Title

How malleable is categorization by race? Evidence for competitive category use in social categorization.

Journal

Journal of personality and social psychology, 107(1)

ISSN

0022-3514

Authors

Klauer, Karl Christoph
Hölzenbein, Fabian
Calanchini, Jimmy
et al.

Publication Date

2014-07-01

DOI

10.1037/a0036609

Peer reviewed

Journal of Personality and Social Psychology
2014, Vol. 107, No. 1, 21–40

0022-3514/14/$12.00 http://dx.doi.org/10.1037/a0036609

How Malleable Is Categorization by Race? Evidence for Competitive Category Use in Social Categorization

Karl Christoph Klauer and Fabian Hölzenbein
Albert-Ludwigs-Universität Freiburg

Jimmy Calanchini and Jeffrey W. Sherman
University of California, Davis

We contrast 3 theoretical viewpoints concerning the factors affecting social categorization by race: (a) the classical theory of social categorization highlighting the role of a priori accessibility and situational factors, (b) the classical theory augmented by a principle of competitive category use, and (c) competition between race (but not gender) and coalition with race (but not gender) encoded only as a proxy to coalition. Study 1 documents a confound that renders important portions of previous research difficult to interpret. In Studies 2 and 3, race categorization was stronger than categorization by more weakly accessible categories when situational support in terms of topic relevance was comparable across categories. A situational focus on race further increased race categorization. Race categorization was reduced in the presence of strongly cued cross-cutting coalitions. Race categorization also was depressed when situational factors promoted comparative processing of cross-cutting categories while cues to potential coalitional divisions were held constant (Study 4). Accessibility, topic relevance, and cuing cross-cutting coalitions had the same effects on gender categorization as found for race categorization (Study 5). Taken together, the results suggest that classical theories of social categorization have to be augmented by a principle of competitive category use that is not limited to a competition between race and coalition.

Keywords: social categorization, race categorization, category accessibility, category fit

Supplemental materials: http://dx.doi.org/10.1037/a0036609.supp

Basic research in category learning has repeatedly demonstrated effects of categorization on perceptual discriminations and similarity judgments so that differences among objects that fall into different categories are exaggerated and/or differences among objects that fall into the same category are minimized (e.g., Goldstone & Hendrickson, 2010; Harnad, 2005; Livingston, Andrews, & Harnad, 1998; Newell & Bülthoff, 2002). According to Tajfel (1982), social categorization similarly leads to the accentuation of intergroup differences and intragroup similarities, increasing the perceived similarity of members of the same category and the perceived differences between members of different categories. Categorization thereby helps perceivers to process information-rich social environments in an efficient manner and to generalize existing knowledge to new stimuli (Bruner, 1957), encouraging the use of information associated with categories rather than with individuating features. For familiar categories such as age and gender, such associations are in part shaped by a lifetime of exposure to stereotypes and prejudice (Banaji & Greenwald, 2013), associations that are so well learned that they are spontaneously activated upon encountering members of the categories (Devine, 1989; Fazio & Dunton, 1997).

For such reasons, categorization figures importantly in interventions designed to reduce stereotyping and prejudice. In the common ingroup identity model, categorizing two separate groups, an ingroup and an outgroup, into an inclusive superordinate ingroup reduces intergroup bias (Gaertner, Rust, Dovidio, Bachman, & Anastasio, 1996). Similarly, making multiple, cross-cutting social categories salient can lead people to perceive a shared social identity with outgroup members, reducing bias (Crisp & Hewstone, 2006). For example, assigning participants to mixed-gender groups on an arbitrary basis reduced gender discrimination (Deschamps & Doise, 1978, Study 2), suggesting that "when a 'natural' and 'strong' categorization is crossed with one which is 'artificial' and 'weak', the effect of the latter is to decrease the discriminatory effects of the former" (Deschamps & Doise, 1978, p. 152; but see Park & Judd, 2005). The same technique has also been found successful in reducing implicit racial bias (Van Bavel & Cunningham, 2009). This suggests that the cuing of a weak categorization cross-cutting a strong categorization may reduce the use of the strong categorization—a possibility that we refer to as competitive category use.

Karl Christoph Klauer and Fabian Hölzenbein, Institut für Psychologie, Albert-Ludwigs-Universität Freiburg; Jimmy Calanchini and Jeffrey W. Sherman, Department of Psychology, University of California, Davis.

The research reported in this article was supported in part by Deutsche Forschungsgemeinschaft Grant Kl 614/31-1 to Karl Christoph Klauer and by the Anneliese Maier Research Award of the Alexander von Humboldt Foundation, Germany, to Jeffrey W. Sherman. We thank Robert Kurzban for making available the discussion statements and giving permission to reproduce them in this article (online supplemental materials).

Correspondence concerning this article should be addressed to Karl Christoph Klauer, Institut für Psychologie, Albert-Ludwigs-Universität Freiburg, D-79085 Freiburg, Germany. E-mail: christoph.klauer@psychologie.uni-freiburg.de

Comparing Three Models of Racial Categorization

The purpose of this article is to contrast three theoretical viewpoints from the literature on social categorization bearing on categorization by race: (a) the classical theory highlighting the roles of accessibility and situational factors (e.g., Blanz, 1999; Oakes, Haslam, & Turner, 1994), (b) the classical theory augmented by a principle of competitive category use (e.g., Crisp & Hewstone, 2006; Klauer, Ehrenberg, & Wegener, 2003; van Knippenberg & Dijksterhuis, 2000), and (c) competition between race (but not gender) and coalition with race (but not gender) encoded only as a proxy for coalition (Cosmides, Tooby, & Kurzban, 2003; Kurzban, Tooby, & Cosmides, 2001).

Classical theories of social categorization assume, following Bruner's (1957) conceptualization in terms of accessibility and fit, that categorization is a function of chronic factors, a priori accessibility, and a host of situational factors (e.g., Blanz, 1999; Brewer, 1988; Fiske & Neuberg, 1990; Oakes et al., 1994; van Knippenberg, van Twuyver, & Pepels, 1994). Accessibility is defined as "the readiness with which a stimulus input with given properties will be coded or identified in terms of a category" (Bruner, 1957, p. 133). Accessibility is a function of frequency and recency of category use. Due to their frequent and almost ubiquitous use in contemporary society, race, gender, and age are therefore traditionally seen as highly accessible dimensions of categorization. Situational factors include variables such as context relevance, distinctiveness, priming, and perceiver goals that can raise accessibility of a categorization above its chronic level in a given situation, as well as variables that regulate the fit or match between a given categorization and the stimulus persons. Fit can be considered at the level of individual target persons as the degree of match between category-associated stereotypical expectancies and the target's attributes (e.g., Brewer, 1988; Fiske & Neuberg, 1990) or at the level of group organization in terms of the meta-contrast ratio. The meta-contrast ratio is high to the extent that average intercategory differences are larger than average intracategory differences (e.g., Blanz, 1999; Oakes et al., 1994). In other words, fit is high in a given situation when the perceived within-category similarities are larger than between-categories similarities.

As pointed out by Macrae, Bodenhausen, and Milne (1995, p. 398), "these factors and a host of others drive the categorization process, determining when one category rather than another assumes dominance in mental life." Thus, category use is competitive in the sense that one category usually comes to dominate the impression-formation process. In the classical theory, it is, however, not addressed whether the competition is a passive one in which one category simply comes to exceed another one in terms of salience and is then preferentially used or whether the competition is accompanied by a more active kind of inhibition so that raising the salience of one category decreases the activation of competing categories. The second theoretical model that we consider augments classical theories by a principle of competitive category use of the latter kind. The principle postulates that the cuing of a weak categorization cross-cutting a strong categorization reduces the use of the strong categorization.

Many of the successful interventions to reduce bias involve coalitional cues, cuing alliances that transcend or cross-cut boundaries of salient categories such as race and gender categories for which prejudice is to be reduced (e.g., Gaertner et al., 1996; Van Bavel & Cunningham, 2009). Recent work by Kurzban and colleagues (2001; Cosmides et al., 2003) suggests that the cuing of cross-cutting *coalitions* may in fact be the decisive factor, at least in reducing race effects. Kurzban and colleagues argued that race encoding is not automatic and mandatory; instead, the construct of race is argued to be one probabilistic cue to a different underlying variable, one that the mind is designed to automatically seek out: coalitional affiliation (Cosmides et al., 2003). In this view, race is encoded only as a proxy for coalition. Consequently, when valid and salient cues to coalitions cross-cutting racial boundaries are offered, race categorization is quickly abandoned in favor of categorization by the cued coalitions, leading to reduced race categorization. Category competition is thus postulated to occur between race and alternative coalitional categorizations. As further explained by Kurzban et al., competitive category use would, however, be expected to be more limited in this view. For example, it would not be expected to occur in gender categorization when cross-cutting coalitions are cued because gender is seen to be a natural and automatically applied category that should not be subject to competition with coalitional categorizations.

Evidence From the "Who Said What?" Paradigm

Empirical evidence for these theoretical viewpoints has primarily been collected in the framework of the "Who said what?" (WSW) paradigm designed by Taylor, Fiske, Etcoff, and Ruderman (1978). The WSW paradigm consists of a discussion phase and an assignment phase. In the discussion phase, participants observe a discussion among members of different social categories. In the subsequent assignment phase, they are asked to match each discussion statement with its speaker. Many errors are typically made in this assignment task, but they often reflect the category boundaries. That is, a statement is more likely to be falsely attributed to a member of the speaker's category than to a member of a different category. For example, a woman's statement is more likely to be erroneously assigned to another woman than to a man. The relative frequency of within-category errors versus between-categories errors (i.e., the error-difference measure) is considered a measure of the amount of social categorization that occurred in processing the observed discussion (a correction for guessing is necessary; see the online supplemental materials).

Consider first the classical theory highlighting the role of accessibility and situational factors. Regarding the role of accessibility, Stangor, Lynch, Duan, and Glass (1992) attempted to raise the accessibility of race by short-term interventions such as priming and instructions to attend to race. But these interventions were not successful, nor were they successful when targeted at gender categorization. This is traditionally attributed to a ceiling effect: It may be difficult to further raise the accessibility of categories such as race and gender that are highly accessible to begin with (Blanz, 1999; Stangor et al., 1992; van Twuyver & van Knippenberg, 1995). A related prediction is that categorization by race should be stronger than categorization by a more weakly accessible dimension when situational factors support (or do not support) either categorization to a comparable extent. But race categorization and categorization by a weaker dimension have never been directly compared within a single study with situational support for both dimensions held constant. On the other hand, comparing across studies, race categorization has generally been strong and robust in

the WSW paradigm and stronger than categorization by weaker dimensions (e.g., Stangor et al., 1992).

Regarding the role of situational factors, Kurzban et al. (2001) and Cosmides et al. (2003) reviewed the results from three papers that addressed the role of situational factors in race categorization in the WSW paradigm. They concluded that "until recently, no context manipulation—whether social, instructional, or attentional—had been able to budge this race effect" (Cosmides et al., 2003, p. 174). We update the literature review in the Appendix, including five additional studies, and conclude that the evidence on the extent to which race effects are malleable by context factors in the WSW paradigm is in fact mixed and inconclusive.

Two papers bear on the possibility of competitive category use. Vescio, Judd, and Kwan (2004, Experiment 1) contrasted conditions in which either race or gender were the only categories that varied (single conditions) with a condition in which race and gender were crossed (crossed condition). They found that both race and gender categorizations were reduced in the crossed condition relative to the respective single condition (but see the online supplemental materials for a discussion of problems with the error-difference measure that may have led to an underestimation of the strength of categorization in the crossed condition relative to the single conditions). Although these results are consistent with competitive category use and perhaps best explained by it, it is difficult to decide in the absence of a neutral baseline whether the results reflect reduced categorization in the crossed condition due to competitive category use or increased categorization in the single conditions due to enhanced salience of a given categorization when it is the only one that is varied while the other one is held constant. Note that Vescio et al.'s studies were not designed with the intention to resolve this issue. Note also that the race categories used by Vescio et al. were Asian and White so that it is unclear how the results relate to the evolutionary approach proposed by Kurzban et al. (2001). According to Kurzban et al., racial categories are encoded if and when they stand as a proxy for coalition, and it is not clear whether Kurzban and colleagues would expect the contrast between Asian and White to function as a proxy for coalition as Black and White are assumed to do.

