Journal of Experimental Social Psychology 104 (2023) 104423

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Do individuation instructions reduce the cross-race effect? A registered replication of Hugenberg, Miller, and Claypool (2007)☆

Francisco Cruz [a,1], Tomás A. Palma [a,*,1], Emil Bansemer [b], Joshua Correll [b], Sara Fonseca [a], Patrícia Gonçalves [a], Ana Sofia Santos [a]

[a] CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Portugal
[b] University of Colorado Boulder, United States

A R T I C L E I N F O

Keywords:
Face recognition
Cross-race effect
Individuation motivation
Culture
Pre-registered replication

A B S T R A C T

People usually have less accurate memory for cross-race (CR) than for same-race (SR) faces, a robust and consequential phenomenon known as the Cross-Race Effect (CRE). In an influential paper, Hugenberg et al. (2007) showed that the CRE can be eliminated when participants are instructed to individuate CR faces in order to avoid displaying this effect. This finding has received widespread attention, and many studies have attempted to replicate it, with mixed results. In the present research, we attempted to replicate the effect of the individuation instructions in eliminating the CRE (Hugenberg et al., 2007) in two pre-registered experiments in two different cultures – the United States and Portugal. The results of both experiments found no evidence that instructing participants to individuate CR faces eliminates or even attenuates the CRE. Additionally, we also examined and failed to find support for the idea that these individuation instructions are more effective for the participants who report greater contact with CR faces (Young & Hugenberg, 2012). Finally, we also did not find evidence that the cultural setting moderates the effect of the individuation instructions of the CRE. We critically discuss the potential reasons for the lack of impact on the individuation instructions in the CRE and its implications for a prominent motivational account of this effect.

In July 2012, Otis Boone, a Black man, went on trial under suspicion of robbery. The prosecution based the case around the testimony of two white victims who had positively identified the suspect. Though the defense filed a request to present evidence against the accuracy of cross-race identifications to the jury, the motion was denied, and Otis was initially found guilty of the robbery charges and sentenced to 25 years in prison. In 2017, an appeal for a retrial was issued, and Otis was exonerated of all charges, with the help of a brief submitted by the American Psychological Association. The brief mentioned the limitations of eyewitness identification, especially those concerning cross-race suspects, and the importance of disclosing the empirical support for this to the jury so that the decision appropriately weights the evidence provided. Otis Boone's case is only one of the several mistrials that occur based on suspects' misidentifications. This phenomenon has been found in both real world and lab settings: when asked to identify the perpetrator of an event from a lineup, participants consistently misidentify cross-race suspects (Meissner & Brigham, 2001).

The tendency for perceivers to have less accurate recognition memory for cross-race (CR) faces than for same-race (SR) faces is known as Cross-Race Effect (CRE; Malpass & Kravitz, 1969). The CRE is one of the best-replicated phenomena in the face perception literature and has been shown to generalize across several research paradigms and participant populations (Meissner & Brigham, 2001).

In addition to being a theoretically compelling phenomenon, the CRE can result in potentially deleterious consequences of misidentifying members of racial outgroups leading them to feel insulted or stereotyped, while the perceiver can experience feelings of embarrassment or social opprobrium (Brigham & Malpass, 1985). These add to the most serious legal consequences detailed on the court case presented above – convictions based on misidentifications of suspects.

Given the social consequences that the CRE brings forth, studying the psychological determinants of this phenomenon is of great interest as it

☆ This paper has been recommended for acceptance by Dr Rachael Jack.

* Corresponding author at: Faculty of Psychology, University of Lisbon, 1649-013 Lisbon, Portugal.
E-mail address: tapalma@psicologia.ulisboa.pt (T.A. Palma).

[1] Contributed equally to this project.

https://doi.org/10.1016/j.jesp.2022.104423
Received 14 January 2021; Received in revised form 14 October 2022; Accepted 14 October 2022
Available online 27 October 2022
0022-1031/

can contribute to reducing its negative impact. An influential explanation for the CRE is that people have reduced motivation to individuate CR faces (Hugenberg, Young, Bernstein, & Sacco, 2010). Consistent with this explanation, several studies have shown that increasing participants' motivation reduces or eliminates the CRE (e.g., Baldwin, Keefer, Gravelin, & Biernat, 2013; Hugenberg et al., 2007; Young & Hugenberg, 2012). The present work further examines the reliability of a commonly used manipulation to increase motivation - individuation instructions (Hugenberg et al., 2007) - through well-powered pre-registered replications.

