol.* 21, 517–528
9 Stangor, C. *et al.* (1992) Categorization of individuals on the basis of multiple social features. *J. Pers. Soc. Psychol.* 62, 207–218
10 Brewer, M. (1988) A dual process model of impression formation. In *Advances in Social Cognition* (Vol. 1) (Srull, T. and Wyer, R., eds) pp. 1–36, Academic Press
11 Fiske, S. and Neuberg, S. (1990) A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In *Advances in Experimental Social Psychology* (Vol. 23) (Zanna, M., ed.), pp. 1–74, Academic Press
12 Hamilton, D. *et al.* (1994) Social cognition and the study of stereotyping. In *Social Cognition: Impact on Social Psychology* (Devine, P.G. *et al.*, eds), pp. 291–321, Academic Press
13 Messick, D. and Mackie, D. (1989) Intergroup relations. *Annu. Rev. Psychol.* 40, 45–81
14 Kelly, R. (1995) *The Foraging Spectrum*, Smithsonian Institution Press
15 Rothbart, M. and Taylor, M. (1990) Category labels and social reality: Do we view social categories as natural kinds? In *Language and Social Cognition* (Semin, G. and Fiedler, K., eds) Sage
16 Hirschfeld, L. (1996) *Race in the Making*, MIT Press
17 Gil-White, F. (2001) Are ethnic groups biological 'species' to the human brain? Essentialism in our cognition of some social categories. *Curr. Anthropol.* 42, 515–554
18 Sperber, D. (1994) The modularity of thought and the epidemiology of representations. In *Mapping the Mind: Domain Specificity in Cognition and Culture* (Hirschfeld, L. and Gelman, S., eds) pp. 39–67, Cambridge University Press
19 Atran, S. (1990) *The Cognitive Foundations of Natural History*, Cambridge University Press
20 Atran, S. (1998) Folk biology and the anthropology of science. *Behav. Brain Sci.* 21, 547–611
21 Kurzban, R. *et al.* (2001) Can race be erased? Coalitional computation and social categorization. *Proc. Natl. Acad. Sci. U. S. A.* 98, 15387–15392
22 Lewicki, P. (1986) *Nonconscious Social Information Processing*, Academic Press
23 Gigerenzer, G. *et al.* (1991) Probabilistic mental models: A Brunswikean theory of confidence. *Psychol. Rev.* 98, 506–528
24 Gallistel, C.R. and Gibbon, J. (2000) Time, rate, and conditioning. *Psychol. Rev.* 107, 289–344
25 Markman, E. (1990) *Categorization and Naming in Children*, MIT Press
26 Medin, D. (1989) Concepts and conceptual structure. *Am. Psychol.* 44, 1469–1481
27 Gelman, S. and Hirschfeld, L. (1999) How biological is essentialism? In *Folkbiology* (Medin, D. and Atran, S., eds) pp. 403–446, MIT Press
28 Levin, D. (2000) Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *J. Exp. Psychol. Gen.* 129, 559–574
29 Keeley, L. (1996) *War Before Civilization*, Oxford
30 Ember, C.R. (1978) Myths about hunter-gatherers. *Ethnology* 27, 239–448
31 Manson, J. and Wrangham, R. (1991) Intergroup aggression in chimpanzees and humans. *Curr. Anthropol.* 32, 369–390
32 Chagnon, N. (1992) *Yanomamo*, 4th edn, Harcourt
33 Smuts, B. *et al.* (1987) *Primate Societies*, University of Chicago Press
34 Kurzban, R. (2001) The social psychophysics of cooperation: Nonverbal communication in a public goods game. *J. Nonverbal Behav.* 25, 241–259
35 Tooby, J. and Cosmides, L. (1988) The evolution of war and its cognitive foundations. *Institute for Evolutionary Studies Technical Report;*

Reprinted in Evolutionary Psychology: Foundational Papers (Tooby, J. and Cosmides, L., eds), MIT Press

36 Sidanius, J. and Pratto, F. (1999) *Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression*, Cambridge University Press

37 Boyer, P. (2001) *Religion Explained: The Evolutionary Origins of Religious Thought*, Basic Books

38 Spelke, E. (1990) Priciples of object perception. *Cogn. Sci.* 14, 29–56

39 Leslie, A. (1994) ToMM, ToBy, and Agency: Core architecture and domain specificity. In *Mapping the Mind: Domain Specificity in Cognition and Culture* (Hirschfeld, L. and Gelman, S., eds) pp. 119–148, Cambridge University Press

40 Graves, J. (2001) *The Emperor's New Clothes: Biological Theories of Race at the Millennium*, Rutgers University Press

41 Marks, J. (1995) *Human Biodiversity: Genes, Races, and History*, Aldine de Gruyter

42 Tooby, J. and Cosmides, L. (1990) On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. *J. Pers.* 58, 17–67

43 Brace, C.L. (1964) A non-racial approach toward the understanding of human diversity. In *The Concept of Race* (Montagu, A., ed.), University of Nebraska Press

44 Dunn, L. and Dobzhansky, T. (1946) *Heredity, Race, and Society*, Mentor Books

45 Brewer, M. *et al.* (1995) Person memory in intergroup contexts: Categorization versus individuation. *J. Pers. Soc. Psychol.* 69, 29–40

46 Nosofsky, R. (1986) Attention, similarity, and the identification-categorization relationship. *J. Exp. Psychol. Gen.* 115, 39–57

47 Nosofsky, R. (1987) Attention and learning processes in the identification and categorization of integral stimuli. *J. Exp. Psychol. Learn. Mem. Cogn.* 13, 87–108

48 Rosch, E. *et al.* (1976) Structural bases of typicality effects. *J. Exp. Psychol. Hum. Percept. Perform.* 2, 491–502

49 Anglin, J. (1976) Les premiers termes de reference de l'enfant. In *La Memoire Semantique* (Ehrlich, S. and Tulving, E., eds) Bulletin de Psychologie

50 Homa, D. and Cultice, J. (1984) Role of feedback, category size, and stimulus distortion on the acquisition and utilization of ill-defined categories. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 83–94