In other related research, Kurzban et al. (2001) crossed race and coalitional categories in a series of studies. Participants followed a heated discussion of two rival basketball teams. The discussion statements were antagonistic and coalitional. In their Experiments 1, 2, 5, and 6, team membership and race were crossed orthogonally. In Experiments 3 and 4, team membership and gender were crossed orthogonally for comparison purposes (with race held constant).

In some of the experiments (Experiments 2, 4, and 6), a visual cue distinguished the two basketball teams: Members of the two teams wore shirts of different colors. In these experiments, there were salient and unambiguous verbal cues (i.e., the coalitional statements) and visual cues to coalition. In the other experiments (Experiments 1, 3, and 5), the members of the two teams were not distinguished by a visual cue, and team membership had to be inferred from statement order: Members of the two teams alternated in making statements. In these experiments, the cuing of cross-cutting coalitions was deemed weaker.

Race categorization was reduced and in one experiment erased (Experiment 6) given strong cues to cross-cutting coalitions, whereas gender categorization was much more robust against cuing cross-cutting coalitions. Kurzban et al. (2001) interpreted these results as showing that race but not gender is a coalitional cue that can by supplanted by stronger cues to coalition. We come back to these findings in Studies 1 and 5.

Interpreting the "Who Said What?" Paradigm

One general problem with the WSW literature is that the interpretation of results rests on an uncritical interpretation of the error-difference measure as a direct and process-pure index of social categorization. Without doubt, the error-difference measure is a function of spontaneous social categorization: If the speaker's category membership is spontaneously activated and used in understanding and interpreting a discussion statement, this should create a link in memory between that statement and the speaker's category. Retrieving this link from memory would later prompt a within-category error in the assignment phase if the speaker herself or himself was forgotten (or, with smaller likelihood, a correct assignment by chance) and prevent a between-categories error, thereby directly amplifying the error-difference measure. However, this explanation is only part of the story because the extent and strength of statement-category links formed in encoding, making sense of, and structuring the discussion statements (henceforth referred to as category memory) are not the only factor that shapes the error-difference measure.

In fact, the error-difference measure is shaped by many factors and processes. In particular, its size hinges

- on the extent to which the different to-be-assigned discussion statements are remembered as having been seen (i.e., on the extent of item memory),
- on the extent to which there is also memory for the speakers of the remembered statements (i.e., on the extent of person memory),
- on the extent to which at least the category of the speaker of a remembered statement is remembered if the speaker is not remembered (i.e., on the extent of category memory), and
- on a number of different strategic guessing processes used to make reasonable assignments in the case of memory gaps reconstructively (e.g., on reconstructive category guessing, which we describe below).

Each of these processes affects the pattern of assignments in the WSW paradigm and the error-difference measure in particular (Klauer & Wegener, 1998). However, only category memory is unambiguously linked to the spontaneous activation and use of the categories in question in encoding discussion statements. Thus, if a manipulation has an effect on the error-difference measure, then it is unclear whether the observed effect was caused by an effect on social categorization (i.e., on category memory) or on one or more of the other involved processes (such as item memory, person memory, or strategic guessing) that are not accounted for. Klauer and Wegener (1998) proposed and validated a mathematical model that is capable of disentangling the contributions of these different processes and, thus, isolating a measure of social categorization that controls for the contribution of these other processes.

The purpose of the present studies was to provide more direct tests of predictions that have the potential to discriminate among the three theoretical viewpoints. We focus on race categorization and, for comparison purposes, on gender categorization and address five questions:

1. *Accessibility:* Is race (gender) categorization stronger than categorization by less familiar coalitional categories when situational factors (cuing, fit, distinctiveness, etc.) support (or do not support) the strong and the weak dimensions to comparable extents?

2. *Context:* Is race (gender) categorization further amplified by situational factors traditionally thought to raise the salience of categories?

3. *Race and coalition:* Is race categorization reduced by cuing coalitional cross-cutting categories?

4. *Limits for race:* Is race categorization robust against cuing cross-cutting categories when cues to potential coalitional divisions are held constant?

5. *Limits for gender:* Is gender categorization robust against cuing coalitional cross-cutting categories?

Questions 1 (accessibility) and 2 (context) test predictions of classical theories. That is, they test for a role of accessibility and situational factors, respectively. Evidence in favor of such a role of accessibility and situational factors would thereby support the first two theoretical points of views, the classical theory and the classical theory augmented by a principle of competitive category use. Question 3 (race and coalition) tests whether Kurzban et al.'s (2001) evidence for competition between race and coalitions can be reproduced when critical confounds in Kurzban et al.'s studies as discussed below are removed. If so, the first theoretical model based on the classical theories of social categorization without competitive category use is ruled out. On the other hand, if there is no evidence for competition between race and coalitions, the second model (classical theory augmented by competitive category use) and third model (race encoded as a proxy to coalition) that predict such competitive category use are questioned. Questions 4 and 5 (limits for race and gender) ask whether competitive category use, if it exists, constitutes a more general principle not limited to race and coalition. Answering Questions 4 and 5 thereby allows one to discriminate between the views that race is encoded only as a proxy to coalition (Model 3) and the classical theory augmented by a more general principle of competitive category use (Model 2), which suggests that all social categorization is subject to competitive category use.

Answering these questions requires removing all of the above confounds and further ones discussed below. This can be achieved by experimental controls coupled with the use of a model of the WSW paradigm that accounts for all the distinct processes that shape responses in the WSW paradigm. Studies 2–5 addressed Questions 1–5. Study 1 probed for a reconstructive guessing confound in Kurzban et al.'s (2001) studies that is especially threatening to their conclusions.

Study 1

As already mentioned, the statements exchanged by the two basketball teams in Kurzban et al.'s (2001) studies were antagonistic and coalitional. Kurzban et al. presented 24 statements in the order shown in the online supplemental materials with alternating team membership of the speakers. That is, one team was assigned the odd-numbered statements, the other team the even-numbered statements. As can be seen in the online supplemental materials, the statements assigned to the two teams differ in gist: The odd-numbered statements frequently refer to a foul by the other team and express complaints that the other team plays too aggressively. In contrast, the even-numbered statements deny that a foul has occurred to begin with and portray the other team as weak and cowardly. The two teams thereby differ in that one team has committed a foul that is discussed by the two teams as well as in team traits (aggressors, complainers) attributed to the teams by each other. If this difference in gist is remembered, then it should often be possible to assign a statement to the correct team even if nothing was encoded about the statement or in relation to it.

In Kurzban et al.'s (2001) Experiments 2, 4, and 6, team membership was signaled by shirt color. There was a gray and a yellow team. Assume that a participant truly has no memory whatsoever of a given statement, let alone of its speaker or his or her team membership or racial identity. If statement content allows the participant to infer the speaker's team membership, whether it likely stemmed from a speaker of the gray team or the yellow team, then it is perfectly rational in this situation to assign the statement to a member of the inferred team. This rational guessing strategy would maximize one's chances of being correct in the absence of helpful specific information from memory.

The problem is that this guessing strategy would lead to a within-team error or, with smaller likelihood, to a correct assignment, but never to a between-teams error. The error-difference measure for team membership is thereby directly inflated if such a rational guessing strategy is used. Klauer and Wegener (1998) termed this guessing strategy *reconstructive category guessing.*

Reconstructive category guessing inflates the error difference for team membership, but it cannot contribute to the error difference for race: Statement content was randomized across speakers' races because, within teams, statements were randomly paired with the men's photos. For this reason, statement content does not systematically covary with the speaker's race, which in turn means that it is not possible to infer the speaker's race from statement content. Taken together, strong team categorization with reduced or no race categorization (Kurzban et al., 2001, Experiments 2 and 6) may reflect the prominent use of a guessing strategy, reconstructive category guessing, rather than category memory.

This confound in Kurzban et al.'s (2001) studies can be stated as follows: The error difference for race is driven by category memory; the error difference for team membership can be driven by both category memory and reconstructive category guessing. Coalition and race are thereby not compared on equal grounds; the error-difference measure is contributed to by substantially different and dissociable processes for the two kinds of categorizations.

There is evidence that respondents in the WSW paradigm tend to use reconstructive category guessing when category membership can be inferred from statement content in order to compensate for memory gaps. The evidence is produced by extending the

assignment phase of the WSW paradigm as proposed by Klauer and Wegener (1998). Intermixed with the discussion statements, new statements called distractors are also shown in this phase. Participants are asked, for each statement, to say first whether it was old (i.e., whether it occurred as part of the discussion) or new. If they respond "old," they are then to assign it to a speaker as in the original WSW paradigm.

Sometimes, a false alarm can occur when a distractor is wrongly classified as old. Because there can be no information about the speaker of a distractor statement in memory (the statement was never made in the discussion), assignment patterns for false alarms are indicative of guessing strategies that occur in the absence of specific memory information. It turns out that when statement content allows one to infer a likely category membership of a potential speaker, then assignments are substantially biased toward the inferred category (Klauer & Ehrenberg, 2005; Klauer & Wegener, 1998; Wegener & Klauer, 2004). The purpose of Study 1 was to test whether team membership can be inferred from statement content in Kurzban et al.'s (2001) studies.

Method

Participants. There were 20 participants (five female, 15 male). Most participants (17) were University of Freiburg students with different majors. Mean age was 21.8 years ($SD = 4.13$). Participants received either partial course credit or a small monetary compensation for participating.

Materials. We used eight pictures of young White men taken from the Internet site of the Clarke University soccer roster. These showed the head and upper part of the body. The men in the photos wore blue team jackets. Photo-editing software was used to remove the team label and to make a parallel version of each photo with jacket color changed to red.

Procedure. Participants took part in individual sessions of about 20 minutes. They read that they were to follow a discussion of two rival basketball teams and that their task was to form an impression of the group and the group members. For each participant, four speakers were randomly sampled from the photos with red jackets and four different speakers from the photos with blue jackets. Participants were told that members of one team wear red jackets and members of the other team blue jackets.

Twelve statements (six odd numbered and six even numbered) were randomly drawn from the set of 24 statements employed by Kurzban et al. (2001), translated into German. Each of these statements was presented for 8.5 s along with a photograph of a member of two basketball teams distinguished by clothing color as in Kurzban et al.'s Experiments 2, 4, and 6 and in the studies described below. As in Kurzban et al.'s studies, members of the two teams alternated in making statements; the six odd-numbered statements were assigned to one team, the six even-numbered statements to the other team, although—unlike in Kurzban et al.—in random order. Following the discussion, participants worked on simple arithmetic problems for 10 minutes. In a subsequent surprise assignment test, the 12 remaining statements that had not been seen were presented one by one, in random order. Participants were asked to assign each new statement to one of the two teams, responding to the question "Which team could have likely made the statement?" All randomizations were made anew for each participant.

Results and Discussion

The 20 participants assigned a total of $20 \times 12 = 240$ statements. Of these, they assigned 147 statements to the team that would indeed have been paired with the statement in Kurzban et al.'s (2001) studies and 93 statements to the other team. That is, even-numbered (odd-numbered) statements were overproportionally assigned to the team that had made other even-numbered (odd-numbered) statements in the brief group discussion.

This difference is significant according to a binomial test ($p < .0003$) as well as according to a two-tailed t test across participants, $t(19) = 2.189$, $p = .04$. Its size is such as to account for much of the team categorization observed by Kurzban et al. (2001) in their Experiment 2. The effect size r reported by Kurzban et al. for team categorization was .79, whereas the effect size of the present effect is .45. Note that in Kurzban et al.'s studies, the component of the error difference due to reconstructive category guessing is likely to be even larger: Participants saw all 24 statements presented in a meaningful order (i.e., in the order reproduced in the online supplemental materials) rather than only a subset of 12 randomly sampled statements presented in a random order.

Taken together, participants are able to infer team membership on the basis of statement content after having followed the group discussion even for statements that were never seen and for which they cannot have encoded a category.[1] In consequence, Kurzban et al.'s (2001) comparison of team membership and race is confounded: The error difference for race is driven by category memory; the error difference for team membership can be driven by both category memory and reconstructive category guessing. As already mentioned, there are additional confounds in the error-difference measure (e.g., in terms of item and person memory), but the confound in terms of reconstructive category guessing is especially threatening given that it produces artifactual differences between the error-difference measures for race and team. Another problem with the error-difference measure that has the potential to compromise comparisons between race and team is described in the online supplemental materials (in the section entitled Correction Factors for Crossed Categories).