In the next paragraphs, we first briefly review the main theoretical accounts of the CRE. Then, we concentrate on the impact of individuation instructions in reducing the CRE, and, finally, we explain the reasons leading to the current replication attempt.

1. What are the causes of the CRE?

Several of the accounts put forward to explain the CRE fall under the designation of expertise theories, which posit that the CRE arises due to differences in the perceptual processing of faces (for an extensive review of the different CRE explanations, see Young, Hugenberg, Bernstein, & Sacco, 2012). Though humans are experts at processing faces, expertise varies across races due to the tendency for people to interact mostly with SR individuals, which will tune the perceptual system to the features found across SR faces. Such an asymmetry in contact leads to differential expertise for SR and CR faces. Differential expertise may explain the CRE either through changes at the level of qualitative processing (i.e., expertise leads to configural processing of SR faces; Gauthier, Williams, Tarr, & Tanaka, 1998; Goffaux & Rossion, 2006; Michel, Caldara, & Rossion, 2006) or representations in memory (i.e., with cross-race faces being more clustered together, hence more likely to be activated together; Leopold, O'Toole, Vetter, & Blanz, 2001; Valentine, 2001). Irrespective of the mechanism, all expertise theories argue that the CRE is due to the lack of adequate perceptual ability to correctly learn CR faces.

However, alternate social cognitive accounts of the CRE have also taken the spotlight. Whereas many of these offer their own unique predictions and explanations as to how the biased processing occurs, they all share one common anchor: the tendency to process CR targets in a categorical manner and SR targets in an individuating manner is the cause of the CRE (Hugenberg et al., 2010). As such, according to social cognitive theories, perceptual learning deficits do not cause the CRE, but rather a difference in the tendency to classify faces by race. Categorical thinking arises when we rely on broad social group membership (e.g., race, sex, age) to process a face, while individuation relies instead on processing characteristics or features that are specific to that target (Young et al., 2012). Categorization impairs the encoding of outgroup faces (i.e., CR faces), leading to a difference in recognition accuracy – the CRE. For example, Levin (2000) attributes this process to an asymmetrical search for features in SR versus CR faces (i.e., differences in quantitative feature search); yet Rodin (1987) argues that this is a result of a differential allocation of attention when encoding faces (i.e., differences on the features searched) – *cognitive disregard model*. Others have posed that, when processing or retrieving CR faces, perceivers are prone to rely on low-effort feelings of familiarity, though they employ more motivated recall strategies when it comes to SR faces (e.g., Meissner, Brigham, & Butz, 2005).

Along with the proposals presented so far, the social cognitive account that has aroused the most interest, instigating a substantial amount of research, is the *categorization-individuation model* (CIM; Hugenberg et al., 2010). According to this perspective, the CRE derives from the tendency to think about outgroup members categorically and search for category-specifying features (e.g., skin tone, nose width, lip thickness), whereas ingroup members are perceived in terms of their individuating features. Thus, if different social cognitions about these two groups are responsible for the CRE (i.e., categorize CR faces but individuate SR faces), then changing these social cognitions may be sufficient to reduce or eliminate the CRE. Specifically, inducing perceivers to individuate CR faces may eliminate the CRE.

Adding to the role of social categorization and individuation motivation, the CIM (Hugenberg et al., 2010) also posits that enhanced perceiver expertise with a group of faces (e.g., SR faces) will facilitate face memory with those faces. However, face expertise translates into accurate recognition only when perceivers are motivated to individuate faces. In this sense, although experience in distinguishing between SR members plays a role in the CRE, both sufficient perceptual expertise and individuation motivation are necessary conditions to successfully discriminate between CR faces.

2. The role of instructions in inducing individuation of CR faces

In an impactful paper (342 citations in Google Scholar at the date of this writing), Hugenberg et al. (2007) tested whether instructing participants to attend to the individuated characteristics and traits of CR faces eliminates the CRE. To the extent that the CRE arises from the tendency to engage in categorical thinking about outgroup members, warning participants of this tendency should motivate them to individuate CR faces and, consequently, lead to the elimination of the CRE. Hugenberg et al. (2007) tested this hypothesis in three experiments.