[1] A reviewer of a previous submission that reported Studies 1–3 suggested an alternative hypothesis in terms of category memory rather than reconstructive category guessing: In assigning new statements, the participants might sometimes confuse a new statement with a statement that had been shown and then assign the new statement on the basis of category memory for the old statement. However, participants were aware of the fact that they were assigning only new statements, so that simple confusion was not possible. Participants might still assign statements on the basis of similarity with other statements that they had been shown and for which category memory had been encoded. To account for the present effect, it is, however, necessary to make an additional assumption: The average pairwise similarity needs to be greater within even-numbered statements and within odd-numbered statements than for pairs comprising one even-numbered and one odd-numbered statement. This is almost, if not exactly, a restatement of the confound in terms of gist-based differences between the two sets of even- and odd-numbered statements that we describe in the body of the text, and it leads to the same conclusions.

Studies 2 and 3

Studies 2 and 3 test whether race categorization is stronger than categorization by less familiar coalitional categories when situational factors support (or do not support) the strong and the weak dimensions of categorization to comparable extents (Question 1), whether race categorization is further amplified by situational factors (here, topic relevance) that have traditionally been found to raise the salience of a given categorization (Question 2), and whether race categorization is reduced by cuing coalitional cross-cutting categories (Question 3). Questions 1 and 2 test predictions of classical theories of social categorization for race categorization. These tests thereby have the potential to falsify or support the classical theory and the classical theory augmented by a principle of competitive category use, that is, the first two theoretical models of racial categorization contrasted here. Question 3 tests for competitive category use involving race and coalitional categories. Finding competitive category use would rule out the classical theory without a principle of competitive category use (Model 1). An absence of competitive category use would question both the classical theory augmented by competitive category use (Model 2) and the idea that race is encoded as a proxy to coalition (Model 3), both of which predict competitive category use for Studies 2 and 3.

As in Kurzban et al. (2001), participants followed a discussion of two rival basketball teams distinguished visually by clothing color. In both studies, there was a condition—termed the condition with high team relevance—in which the discussion statements were antagonistic and coalitional, referring to a recent match between the two teams. However, statement content did not allow one to infer team membership, removing the reconstructive category guessing confound. In addition, we applied Klauer and Wegener's (1998) mathematical model to disentangle social categorization (i.e., category memory) from the other processes (i.e., item memory, person memory, guessing processes) that shape the pattern of assignment errors in the WSW paradigm.

Situational factors favor team categorization. For example, the topic of the discussion is relevant for team membership but not for race, and topic relevance has been identified as one situational factor (among others such as fit, distinctiveness, etc.; see Brown, 2010, Chapter 3, for an overview) that amplifies categorization (e.g., Klauer et al., 2003; van Knippenberg et al., 1994). In both studies, we also implemented conditions with high race relevance in which there was topic relevance for race. In these conditions, the discussion topic was switched to issues of racial equality in society. Finally, in control conditions, participants discussed topics with comparatively less relevance for race and team membership: either gender roles in society or the situation at the university. The control conditions were termed *conditions with low relevance*.

To summarize: There are three conditions in each of Studies 2 and 3, a condition with high team relevance, a condition with low relevance, and a condition with high race relevance. Topic relevance is manipulated by the discussion topic. In Study 2, the three topics are, in order, (a) the merits of one's own team and weaknesses of the other team, (b) gender roles in society, and (c) issues of racial equality in society. In Study 3, the topic with low relevance (Topic 2) is conditions at the university. The conditions with high team relevance and low relevance provide a conceptual replication of Kurzban et al.'s (2001) coalitional manipulation, that is, they differ in the strength of cues to cross-cutting coalitions (i.e., teams). Just as in the analogous conditions in Kurzban et al., these two conditions lack verbal cues to race. The key differences in those two conditions are the extents to which cues to cross-cutting team categories are available.

Method

Because both studies followed the same procedures with few exceptions, it is efficient to describe them together.

Participants. In Studies 2 and 3, 177 (120 female, 57 male) and 121 (92 female, 29 male) participants, respectively, took part and were randomly assigned to one of the three relevance conditions so that there were 59 and 40–41 participants per condition in Studies 2 and 3, respectively. Most participants (149 and 114 in Studies 2 and 3, respectively) were University of Freiburg students with different majors. Mean age was 23.9 years ($SD = 4.79$) and 22.5 years ($SD = 4.71$) in Studies 2 and 3, respectively. Participants received either partial course credit or a small monetary compensation for participating.

Materials. There were four large pools of 192 statements each, one on gender roles (see Klauer et al., 2003), one on conditions at the university (see Klauer et al., 2003; Klauer & Wegener, 1998), one on issues of racial equality in society, and one on the merits of one's own team and the weaknesses of the other team in respect to the last match between the two teams.

Half of the gender-role statements express a more conservative point of view, the other half a more progressive point of view. Conservative and progressive statements were constructed in similarly worded pairs so that, for each statement, there is a parallel one expressing the alternative attitude toward the same or a very similar aspect of the gender-role issue. For example, the progressive statement "It is only natural that a married man has to iron his shirts himself" finds its more conservative counterpart in "It is somewhat strange if a married man has to iron his shirts himself."

This same principle was followed for all statement pools. Thus, the university statements come in pairs of parallel statements, one member of which expresses satisfaction, the other member dissatisfaction with the same or a very similar aspect of the university, using almost the same words. In the team statements, one member of each pair of parallel statements claims a certain strength for one's own team relative to the other team (e.g., "We were simply in better shape physically than you"), whereas the other member claims the absence of that strength or a weakness in that respect for the other team (e.g., "You were simply in worse shape physically than us") using almost the same words. In the race statements, one member of each parallel pair of statements claims the absence of race-based differences in how a person is treated or perceived in contemporary society (e.g., "Black and White persons doing the same job are usually paid the same"); the other member claims the existence or persistence of such differences using almost the same words (e.g., "Black and White persons doing the same job are often not paid the same"). The race statements are usually cautious and often hedged matter-of-fact statements and opinion statements describing the current state of affairs in racial relations, none of them reflecting strongly prejudiced points of view or

claims (see Maddox & Chase, 2004, for a similar statement pool). Let us refer to statements of the first kind (i.e., for Topics 1–4, in order, conservative, satisfied, own-team aggrandizing, and claiming the absence of race-based differences) as Type 1 statements and to statements of the second kind as Type 2 statements. In addition, the statements in each pool address four different subtopics with 24 statement pairs per subtopic.

Each speaker made four statements that were randomly sampled (for each participant anew) with the restrictions that the speaker made two Type 1 and two Type 2 statements, one statement on each of the four different subtopics, and that the two members of a statement pair never both occurred in the discussion. This and the fact that each speaker made four statements addressing each of the four subtopics ensured that no speaker made contradictory statements on the same or similar issues (e.g., expressing both satisfaction as well as dissatisfaction with the services of the university library) and that speakers' category memberships were not systematically related to statement content, preventing reconstructive category guessing from contributing to the error difference in the WSW paradigm.

For the assignment phase, additional distractor statements were sampled with these same restrictions (i.e., equal numbers of statements from each subtopic, equal numbers of Type 1 and Type 2 statements, the two statements of each pair were never both sampled) from the statement pools from which the statements sampled for the discussion as well as their paired parallel statements had been removed. Note in particular that the two members of a statement pair were never both seen in the course of any participant's experimental session.

We used 24 pictures of young White and Black men taken from the official Internet site of the Mississippi Bulldogs track and field 2011 roster, 12 White and 12 Black. These showed the head and upper part of the body. The men in the photos wore red team jackets. Photo-editing software was used to remove the team label and the producer's label from the jackets and to make a parallel version of each photo with jacket color changed to blue. For Study 3, red jackets were changed to a light green in an attempt to increase the visual discriminability of the two teams.

Each participant saw a random sample of six White and six Black speakers from these photos. Within race, three speakers were randomly sampled from the photos with red (in Study 3, green) jackets, and three different speakers from the photos with blue jackets.

Procedure. Participants took part in individual sessions of about 30 minutes. They read that they were to follow a discussion of two rival basketball teams (in Study 3, two rival basketball teams of the same university)[2] and that their task was to form an impression of the group and the group members. They were told that members of one team wear red jackets (in Study 3, green jackets), whereas members of the other team wear blue jackets, and that each team comprises players with White and Black skin color. Participants were told that the two teams had competed against each other in the last season and were now engaged in a discussion.

Participants then watched the simulated discussion. Each contribution consisted of a written statement along with its speaker's photo. The photo was first presented for 1.5 s, followed by the presentation of the written statement along with the photo for 7.5 s and an interstimulus interval of 0.1 s. Speakers made statements in turns so that, in a first round, all 12 discussants made their first statement followed by a second round of statements by each speaker, and so forth. The order of speakers was randomized in each round anew.

After observing the discussion, participants were shown, in random order, all 48 discussion statements mixed with 48 distractor statements. Their task was to decide first whether the statement had been previously presented by clicking on screen fields labeled *old* and *new*. If they decided old, the 12 speakers were presented in random order in two rows of six speakers, and participants were asked to assign the statement to its speaker by clicking on one of the 12 speakers. After a new decision or a person decision, the next statement appeared and was to be classified.

Data analyses. The data were analyzed using the mathematical model for WSW data with crossed categories proposed by Klauer et al. (2003). The model allows one to disentangle social categorization in the form of category memory from several other memory and guessing processes that also contribute to the error-difference measure and may systematically bias it (Klauer & Wegener, 1998). The modeled processes and parameters are the following:

- Item memory I: the probability that a discussion statement is detected as an old statement or a distractor statement as a new statement.
- Person memory c: the probability that the statement's speaker is remembered for a discussion statement correctly detected as old.
- Category memory d(race), d(team|race), and d(team|not race): in order, the probabilities that race is remembered, team is remembered given that race is remembered, and team is remembered given that race is not remembered with regard to the speaker of a discussion statement correctly detected as old, for which there is no person memory. This is one of several equivalent parameterizations of the 2×2 table for category memory crossing the factors race remembered (yes vs. no) and team remembered (yes vs. no). Category memory for team can be computed as $d(\text{team}) = d(\text{team}|\text{race})d(\text{race}) + d(\text{team}|\text{not race})(1 - d(\text{race}))$.
- Item status guessing b: the probability that a statement is guessed to be old if there is no item memory for the statement.
- Category guessing a(Team 1) and a(Black): the probability of guessing that the speaker was from Team 1 rather than Team 2 and the probability of guessing that the speaker was Black rather than White, respectively, in guessing without memory. The probabilities to guess

[2] In Study 3, the rival teams were said to be teams at the same university, competing against each other within the university, rather than teams from different universities. In Study 3, we used the university-statement pool, which might have made the contrast between two different universities salient, and we wanted to avoid superimposing a possibly salient categorical contrast in terms of two different universities on the contrast between the two rival teams.

Team 2 and White are given by 1 − *a*(Team 1) and 1 − *a*(Black), respectively.

In applying Klauer et al.'s (2003) mathematical model for crossed categories, a number of decisions need to be made. For example, for each kind of memory parameter, the same or separate parameters can be used as a function of the speaker's team, as a function of his race, and as a function of statement type (Type 1 vs. Type 2), as further explained by Klauer et al. To limit researcher degrees of freedom, we used the simplest model, in which all memory parameters (*I*, *c*, and *d*) are set equal a priori across team and race as well as statement type. At first sight, this conflicts with the frequent finding of own-race bias in face recognition (Meissner & Brigham, 2001), given that all our participants were native Europeans: Own-race faces are better recognized than faces of another, less familiar race. Recognition memory for individuating features of faces is captured by person memory in the WSW paradigm (Klauer, Ehrenberg, & Wegener, 2004), and person memory is typically small to begin with in the WSW paradigm. That is, there is little memory for the individual speaker, whether he or she stems from one's own race or another one, so that there will not be a pronounced own-race bias in the typical WSW study due to a floor effect in person memory.

As in Klauer et al. (2003), only *a* parameters were allowed to vary as a function of statement type in order to capture possible reconstructive category guessing on the basis of stereotypical a priori associations. For example, statements from the race statement pool claiming the absence of race-based differences in how a person is treated or perceived in contemporary society might predominantly be assigned to White rather than Black speakers in category guessing. Equating parameters across statement type was, however, not possible for the *b* parameters for guessing old/new statement status: We found that participants in the condition with high race relevance in Study 2 were more reluctant, when guessing, to call statements old if the statement claimed the persistence of race-based differences than to call statements old if the statement claimed the absence of such differences (Klauer et al., 2003, did not use this pool of statements). Hence, we allowed the *b* parameters to differ between the two kinds of statements in each context. This baseline model was applied in all our studies.

Figure 1 shows a graphical depiction of part of the model. The figure shows the processing tree for Type 1 statements made by a Black speaker from Team 1. It depicts the processing events leading to the six response categories that can arise for Type 1 statements made by Black speakers from Team 1. Analogous processing trees account for the assignments of statements of other types and from other subgroups. The response categories are displayed as rectangular boxes on the right-hand side of the graph. The responses are classified as either a correct assignment when

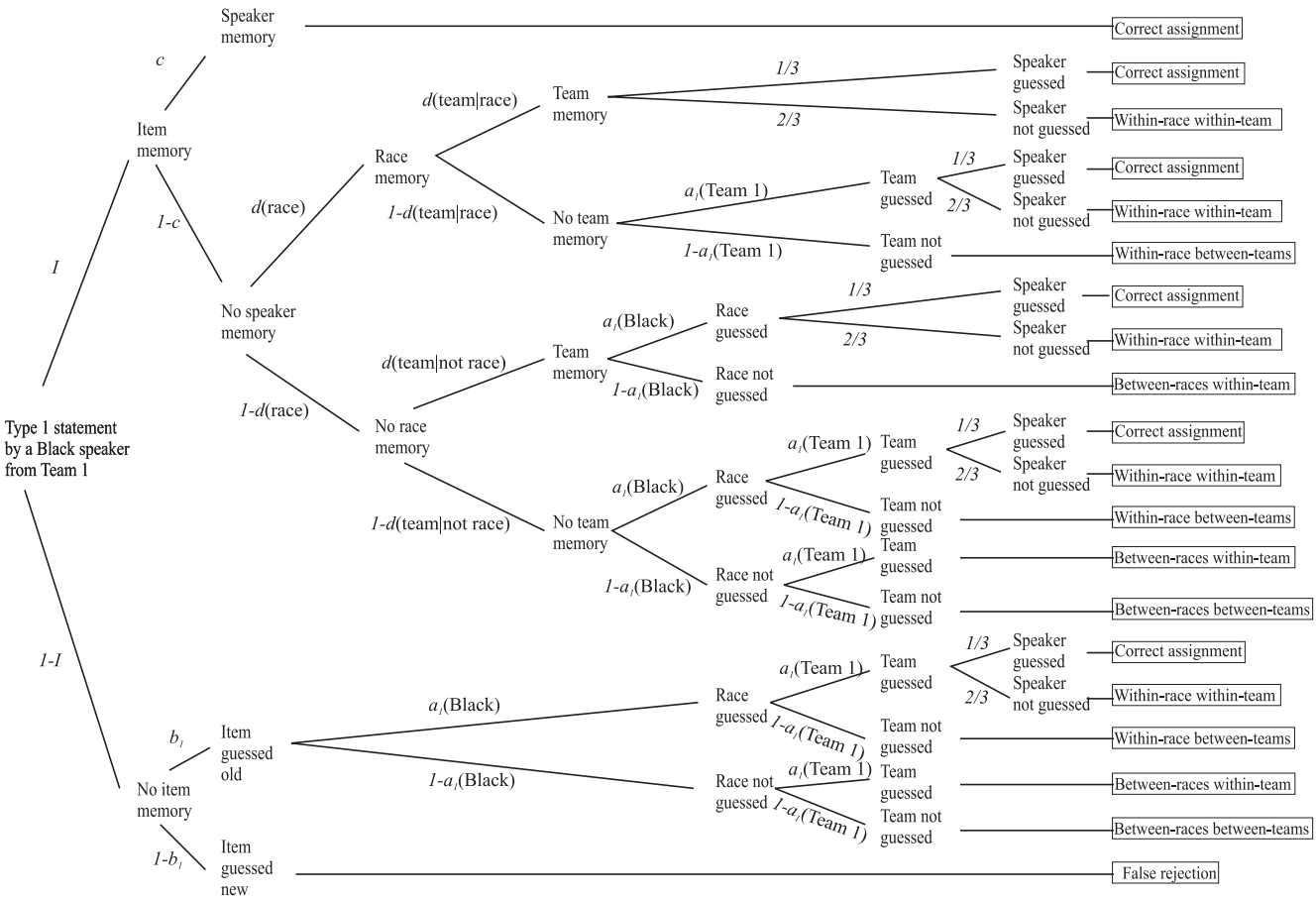

Figure 1. The tree for statements made by a Black member of Team 1. Response categories are shown in rectangles to the right.

the correct speaker is chosen, within-race within-team errors, within-race between-teams errors, between-races within-team errors, between-races between-teams errors, and false rejections (the old statement is falsely judged new). The frequencies with which the different response categories occur are counted separately for each processing tree.

An item will be recognized as old with probability I for item memory or not with probability $1 - I$. In the event of item memory, its speaker will be remembered with probability c for person memory, in which case the statement is assigned to the correct person. In case the correct person is not remembered (with probability $1 - c$), there might still be category memory for the speaker's race with probability d(race) or not with probability $1 - d$(race). In the event of race memory, the speaker's team membership is retrieved with probability d(team|race), and the correct speaker can still be guessed with a fixed probability of 1/3 since there are three discussants in each subgroup in the present studies. If the correct speaker is not guessed with probability 2/3, a within-race within-team error occurs. If team membership is not retrieved with probability $1 - d$(team|race), on the other hand, participants are disposed to assign Type 1 statements to Team 1 rather than Team 2 with probability a_1(Team 1). A value of a_1(Team 1) = .5 corresponds to the absence of a systematic bias to prefer one or the other team in guessing a speaker; a value larger than .5 indicates a preference for assigning Type 1 statements to a speaker from Team 1; a value smaller than .5 indicates a preference for Team 2. With probability $1 - a_1$(Team 1), one of the speakers from Team 2 is guessed. Given that a speaker from the correct team (Team 1) is guessed, it will be the correct speaker with probability 1/3, whereas a within-race within-team error occurs otherwise. Given that a speaker from the false team (Team 2) is guessed, a within-race between-teams error occurs.

Traversing the tree in this way accounts for all the different processing paths by which one of the different responses can occur. The processing tree for distractors is much simpler, as there are no category-memory and person-memory processes involved in assigning items that were not presented. Responses are classified by statement type, source (Black Team 1, Black Team 2, White Team 1, White Team 2, distractor), and the six response categories shown on the right of Figure 1, for a total of 58 response categories (note that there are only five response categories for distractor trees given that a correct assignment to a speaker cannot occur for these).

The model specifies the probabilities with which the different response categories are expected to occur given numerical values for the different model parameters. For example, a correct assignment of a Type 1 statement by a Black Team 1 member is expected to occur with probability

$$\begin{aligned} P(\text{correct assignment} \mid \text{Type 1statement by Black Team 1 member}) = {} & Ic + \\ & I(1-c)d(\text{race})d(\text{team} \mid \text{race})\frac{1}{3} + \\ & I(1-c)d(\text{race})(1-d(\text{team} \mid \text{race}))a_1(\text{Team 1})\frac{1}{3} + \\ & I(1-c)(1-d(\text{race}))d(\text{team} \mid \text{not race})a_1(\text{Black})\frac{1}{3} + \\ & I(1-c)(1-d(\text{race}))(1-d(\text{team} \mid \text{not race}))a_1(\text{Black})a_1(\text{Team 1})\frac{1}{3} + \\ & (1-I)b_1a_1(\text{Black})a_1(\text{Team 1})\frac{1}{3}, \quad (1) \end{aligned}$$

each row of the equation corresponding to one of the six paths of the tree in Figure 1 that lead to a correct assignment.

The observed relative response frequencies estimate these response probabilities. On the basis of the estimated response probabilities and the model equations, model parameters can be estimated. Parameter values are chosen so as to reproduce the estimated probabilities as closely as possible. This is usually done via maximum-likelihood estimation aggregating across items (i.e., statements) and participants (Batchelder & Riefer, 1999).

As recently explained by Judd, Westfall, and Kenny (2012), substantial biases are, however, inherent in analyses that ignore one or the other of these random factors (i.e., items or participants), and the standard approach ignores both. We therefore used a Bayesian approach proposed by Klauer (2010) and Matzke, Dolan, Batchelder, and Wagenmakers (2013) to fit a multilevel extension of the model that treats participants and items as random factors for each model parameter with relevance condition as a fixed factor (see the online supplemental materials for more detail and Rouder & Lu, 2005, and Lee & Wagenmakers, in press, for introductions to Bayesian hierarchical modeling).

This method naturally provides three model checks to assess goodness of fit (Klauer, 2010): Statistic T_1 summarizes how well the model accounts for the pattern of 58 observed assignment frequencies aggregated across items and participants within each condition, corresponding to the goodness-of-fit statistic G^2 used in the traditional approach (Batchelder & Riefer, 1999); T_{2a} summarizes how well the model accounts for the variances and correlations of these frequencies computed across participants as units within each condition; T_{2b} summarizes how well the model accounts for the variances and correlations of these frequencies computed across items as units within each condition. Furthermore, the Bayesian approach yields credible intervals (CIs) for the parameter estimates that can be interpreted much as classical confidence intervals, and it permits one-tailed hypothesis tests for equality between any two parameters. Note that the hypothesis tests reported below all rely on directed hypotheses.

The model checks were satisfactory in Study 2, with none of the test statistics suggesting significant deviations of the model from the data (i.e., all ps $> .05$):

- $T_1^{\text{observed}} = 149.4,\ T_1^{\text{predicted}} = 138.0,\ p = .26;$
- $T_{2a}^{\text{observed}} = 5,904,\ T_{2a}^{\text{predicted}} = 5,575,\ p = .20;$
- $T_{2b}^{\text{observed}} = 3,967,\ T_{2b}^{\text{predicted}} = 3,758,\ p = .10.$

In Study 3, the variance–covariance structure of assignment frequencies across participants was not well accounted for by the multilevel model, indicating significant deviations of the model predictions from the observed data: $T_{2a}^{\text{observed}} = 7{,}612$, $T_{2a}^{\text{predicted}} = 5{,}525$, $p < .01$. When we excluded the participant with the individually worst model fit as determined via traditional methods—the participant's classical goodness-of-fit statistic was $G^2(37) = 88.73$, next lower $G^2(37) = 60.10$—the multilevel model fit the data satisfactorily:

- $T_1^{\text{observed}} = 154.5,\ T_1^{\text{predicted}} = 137.5,\ p = .18;$
- $T_{2a}^{\text{observed}} = 5,830,\ T_{2a}^{\text{predicted}} = 5,416,\ p = .21;$

- $T_{2b}^{\text{observed}} = 2,439,\ T_{2b}^{\text{predicted}} = 2,325,\ p = .12.$

The participant was therefore excluded from all analyses. Including the participant does not otherwise change any of the results.

Results and Discussion

Figure 2 presents team categorization, *d*(team), and race categorization, *d*(race), for each condition in Studies 2 and 3. Results for all model parameters (estimates and CIs) are reported in the online supplemental materials for all studies in the present article. As can be seen in the figure, all of the CIs for race categorization fall well above zero so that there is evidence for race categorization in all conditions in both studies. Based on the CIs, there is also evidence for team categorization in the conditions with high team relevance in both studies and in the conditions with low relevance and high race relevance in Study 3.

Regarding the impact of accessibility on category memory (Question 1), the studies' design affords five comparisons per study between race and team with situational support in terms of topic relevance held constant. For example, in the group with high team relevance, there is little situational support for race categorization. Conversely, in the group with high race relevance, there is little situational support for team categorization. Thus, comparing race categorization in the group with high team relevance and team categorization in the group with high race relevance is a comparison between the two categorizations in which situational factors do not lend strong support to either categorization. There are four such comparisons (1–4) with low situational support and one comparison (5) in which both categorizations are made salient via high topic relevance. For brevity, we refer to the two categorizations as *race* and *team* in the list of comparisons below:

1. Race in the group with high team relevance versus team in the group with high race relevance.
2. Race in the group with high team relevance versus team in the group with low relevance.
3. Race versus team compared within the group with low relevance.
4. Race in the group with low relevance versus team in the group with high race relevance.
5. Race in the group with high race relevance versus team in the group with high team relevance.

In both studies, all five comparisons point in the direction consistent with a role for accessibility (i.e., more race categorization than team categorization; see Figure 2). In both studies, the last three of these comparisons are significant, with the largest Bayesian *p*-value smaller than .02. The nonsignificant comparisons (i.e., Comparisons 1 and 2) are those with race categorization assessed in the group with high team relevance, where competitive category use would be expected to depress race categorization. Nevertheless, in both studies, these two comparisons are descriptively in the expected direction, and in Study 2, both are associated with *p*-values below .10.

Regarding the impact of situational support in terms of topic relevance on race categorization (Question 2), we compared race categorization in the condition with low relevance to the condition with high race relevance. As can be seen in Figure 2, in both studies, race categorization is amplified by topic relevance, although statistically, the evidence is stronger in Study 3 ($p = .02$) than in Study 2 ($p = .09$).

Regarding the impact of coalitional cues on race categorization (Question 3), race categorization is reduced given visual and verbal cues to coalition (the condition with high team relevance) compared to the case where there are only visual cues (the control condition with low relevance) in both Studies 2 and 3. In Study 2, the Bayesian *p*-value for the reduction is $p = .02$; in Study 3, $p < .01$.

As usual, there were moderately sized differences in the memorability of different statement pools (e.g., Klauer et al., 2003). Specifically, the estimates (CIs in parentheses) of item memory in Study 2 for the conditions with high team relevance, low relevance, and high race relevance, in order, were 0.53 (0.46, 0.59),

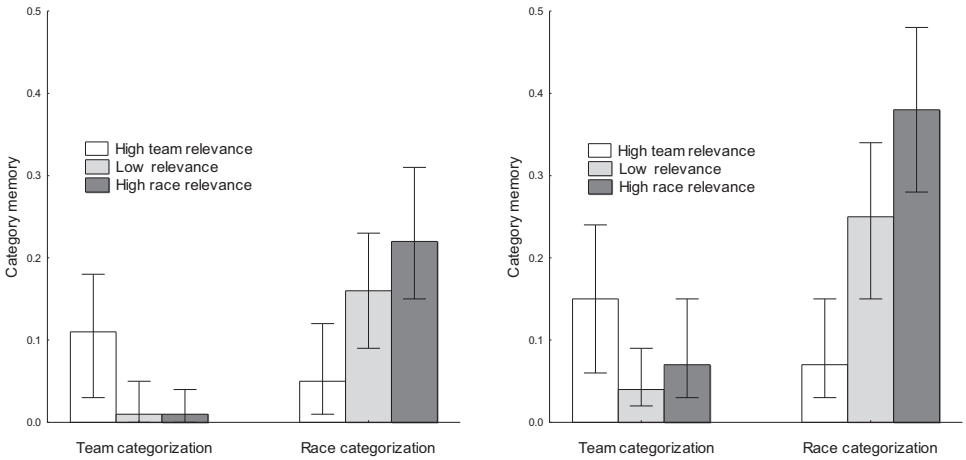

Figure 2. Categorization by team membership and race in Study 2 (left panel) and Study 3 (right panel). The error bars give the 95% credible intervals.

0.70 (0.64, 0.76), and 0.59 (0.52, 0.68); in Study 3, they were 0.50 (0.42, 0.59), 0.55 (0.45, 0.65), and 0.53 (0.44, 0.62).

Taken together, there was evidence for an impact of accessibility and situational factors in race categorization as highlighted in traditional theories of social categorization. In addition there was evidence for competitive category use, suggesting that the classical theories must be augmented by a principle of competitive category use. The evidence for an impact of accessibility was weakest for the comparisons involving race categorization in the presence of strongly cued cross-cutting coalitions, underlining the need to augment the classical theories by a principle of competitive category use.

The findings also agree with Kurzban et al.'s (2001) results in that race categorization was reduced given strong and unambiguous cues to cross-cutting coalitions. Although their evolutionary theory does not necessarily predict the other effects found here, it also is not inconsistent with them. Thus, Studies 2 and 3 rule out classical theories without a role for competitive category use (Model 1), but they are consistent with classical theories cum competitive category use (Model 2) and the coalitional theory of race categorization (Model 3). For example, to account for the effect of topic relevance, the coalitional theory could argue that discussing a race-related topic further raises the preexisting weight given to race as a proxy to coalition.

Studies 4 and 5 investigate the question of whether competitive category use is limited to competition between race and coalition (Questions 4 and 5). For this purpose, Study 4 crossed race with a categorization in terms of prison membership while holding cues to coalitions constant, and Study 5 crossed coalitional categories with gender categories for which Kurzban et al. (2001) predicted and found little competition.

Study 4

In Study 4, we ask whether race categorization is robust against cuing cross-cutting categories when cues to potential coalitional divisions of the speaker group are held constant (Question 4). For this purpose, race categorization was crossed with a weak and unfamiliar categorization in terms of prison membership, and we manipulated the extent of comparative processing directed toward prison categories while holding constant cues to potential coalitional divisions of the speaker group across conditions. Evidence for competitive category use with race and prison categories would support the classical theory augmented by a principle of competitive category use (Model 2). Robustness of race categorization against comparative processing of prison categories on the other hand would support the idea that competition is limited to race and alternative coalitional categories (Model 3).

The coalitional statements used in the previous studies and by Kurzban et al. (2001) promote a comparative focus, so that category differentiation (Doise, 1978) in terms of team membership provides a meaningful and clear partitioning of the group members. This is fostered by a couple of structural features built into these statements. One is that most of the coalitional statements implicitly or explicitly contrast the two teams, implying, for example, that the other team behaves more aggressively than the speaker's team. Moreover, the sentences themselves referred to the teams using pronouns such as *we* and *you*. In order to understand what the sentence says about which team, it is therefore necessary to make oneself aware of the speaker's team membership. This implies that some amount of attention has to be expended on the speaker's team membership for each and every statement. Finally, the statements implement what has been called comparative or structural fit (Oakes, 1987) for the team categorization, meaning that statement content correlates with team membership: Each speaker speaks in favor of his or her own team and against the other team. Fit is even normative in the sense that statement content thereby conforms with stereotypical expectations associated with team membership (i.e., one would a priori expect team members to be loyal to their own team in their statements and less favorable toward members of a rival team). Comparative and normative fit are known to promote category differentiation (Blanz, 1999; Oakes, 1987; Wegener & Klauer, 2004).

Comparative processing in terms of team membership is facilitated where clear visual cues make it easy to identify the speaker's team membership, as in Kurzban et al.'s (2001) Experiments 2, 4, and 6. Comparative processing is hampered where it is difficult to track the speaker's team membership in the first place, as in Kurzban et al.'s Experiments 1, 3, and 5.[3] In Study 4, we implemented a more direct manipulation of comparative processing using cross-cutting categories while holding cues to potential coalitional divisions of the speaker group constant. The question was whether a comparative focus on the cross-cutting categories would nevertheless be sufficient to reduce race categorization.

Comparative focus was manipulated with respect to a categorization in terms of prisons. The speakers were inmates of two juvenile detention centers in the United Kingdom. Half of the inmates were Black, the other half White. Orthogonally, half of the inmates came from the Manchester prison, the other half from the Wakefield prison. The speakers' prison membership was indicated by clothing color.

The interviewed inmates were said to have recently been transferred from the Wakefield prison to the Manchester prison or vice versa due to recent remodeling. They thereby had experience with both prisons. Their statements expressed satisfaction or dissatisfaction with aspects of the prisons. In a condition without comparative focus, participants were asked, on the basis of the seen statements, to form an impression of the conditions in penitentiaries of this kind, thinking about things like the following: In which areas are there problems? How content are the inmates overall? What should be improved in prisons of this type? Thus, a common impression for the type of detention center exemplified by the Manchester and Wakefield prisons was to be formed. In line with this, there was no comparative fit for the categorization in terms of prisons: Each inmate made two statements expressing satisfaction and two statements expressing dissatisfaction with his current prison, one statement from each of four subtopics.

In a second condition, participants were given a comparative focus. Their task was to form an impression of the differences between the Wakefield prison and the Manchester prison, thinking about things like the following: In which areas are there differences? Are there differences in overall contentment between the Wakefield prison and the Manchester prison? What problems are unique to one prison and do not exist in the other prison? Further-

[3] Note that use of category memory for team categories is also hampered in the assignment phase of these experiments: Without visual cues, participants have to remember for at least a few of the speakers which speaker was on which team, a requirement that is not in force in the conditions with visual cues.

more, as with the coalitional team statements, a contrastive perspective was built into the statements shown in this condition, as described below in the Method section.

We implemented several measures to ensure that cues to potential coalitional division of the speaker group were held constant across conditions in Study 4: (a) Choosing prisons as categories cross-cutting racial boundaries made it unlikely in either condition of Study 4 that an inmate would develop a feeling of loyalty to his prison in the same way one might develop a feeling of loyalty for one's sports team. (b) In addition, participants were told that "there are several gangs in each prison that hate each other. They band together based on the type of offense the inmate has committed (e.g., robbery, assault, etc.). One inmate from each gang was interviewed in individual sessions." This provided participants with a clear coalitional structure for speakers (the interviewed inmates) within prisons, making it extremely unlikely that the speakers would be perceived as forming coalitions within prisons in either condition of Study 4 because the speakers were all from different gangs within each prison. (c) Nevertheless, the inmates might be perceived as forming ad hoc coalitions in the course of a group discussion. To counteract such perceptions, the statements were not presented as part of a group discussion but as statements compiled from individual interviews obtained in a field study about the situation at the prisons in both conditions of Study 4. (d) The prisons were furthermore described as high-security institutions in which contact between inmates was restricted and speakers from different prisons had no contact with each other. (e) Finally, the discussion statements spoke to many aspects of the prisons and prison life, but none of them suggested that race or newcomer status was a basis for coalitions in either condition of Study 4.

Because of the very low a priori accessibility of the prison categories, we expected at best a small amount of categorization by prison that might however be significant due to topic relevance for the prisons. As per Question 4 (limits for race), the central question was whether race categorization would be reduced by promoting a comparative focus on the cross-cutting prison categories.

Method

Participants. In Study 4, 82 (68 female, 14 male) participants took part. Participants were students at the University of California, Davis. As in Kurzban et al.'s (2001) studies, participants were primarily European American, Hispanic American (16 Caucasian, 16 Hispanic), and Asian American (43); one participant was African American; six participants chose the category *other*. Mean age was 19.52 years ($SD = 1.47$).

Materials and procedure. The procedures followed those of Study 3. Participants were randomly assigned to two conditions that we termed the condition with comparative focus and the condition without comparative focus. The conditions differed in instructions as already described.

They also differed in the statement pools. In the condition without comparative fit, the statements expressed satisfaction or dissatisfaction with aspects of the speaker's current prison (e.g., dissatisfied: "The food is much too fatty, and it is downright disgusting"). The satisfied and dissatisfied statements were again organized in parallel pairs (e.g., satisfied: "The food is not too fatty, I am happy to say. Anything else would be disgusting."). As before, the statements addressed four subtopics.

Remember that the inmates were said to have recently been transferred from Manchester to Wakefield or vice versa. In the condition with comparative focus, each statement was edited to express dissatisfaction with the new prison relative to the old prison (e.g., "The food in the new prison is much more fatty and disgusting than in the old one") or, for a paired parallel statement, to express satisfaction with the old prison relative to the new one (e.g., "The food in the old prison was much less fatty and disgusting than in the new one"). This implements the structural features pointed to above that exist in the coalitional statements. Specifically, the statements were contrastive, comparing the two prisons, so the reference to an old and a new prison made it necessary to make oneself aware of the speaker's current prison membership in order to understand what the statement says about each prison. To aid in this, a reminder of the mapping of clothing colors on prison membership was visible on the screen (green = Wakefield, formerly Manchester; blue = Manchester, formerly Wakefield) throughout the presentation phase for this condition. Finally, there was comparative fit for the categorization in terms of prison: Inmates of the Wakefield prison effectively expressed a relative preference for the Manchester prison over the Wakefield prison, and inmates of the Manchester prison a relative preference for the Wakefield prison over the Manchester prison (because each statement expressed a relative preference for the speaker's old prison). We acknowledge that this fit manipulation was subtle and perhaps not very meaningful. Nevertheless, it is analogous structurally to the coalitional situation in which each speaker expresses different preferences for the two teams as a function of his team membership. Note also that statement content alone did not allow one to infer the speaker's prison.

Results and Discussion

Model fit is very good:

- $T_1^{\text{observed}} = 89.34,\ T_1^{\text{predicted}} = 92.25,\ p = .57;$
- $T_{2a}^{\text{observed}} = 3,142,\ T_{2a}^{\text{predicted}} = 3,072,\ p = .39;$
- $T_{2b}^{\text{observed}} = 1,561,\ T_{2b}^{\text{predicted}} = 1,528,\ p = .32.$

Figure 3 shows prison and race categorization for each condition. As can be seen in the figure, all of the CIs fall well above zero so that there is evidence for categorization by both dimensions, prison and race, in all conditions.

Regarding the robustness of race categorization against the cuing of cross-cutting categories when cues to potential coalitional divisions of the speaker group are held constant (Question 4), race categorization is significantly reduced in the condition with comparative focus relative to the condition without comparative focus (Bayesian $p < .01$). Considering prison categorization, there are no significant differences between the two conditions ($p = .77$).

Implementing comparative focus with respect to an unfamiliar categorization in terms of prisons thus significantly reduces categorization by race (Question 4). This reduction, $\Delta d(\text{race}) = .11$, was of a similar magnitude as in Study 2 (.11) and Study 3 (.18). This supports the classical theory augmented by a principle of competitive category use (Model 2) but does not agree with the idea that race is encoded as a proxy to coalition (Model 3) inasmuch as the results suggest that competitive category use may not be limited to race and coalition.

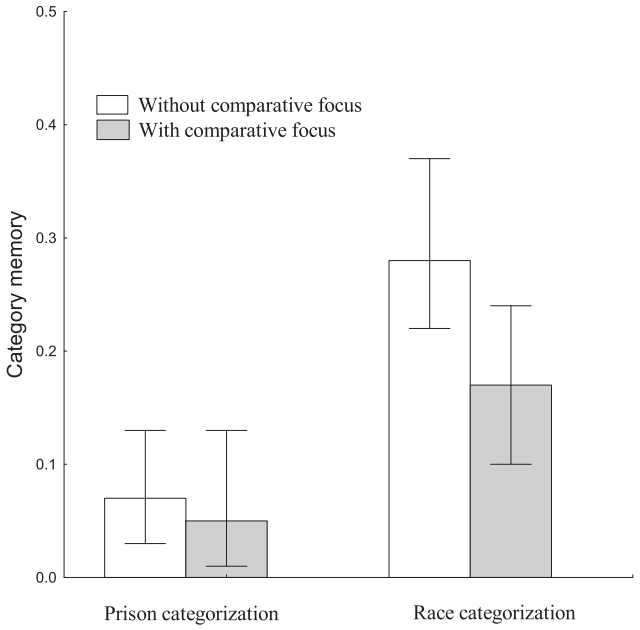

Figure 3. Categorization by prison and race in Study 4. The error bars give the 95% credible intervals.

There was topic relevance for the prison categories in both our conditions, as the statements referred to the prisons in both conditions. This may account for the finding that we observed a significant amount of categorization by prison in all conditions despite the low a priori accessibility of prison categories. The absence of a significant effect of comparative focus on prison categorization agrees well with the idea that perceived comparative fit should have only small effects for categories with low accessibility (Blanz, 1999; see also Oakes, 1987). We return to this issue in the General Discussion. Memorability of the statement pools used in the conditions without and with comparative focus was 0.84 with CI (0.79, 0.88), and 0.66 with CI (0.58, 0.73), respectively.

The results of Study 4 suggest that competitive category use—defined here as the possibility that the cuing of a weak categorization cross-cutting a strong categorization reduces the use of the strong categorization—may not be limited to race and coalition. Study 5 followed up on this suggestion by studying gender and coalition.

Study 5

Study 5 tests whether gender categorization is stronger than categorization by less familiar coalitional categories when situational factors support (or do not support) the strong and the weak dimensions to comparable extents (Question 1), whether gender categorization is further amplified by situational factors (here topic relevance) traditionally thought to raise the salience of categorizations they are targeted at (Question 2) and whether gender categorization is robust against cuing coalitional cross-cutting categories (Question 5). Evidence for an impact of accessibility and situational factors would support the classical theory with and without a principle of competitive category use (i.e., Models 1 and 2). Evidence for a reduction in gender categorization as a function of cues to coalition would question the classical theory without such a principle (Model 1) and would question the coalitional approach (Model 3) in which competitive category use is conceptualized to be more limited.

Kurzban et al. (2001) indeed found that gender categorization, although decreasing significantly, remained far stronger than race categorization even in the presence of cues to cross-cutting coalitions as per their Prediction 5: "Sex will be encoded far more strongly than race, even in contexts in which it is irrelevant to coalition and task" (Kurzban et al., 2001, p. 15388). In fact, gender categorization remained close to ceiling in Kurzban et al.'s condition with strong cues to coalitions cross-cutting gender categories (see Figure 2 in Kurzban et al., 2001). But the basketball context employed by Kurzban et al. may have introduced an additional confound in terms of distinctiveness. The most salient basketball teams, watched by millions of viewers, are male professional and Olympic teams that are almost always racially mixed but are never mixed in terms of gender. Mixed-gender basketball teams being more unusual than mixed-race basketball teams may have endowed gender with a higher level of distinctiveness than race in the basketball context. This may have contributed to the finding that gender categories continued to receive attention in Kurzban et al.'s Experiments 3 and 4 even in the presence of a strongly cued cross-cutting categorization (Experiment 4), where race did not (Experiment 6) or did to a much smaller extent (Experiment 2). Note that distinctiveness is one of the situational variables traditionally seen as modulating the salience of categories (e.g., Higgins, 1996; Taylor, 1981). Even so, the significant decrease in gender categorization that was observed as a function of cues to cross-cutting coalitions is difficult to account for in terms of the coalitional theory, which proposes that gender is a primary and automatically applied category that should not be subject to competitive categorization.

To remove a possible distinctiveness confound, Study 5 was couched in a dodgeball rather than basketball context. Dodgeball is a traditional team sport that most of our German participants are familiar with by having played the game as part of obligatory sports activities in school and as a recreational sport. It routinely and usually involves mixed-gender teams (at least in Germany), removing any special distinctiveness for gender categories that may exist in a basketball context.

Method

The method and procedures closely followed those of Study 3 except where stated otherwise.

Participants. There were 122 (82 female, 30 male) participants who were randomly assigned to one of the three relevance conditions so that there were 40 or 41 participants per condition. Most participants were University of Freiburg students with different majors. Mean age was 22.3 years ($SD = 4.52$). Participants received either partial course credit or a small monetary compensation for participating.

Materials. For the condition with high team relevance, we used the coalitional statements already employed in Studies 2 and 3, edited to accommodate differences between basketball and dodgeball where necessary. A condition with low relevance was

based on the university statements, and a condition with high-gender relevance was based on the gender-role statements.

We used 24 pictures of young men and women taken from the Internet site of the George Mason University track and field teams, 12 men and 12 women. These showed the head and upper part of the body. Photo-editing software was used to change the jacket colors from dark green to a lighter green and to make a parallel version of each photo with jacket color changed to blue.

Results and Discussion

Model fit is good:

- $T_1^{\text{observed}} = 153.0,\ T_1^{\text{predicted}} = 136.15,\ p = .17;$
- $T_{2a}^{\text{observed}} = 5,886,\ T_{2a}^{\text{predicted}} = 5,728,\ p = .40;$
- $T_{2b}^{\text{observed}} = 2,394,\ T_{2b}^{\text{predicted}} = 2,331,\ p = .25.$

Figure 4 shows team and gender categorization in each condition. As can be seen in the figure, all of the CIs for gender categorization fall well above zero so that there is evidence for gender categorization in all conditions. Based on the CIs, there is evidence for team categorization in the conditions with high team relevance and high gender relevance. In the condition with high team relevance, team categorization is as strong as gender categorization. Gender categorization is further increased in the other conditions.

Regarding the impact of accessibility on category memory (Question 1), we tested the five contrasts specified in the results section of Studies 2 and 3 for a role of accessibility. For the last four of these comparisons, gender categorization was stronger than team categorization as expected (largest Bayesian $p < .01$), but as in Studies 2 and 3, this was not the case for the first comparison. This comparison involves gender categorization in the presence of strong cues to a cross-cutting categorization, where competitive category use would be expected to depress gender categorization.

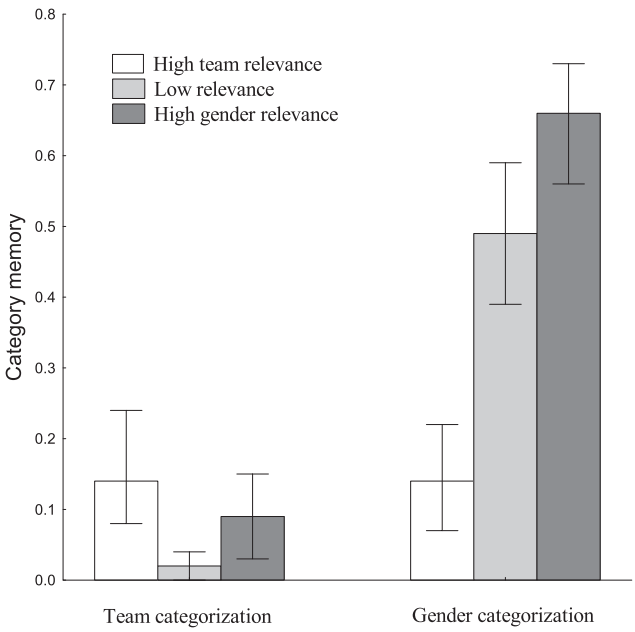

Figure 4. Categorization by team membership and gender in Study 5. The error bars give the 95% credible intervals.

Gender categorization was boosted by topic relevance relative to the condition with low relevance with $p < .01$ (Question 2). But, gender categorization is substantially and significantly reduced given visual and verbal cues to coalition (the condition with high team relevance) compared to the case where there are only visual cues (the condition with low relevance), with $p < .01$ (Question 5).

There were again moderately sized differences in the memorability of different statement pools (e.g., Klauer et al., 2003). Specifically, the estimates (CIs in parentheses) of item memory in Study 5 for the conditions with high team relevance, low relevance, and high gender relevance, in order, were 0.55 (0.48, 0.62), 0.60 (0.53, 0.67), and 0.65 (0.59, 0.73).

Study 5 adduced evidence for competitive category use. Gender categorization was strongly and substantially reduced in the presence of strong cues to cross-cutting coalitions, as was race in Studies 2 and 3. The relative robustness of gender categorization against such cues in Kurzban et al. (2001) lent strong support to the idea that race and gender differ qualitatively, the former being encoded as a proxy to coalition and the latter representing a true and stable primary dimension of categorization. However, once a distinctiveness confound drawing attention to gender is removed, it turns out that there is, in fact, very little basis for claiming such differences between race and gender. Instead, gender categorization decreased just as strongly and to similarly low levels as race categorization in the present studies (see Figures 2 and 4), whereas, in Kurzban et al., it remained far stronger than race categorization and, in fact, close to ceiling in line with their Prediction 5. The present pattern of effects thereby supports Model 2 (comprising a domain-general principle of competitive category use) but is difficult to account for in terms of the coalitional model that postulates domain-specific category competition that is limited to competition between alternative coalitional divisions of the speaker group.

In addition, there was evidence for a role of accessibility and situational factors as highlighted in traditional theories of social categorization. As in Studies 2 and 3, the evidence for accessibility was, however, weakest for a comparison involving gender categorization in the presence of a strongly cued cross-cutting categorization, underlining the need to augment the classical theories by a principle of competitive category use. Gender categorization is stronger overall than race categorization, as was the case in Stangor et al.'s (1992) studies, suggesting that it is somewhat higher in chronic accessibility.

Taken together, the results agree well with classical theories augmented by a principle of competitive category use (Model 2), but they are consistent with neither classical theories without such a principle (Model 1) nor the coalitional theory of categorization (Model 3), which would predict competitive category use to be restricted to a competition between race and coalition.

Correlational Evidence for Competitive Category Use

Competitive category use should also lead to negative correlations within conditions between the strengths of categorization by the competing categories. The multilevel modeling approach pur-

sued here provides estimates of the correlations between categorization by the strong dimension (race, gender) and categorization by the weaker dimension (team, prison) in each study. Specifically, based on the model parameters, two correlations are estimated between (a) the strong categorization (e.g., *d*(race)) and the weak categorization, given that the strong category membership was encoded (e.g., *d*(team|race)), and between (b) the strong categorization (e.g., *d*(race)) and the weak categorization, given that the strong category membership was not encoded (e.g., *d*(team|not race)). The estimates are corrected for the main effects of condition and disattenuated for measurement error. These two correlations, a and b, are estimated across participants and separately across items, yielding four correlations per study. Across the four studies, there are thus 16 estimated correlations between the competing categories.

The correlations came with large credible intervals, and none of them was individually different from zero. However, in line with competitive category use, 11 of 16 correlations were negative, and their mean, −.12 (*SD* = .21), was significantly negative, $t(15) = -2.24$, two-tailed $p = .04$. In other words, the more participants and/or items activate team or prison categories, the less participants and/or items activate race or gender categories.

General Discussion

The present studies addressed three theoretical viewpoints on race categorization: (a) the classical theory of social categorization highlighting the roles of accessibility and situational factors, (b) the classical theory augmented by a principle of competitive category use, and (c) category competition between race (but not gender) and coalition with race (but not gender) encoded only as a proxy to coalition. The studies were organized around five questions.

Questions 1–5

Studies 2, 3, and 5 tested for an impact of accessibility (Question 1) by comparing categorization by race (Studies 2 and 3) and gender (Study 5) on the one hand with categorization by less familiar categories in terms of team membership on the other hand across conditions in which both the strong and the weak categorizations received roughly equal situational support in terms of how relevant they were in relation to the topic of the group discussion. All three studies found evidence for stronger categorization by race and gender than in terms of team membership when situational factors support (or do not support) both categorizations to comparable extents. The evidence was, however, consistently weakest and in fact mostly nonsignificant in conditions in which the strong categorization can be expected to be reduced due to competitive category use (i.e., where context supported categorization in terms of team membership).

The role of situational factors (Question 2) was also addressed in these studies. The question was whether race and gender categorization would be further amplified by topic relevance relative to a control condition (the conditions with low relevance). In all three studies, there was evidence for an amplification due to topic relevance, although the evidence was weak in Study 2 (with $p = .09$). As discussed in the Appendix, Hewstone, Hantzi, and Johnston (1991) did not find an effect of topic relevance in race categorization, but two subsequent studies did (Cabecinhas & Amâncio, 1999; Maddox & Chase, 2004). Cabecinhas and Amâncio (1999) did not find an effect of topic relevance in gender categorization. Note, however, that topic relevance is manipulated by using different statement pools, and statement pools regularly differ in the amount of item memory that they elicit. The error-difference measure is, however, influenced by the extent of item memory, making it difficult to unambiguously attribute effects or the absence of effects to social categorization in these studies. Note also that the coalitional theory can account for effects of topic relevance on race categorization if and when these are confounded with variations in coalitional cues that are sufficiently "obvious and relevant" (Kurzban et al., 2001, p. 15389) to affect the pre-existing weight given to race as a proxy for coalition in the direction consistent with the observed effects of topic relevance.

In Studies 2 and 3, we also tested whether the reduction in race categorization due to the cuing of cross-cutting coalitions found by Kurzban et al. (2001) could be reproduced when critical confounds identified in Study 1 were removed (Question 3). Both studies found race categorization to be significantly and substantially reduced in the presence of strongly cued cross-cutting coalitions.

Study 4 probed whether a similar reduction in race categorization could also be achieved by promoting comparative processing of a weak categorization, prison membership, cross-cutting race when cues to potential coalitional divisions of the speaker group were kept constant (Question 4). Race categorization was significantly and substantially reduced when comparative processing in terms of prison membership was fostered. The coalitional account (Model 3) can explain reduced race categorization when cues to coalitional divisions of the speaker group that cross-cut race are amplified or, alternatively, when cues to race as basis of coalitions are weakened. But a domain-general principle of category competition (as in the present Model 2) is required to account for reduced race categorization when coalitional cues of either kind are held constant, as in Study 4.

The cognitive mechanism driving the results of Study 4 may be attentional: The condition with comparative focus on prison membership may have directed processing away from speaker cues toward the content of the statements that needed to be processed in relation to prison membership. Note, however, that the speaker of each statement was visible on screen for 1.5 s prior to each statement to ensure that speaker cues were fully perceived and processed. In addition, a partial redeployment of attentional resources from race cues to prison cues is, in fact, one mechanism by which competitive category use may operate (i.e., as an attention-directing mechanism), and so, a process explanation based on it is compatible with Model 2 (the classical theory augmented by a principle of competitive category use).

Study 5 probed whether the robustness of gender categorization against strong cues to cross-cutting coalitions found by Kurzban et al. (2001) could be reproduced when critical confounds were removed (Question 5). Gender categorization was significantly and substantially depressed in the presence of strongly cued cross-cutting coalitions in line with the prediction by competitive category use and in contrast to the prediction of the coalitional theory, according to which gender categorization should have remained far stronger than race categorization in all conditions of Study 5.

Taken together, the results support the idea that classical theories of social categorization need to be augmented by a principle of

competitive category use. Competitive category use appears not to be limited to a competition between race and coalition. Its strength is such that it can level a priori differences in accessibility between the weak and the strong categorizations (Studies 2, 3, and 5).

The classical theory cum competitive category activation, our Model 2, paints a more optimistic picture regarding the malleability of race categorization and the possibility to reduce racial stereotyping and prejudice via reduced categorization than does the coalitional theory. In the coalitional theory, the route to down-regulating race categorization is to provide valid cues to alternative coalitions not aligned with racial boundaries. This route is also open for Model 2. In addition, many other factors should be successful according to Model 2, such as manipulations of structural and normative fit for race categories, (effective) manipulations of accessibility, and providing valid cues to alternative, weaker categories not aligned with racial boundaries (see the Appendix for a review of malleability effects in race categorization).

Asymmetry in Competitive Category Use

The present studies addressed race and gender categorization and found evidence for competitive category use, defined as the possibility that the cuing of a weak categorization cross-cutting a strong categorization may reduce the use of the strong categorization. The effects on the weak categorization (team and prison), on the other hand, frequently differed in magnitude from the effects on the strong categorization.

For example, comparing the conditions with high team relevance and the control conditions (low relevance for team and race) in Study 3, the decrease in race categorization was Δd(race) = .18, but the parallel increase in team categorization was only Δd(team) = .11. This asymmetry was stronger in Study 5 with a decrease in gender categorization of Δd(gender) = .35 and a parallel increase in team categorization of only Δd(team) = .13. Similar asymmetries occurred in Study 2. In Study 4, a significant reduction in race categorization was not accompanied by a significant effect on prison categorization. Finally, significant increases in race and gender categorization due to topic relevance were not accompanied by significant effects on team categorization.

Some of these asymmetries are likely due to ceiling and floor effects, but in Studies 2–5, there is a repetitive pattern: Situational support for a relatively weak categorization (team and, even weaker, prison) decreases categorization along a strong dimension (race or gender) more strongly than it increases categorization along the weak dimension.

One explanation can be based on findings from memory research: Participants have better memory for presented stimuli the more familiar they are with that kind of stimuli. Chess experts, for example, can memorize more pieces of a chess game than a novice chess player (Chase & Simon, 1973). Similarly, in levels-of-processing theory (Craik & Lockhart, 1972), the memorability of information is a function of the elaborateness of encoding operations. In particular, elaborate, deep processing implies semantic-associative processing, linking the new information with old knowledge in memory. Deep processing of this kind leans on the existence of a rich knowledge base in memory to which the new information can be related (Olson, 1980). In terms of social categorization in the WSW paradigm, these ideas suggest the hypothesis that encoding of categories is deeper the more familiar and accessible the categories were a priori, other things—such as the amount of attention devoted to the categories—being equal. Consequently, a dimension of categorization associated with a rich set of stereotypes stored in memory would be harmed more from having attention withdrawn from its encoding than a dimension of categorization that is more unfamiliar would profit from having attention directed to it. From a broader perspective, this hypothesis is consistent with the hypothesis that accessibility and situational fit multiplicatively determine the strength of categorization (Blanz, 1999; see also Oakes, 1987), implying that the effects of situational support are larger the more accessible the categories are.

Alternatively, findings from basic research on category learning suggest another explanation of the asymmetry. A number of results on category learning indicate that, for unfamiliar objects and categories, a certain amount of learning is required for the typical effects of within-category compression and between-category differentiation to occur (see Newell & Bülthoff, 2002, for a summary). Thus, categorization effects for less familiar categories may emerge only gradually and sluggishly in the WSW paradigm, even when these are cued as relevant. On the other hand, disengaging from categorizing by familiar categories when cues suggest that another set of categories is much more relevant may be a relatively fast process, accounting for the above asymmetry.

Note, however, that further research is necessary to determine to what extent these hypotheses receive empirical support after possible floor and ceiling effects have been taken into account. Such research also needs to control for additional variables that have the potential to mask the relationship, such as the extent to which perceivers process others in an individuating mode and the overall extent of category-based processing (Klauer et al., 2003).

A Possible Limit on Competitive Category Use

One limitation on competitive category use was already suggested by the discrepancy between Kurzban et al.'s (2001) results and the results of Study 5 regarding gender categorization. Cuing the weak categorization (team membership) effectively reduced the use of the strong categorization (race or gender) where crossing the two categorizations was not unusual (i.e., basketball teams are frequently mixed racially, dodgeball teams are frequently mixed in terms of gender). In contrast, the strong categorization (gender in Kurzban et al.'s, 2001, Experiments 3 and 4) remained distinct where crossing it with the weak categorization ran counter to stereotypical expectancies (i.e., basketball teams are rarely mixed in terms of gender). As already explained, these effects are accounted for by the classical theory cum competitive category use (Model 2) as effects of the situational factor distinctiveness.

In concluding, we speculate that another limit on competitive category use arises in situations where crossing strongly accessible and familiar categorizations defines subgroups that are themselves familiar, so that categorization may directly operate at the subgroup level. For example, considering age and gender, it seems likely that perceivers frequently activate the subgroup identities of target persons, such as whether the person is an old man, a young woman, and so forth (Brewer, 1988; Fiske & Neuberg, 1990; Stroessner, 1996).

Klauer et al. (2003; see also Stangor et al., 1992) found a substantial amount of categorization at the subgroup level when age and gender were crossed. Of course, some amount of categorization at the subgroup level is to be expected if both sets of categories are encoded independently of each other and thereby

sometimes jointly. But the amount of categorization found at the subgroup level exceeded that to be expected on the basis of an independent encoding of both dimensions of categorization in Klauer et al.'s studies. This suggests that perceivers frequently categorized directly in terms of the subgroup identities of the target persons, which would counteract a possible competition between the superordinate categories. For such reasons, we suspect evidence for category competition to emerge most readily for dimensions of categorization (such as race and team or prison membership) for which the subgroups (e.g., Black members of the blue team) are not themselves highly familiar and accessible.

References

Banaji, M. R., & Greenwald, A. G. (2013). *Blindspot: Hidden biases of good people*. New York, NY: Random House.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial processing tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86. doi:10.3758/BF03210812

Biernat, M., & Vescio, T. K. (1993). Categorization and stereotyping: Effects of group context on memory and social judgment. *Journal of Experimental Social Psychology, 29,* 166–202. doi:10.1006/jesp.1993.1008

Blanz, M. (1999). Accessibility and fit as determinants of the salience of social categorizations. *European Journal of Social Psychology, 29,* 43–74. doi:10.1002/(SICI)1099-0992(199902)29:1<43::AID-EJSP913>3.0.CO;2-Z

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. *1,* pp. 1–36). doi:10.1521/soco.1993.11.1.150

Brewer, M. B., & Harasty Feinstein, A. S. (1999). Dual processes in the cognitive representation of persons and social categories. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 255–270). New York, NY: Guilford Press.

Brown, R. (2010). *Prejudice: Its social psychology* (2nd ed.). Chichester, England: Wiley.

Brown, R. J., & Turner, J. C. (1981). Interpersonal and intergroup behavior. In J. C. Turner & H. Giles (Eds.), *Intergroup behaviour* (pp. 33–64). Oxford, England: Blackwell.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64,* 123–152. doi:10.1037/h0043805

Cabecinhas, R., & Amâncio, L. (1999). Asymmetries in the perception of other as a function of social position and context. *Swiss Journal of Psychology, 58,* 40–50. doi:10.1024//1421-0185.58.1.40

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81. doi:10.1016/0010-0285(73)90004-2

Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Science, 7,* 173–179. doi:10.1016/S1364-6613(03)00057-3

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684. doi:10.1016/S0022-5371(72)80001-X

Crisp, R. J., & Hewstone, M. (2006). Multiple social categorization: Context, process, and social consequences. In R. J. Crisp & M. Hewstone (Eds.), *Multiple social categorization: Processes, models and applications* (pp. 3–22). New York, NY: Psychology Press.

Deschamps, J.-C., & Doise, W. (1978). Crossed category memberships in intergroup relations. In H. Tajfel (Ed.), *Differentiation between social groups* (pp. 142–158). London, England: Academic Press.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18. doi:10.1037/0022-3514.56.1.5

Doise, W. (1978). *Groups and individuals: Explanations in social psychology*. Cambridge, England: Cambridge University Press.

Fazio, R. H., & Dunton, B. C. (1997). Categorization by race: The impact of automatic and controlled components of racial prejudice. *Journal of Experimental Social Psychology, 33,* 451–470. doi:10.1006/jesp.1997.1330

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23,* 1–74. doi:10.1016/S0065-2601(08)60317-2

Frable, D. E., & Bem, S. E. (1985). If you are gender schematic, all members of the opposite sex look alike. *Journal of Personality and Social Psychology, 49,* 459–468. doi:10.1037/0022-3514.49.2.459

Gaertner, S. L., Rust, M. C., Dovidio, J. F., Bachman, B. A., & Anastasio, P. A. (1996). The contact hypothesis: The role of a common ingroup identity on reducing intergroup bias among majority and minority group members. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 230–260). Newbury Park, CA: Sage.

Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 1,* 69–78. doi:10.1002/wcs.26

Harnad, S. (2005). To cognize is to categorize: Cognition is categorization. In C. Lefebvre & H. Cohen (Eds.), *Handbook of categorization* (pp. 19–43). Amsterdam, the Netherlands: Elsevier.

Hewstone, M., Hantzi, A., & Johnston, L. (1991). Social categorization and person memory: The pervasiveness of race as an organizing principle. *European Journal of Social Psychology, 21,* 517–528. doi:10.1002/ejsp.2420210606

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York, NY: Guilford Press.

Judd, C., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103,* 54–69. doi:10.1037/a0028347

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika, 75,* 70–98. doi:10.1007/s11336-009-9141-0

Klauer, K. C., & Ehrenberg, K. (2005). Social categorization and fit detection under cognitive load: Efficient or effortful? *European Journal of Social Psychology, 35,* 493–516. doi:10.1002/ejsp.266

Klauer, K. C., Ehrenberg, K., & Wegener, I. (2003). Crossed categorization and stereotyping: Structural analyses, effect patterns, and dissociative effects of context relevance. *Journal of Experimental Social Psychology, 39,* 332–354. doi:10.1016/S0022-1031(03)00017-9

Klauer, K. C., Ehrenberg, K., & Wegener, I. (2004). Components of homogeneity: A multiple-process model of social categorization. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp. 147–160). Hove, England: Psychology Press.

Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the "Who said what?" paradigm. *Journal of Personality and Social Psychology, 75,* 1155–1178. doi:10.1037/0022-3514.75.5.1155

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *PNAS: Proceedings of the National Academy of Sciences, USA, 98,* 15387–15392. doi:10.1073/pnas.251541498

Lee, M. D., & Wagenmakers, E. J. (in press). *Bayesian data analysis for cognitive science: A practical course*. New York, NY: Cambridge University Press.

Livingston, K. R., Andrews, J. K., & Harnad, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 732–753. doi:10.1037/0278-7393.24.3.732

Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology, 69,* 397–407. doi:10.1037/0022-3514.69.3.397

Maddox, K. B., & Chase, S. G. (2004). Manipulating subcategory salience: Exploring the link between skin tone and social perception of Blacks. *European Journal of Social Psychology, 34,* 533–546. doi:10.1002/ejsp.214

Matzke, D., Dolan, C. V., Batchelder, W. H., & Wagenmakers, E.-J. (2013). Bayesian estimation of multinomial processing tree models with heterogeneity in participants and items. *Psychometrika*. Advance online publication. doi:10.1007/s11336-013-9374-9

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7,* 3–35. doi:10.1037/1076-8971.7.1.3

Newell, F. N., & Bülthoff, H. H. (2002). Categorical perception of familiar objects. *Cognition, 85,* 113–143. doi:10.1016/S0010-0277(02)00104-X

Oakes, P. J. (1987). The salience of social categories. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorization theory* (pp. 117–141). Oxford, England: Blackwell.

Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotypes and social reality*. Oxford, England: Blackwell.

Olson, J. C. (1980). Encoding processes: Levels of processing and existing knowledge structures. *Advances in Consumer Research, 7,* 154–160.

Park, B., & Judd, C. (2005). Rethinking the link between categorization and prejudice within the social cognition perspective. *Personality and Social Psychology Review, 9,* 108–130. doi:10.1207/s15327957pspr0902_2

Reynolds, K. J., & Oakes, P. J. (2000). Variability in impression formation: Investigating the role of motivation, capacity, and the categorization process. *Personality and Social Psychology Bulletin, 26,* 355–373. doi:10.1177/0146167200265008

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review, 12,* 573–604. doi:10.3758/BF03196750

Stangor, C., Lynch, L., Duan, C., & Glass, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62,* 207–218. doi:10.1037/0022-3514.62.2.207

Stroessner, S. J. (1996). Social categorization by race or sex: Effects of perceived non-normalcy on response times. *Social Cognition, 14,* 247–276. doi:10.1521/soco.1996.14.3.247

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology, 33,* 1–39. doi:10.1146/annurev.ps.33.020182.000245

Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 83–114). Hillsdale, NJ: Erlbaum.

Taylor, S. E., Fiske, S. T., Etcoff, N. J., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36,* 778–793. doi:10.1037/0022-3514.36.7.778

Van Bavel, J. J., & Cunningham, W. A. (2009). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin, 35,* 321–335. doi:10.1177/0146167208327743

van Knippenberg, A., & Dijksterhuis, A. (2000). Social categorization and stereotyping: A functional perspective. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 11, pp. 105–144). Chichester, England: Wiley.

van Knippenberg, A., van Twuyver, M., & Pepels, J. (1994). Factors affecting social categorization processes in memory. *British Journal of Social Psychology, 33,* 419–431. doi:10.1111/j.2044-8309.1994.tb01038.x

van Twuyver, M., & van Knippenberg, A. (1995). Social categorization as a function of priming. *European Journal of Social Psychology, 25,* 695–701. doi:10.1002/ejsp.2420250608

Vescio, T. K., Judd, C. M., & Kwan, V. S. Y. (2004). The crossed categorization hypothesis: Evidence of reductions in the strength of categorization, but not intergroup bias. *Journal of Experimental Social Psychology, 40,* 478–496. doi:10.1016/j.jesp.2003.09.005

Wegener, I., & Klauer, K. C. (2004). Inter-category versus intra-category fit: When social categories match social context. *European Journal of Social Psychology, 34,* 567–593. doi:10.1002/ejsp.217

Appendix

Is Race Categorization Malleable by Context Factors?

Klauer and Wegener (1998, Table 1) tabulated 30 papers comprising 50 studies on social categorization in the "Who said what?" (WSW) paradigm. Of these, only five papers and nine studies used race as social category compared to, for example, 15 papers comprising 21 studies on gender as social category. We do not know why race has received this relatively small amount of attention, but the picture has not changed noticeably since then. We organize the literature review along the findings discussed by Cosmides et al. (2003) in their Box 2.

1. Topic relevance did not affect race categorization in Hewstone et al. (1991, Experiment 1), as pointed out by Cosmides et al. (2003). There are, however, two studies (Cabecinhas & Amâncio, 1999, Experiment 1; Maddox & Chase, 2004, Experiment 1) that found such an effect of topic relevance. We thank Leda Cosmides for drawing our attention to these two studies.

2. Whether the participant is Black or White had little effect on the pattern of assignment errors for race categories in Hewstone et al. (1991, Experiment 1). Cosmides et al. (2003) also referred to Taylor et al. (1978) to make this point, but Taylor et al.'s participants were all White. It is possible that Cosmides et al. had in mind that there was no evidence for effects of speaker race on the strength of race categorization in Taylor et al.'s Experiment 1: The error difference was about the same size when computed separately for statements from Black and White speakers. Frable and Bem (1985), using mostly White participants (92 of 96 participants were White), reported a large study that did find an ingroup bias for race categorization with more within-race errors than between-races errors for Black than White speakers. Similar indications of effects of participant and speaker race were reported by Cabecinhas and Amâncio (1999), who focused on analyzing within-race errors as a function of participant and speaker race, however, rather than on analyzing the error-difference measure. Relatedly, Vescio et al. (2004, Experiment 1) crossed race and gender and found evidence for a moderation of race categorization as a function of whether or not speakers were in the participants' ingroup with respect to gender. Taken together, results in the very limited literature on the effects of participant and speaker race in race categorization are mixed.

3. Whether or not the participant was led to believe that he or she would soon be interacting with the targets had no effect in Hewstone et al.'s (1991) Experiment 2. Anticipated interaction has been argued to enhance the use of an individuating mode of person perception that early accounts considered to be antagonistic to a category-based mode of processing (e.g., Brewer, 1988; Brown & Turner, 1981; Fiske & Neuberg, 1990). If so, increased individuation should have been accompanied by decreased categorization, which was not found. Note, however, that later approaches in the traditional framework such as the one by Brewer and Harasty Feinstein (1999) or the one by Reynolds and Oakes (2000) abandoned the assumption of strictly opposed principles (see also Wegener & Klauer, 2004).

4. Cosmides et al. (2003) stated, with reference to Hewstone et al. (1991), that race categorization is not affected by whether or not participants work under cognitive load. But neither Experiment 1 nor Experiment 2 by Hewstone et al. comprised a manipulation of cognitive load. It is possible, however, that Hewstone et al.'s manipulation of anticipated interaction just discussed was inadvertently confounded with a manipulation of cognitive load.

5. Whether the participant was told that there would be a memory test or not had no effect on race categorization in Taylor et al.'s (1978) Experiment 1. Note, however, that the traditional views of social categorization do not appear to be committed to strong predictions on whether or not there should be an instructional effect of this kind.

6. Cosmides et al. (2003) stated, with reference to the studies by Stangor et al. (1992), that including a competing dimension (e.g., targets differ in both race and gender) had little effect. But in none of the studies reported by Stangor et al. was there a control condition without competing dimension so that it is difficult to judge what the effect of including a competing gender dimension had on the strength of race categorization.

In a methodologically strong study published after Cosmides et al.'s (2003) paper, Vescio et al. (2004, Experiment 1) did in fact implement control conditions in which race and gender were the only categories that varied (single conditions) and contrasted these with a condition in which race and gender were crossed (crossed condition). They found that both race and gender categorizations were reduced in the crossed condition relative to the respective single condition. We further discuss this study in the body of the article.

7. Asking participants to attend to either race or gender did not influence categorization by race in Stangor et al.'s (1992) Experiment 2. Nor, one should add, was gender categorization in any way affected by the instruction. But it is unclear to what extent participants heeded the instruction in the course of the experiment, and there was no manipulation check. Perhaps the manipulation was simply too weak to exert any effect.

Similarly, a priming manipulation targeted at either race or gender, administered prior to undergoing the WSW paradigm, had no effect in their Experiment 1. As discussed by Stangor et al. (1992), the manipulation may have been too weak to raise the accessibility of race or gender. Another possibility also discussed by Stangor et al. (see also Blanz, 1999; van Twuyver & van Knippenberg, 1995) is that it may be difficult to further raise the accessibility of categories such as race and gender categories that are highly accessible to begin with, suggesting a ceiling effect in accessibility.

(*Appendix continues*)

Finally, Stangor et al. (1992) reported evidence for an effect of level of prejudice on race categorization (Experiment 3). Specifically, although participants did not independently categorize by race, there was evidence that at least the highly prejudiced participants did: The error difference for race was significant among the between-genders errors for the high-prejudice participants, whereas it was not significant for the low-prejudice participants. This difference between high- and low-prejudice participants was itself, however, only marginally significant.

Moreover, there are two studies by Biernat and Vescio (1993) in which racial categories and a discussion topic relevant for race were used. Biernat and Vescio's results are difficult to interpret because their design invites the use of reconstructive category guessing based on statement content, a confound that is analogous to the one further elaborated on in the present Study 1.

In summary, even to date, the pool of studies using race in the WSW paradigm seems much too small to base strong conclusions on it. Moreover, the evidence for the malleability versus nonmalleability of race categorization is in some cases mixed (effects of topic relevance, effects of participant and speaker race), whereas, in other cases (effects of participant and speaker race, memory instructions), traditional views do not appear to be strongly committed to predicting effects of the reviewed manipulations on race categorization. Stangor et al. (1992) reported two failures to manipulate the salience of race (by priming and an attentional instruction), but these manipulations when targeted at raising the salience of gender also failed to affect gender categorization. The most parsimonious explanation of these null findings would probably be that the employed manipulations were too weak to engender strong effects.

Received July 28, 2013
Revision received February 12, 2014
Accepted March 10, 2014 ■