In their two initial experiments, Experiments 1A (N = 30) and Experiment 1B (N = 146), the authors randomly assigned participants to one of two conditions – individuation and control condition. All participants were told that they would be participating in a face recognition experiment and, thus, they would be seeing 40 faces on the screen and should pay close attention in order to recognize them later on. Participants in the individuation condition, however, were given the following additional instructions designed to induce individuation of CR faces:

> Previous research has shown that people reliably show what is known as the Cross Race Effect (CRE) when learning faces. Basically, people tend to confuse faces that belong to other races. For example, a White learner will tend to mistake one Black face for another. Now that you know this, we would like you to try especially hard when learning faces in this task that happen to be of a different race. Do your best to try to pay close attention to what differentiates one particular face from another face of the same race, especially when that face is not of the same race as you…
> Remember, pay very close attention to the faces, especially when they are of a different race than you in order to try to avoid this Cross Race Effect.

The learning phase was initiated after the instructions and terminated after participants viewed all the target faces (20 Black, 20 White). This phase was followed by a 5–7 min unrelated distractor task, meant to clear working memory. Then, participants completed a recognition task in which they were presented with the previously seen target faces together with the same number of never-seen-before face lures (20 Black, 20 White) and had to indicate whether they saw each face in the learning phase or not.

Hugenberg et al. (2007) obtained an ordinal interaction between instruction and race. That is, they successfully replicated the CRE in the control conditions across both experiments, as they found that participants demonstrated greater sensitivity (d'; Green & Swets, 1966) when recognizing SR (white) than CR (black) faces. However, they successfully eliminated the CRE in the individuation condition (i.e., no significant difference in memory sensitivity between SR and CR faces), suggesting that the instructions given before learning improved memory for CR faces. It is, nonetheless, important to note that the *p*-value of the interaction between instruction condition and face race in Experiment 1A did not reach the conventional significance level of 0.05 (Experiment 1A, $p = .056$, $\eta_p^2 = 0.12$, 90% CI [0.00, 0.31]; Experiment 1B, $p = .023$, $\eta_p^2 = 0.04$, 90% CI [0.003, 0.10]).

Additionally, Hugenberg et al. (2007) devised Experiment 2 ($N = 55$) to rule out the alternative hypothesis that the previous results might be due to a general accuracy motivation instead of a motivation to individuate CR faces. They added a third condition to the design in which they asked participants to pay close attention to all of the presented stimuli, without ever mentioning the CRE:

> Previous research has shown that people are often unreliable when it comes to recognizing previously seen faces. Because of this, eyewitness testimonies are often discredited. Now that you know this, we would like you to try especially hard when learning faces in this task. Do your best to try to pay close attention to what differentiates one particular face from another face…

They found that the CRE was only eliminated in the individuation condition, suggesting that the previously obtained results cannot be attributed to increased processing motivation. However, as in Experiment 1A, the interaction did not reach significance ($p = .055$), which calls for caution in the interpretation of the results.

Hugenberg et al. 's (2007) findings constitute a major theoretical contribution to the literature as they question the central assumption of expertise accounts that the CRE is the consequence of a perceptual inability to adequately learn CR faces. Moreover, this paper inspired the development of the influential *categorization-individuation model* (Hugenberg et al., 2010), which, at the date of this writing, has 550 citations in Google Scholar. Such theoretical relevance sparked several replication studies in the past years, with mixed findings. Specifically, from the 10 studies encountered in a literature review,[2] 2 replicated the critical interaction between instruction condition and face race, while 8 did not replicate such interaction (see Table 1). In the next section, we reflect on these conflicting results and explain why we believe a registered replication of Hugenberg et al.'s (2007) study is necessary for understanding the role of individuation instruction in the CRE.

3. Considerations on conflicting replications

Previous replication studies differ from each other and from the original study in ways that may have contributed to the conflicting results (see Table 1). First, we should note that the primary goal of existing replication studies was to examine the boundary conditions and mechanisms underlying the effect of motivation individuation on the CRE and not to closely replicate Hugenberg et al.'s (2007) study. As such, some of these studies manipulated additional independent variables other than instructions and face race, and some measured more outcomes than just recognition performance.

Specifically, Rhodes et al. (2009, Experiment 2) contrasted the original control and individuation instructions with a third type of instruction - race categorization. They found no interaction between instructions and face race, suggesting that none of the instruction conditions significantly decreased the CRE.[3] Both Young et al. (2010, Experiment 1) and Bornstein et al. (2013, Experiment 2) examined the effectiveness of the individuation instructions when given pre- versus post-encoding of faces. While the former study found that only the pre-encoding instructions eliminated the CRE, in the latter study, the CRE was unaffected by pre- or post-encoding instructions. Young and Hugenberg (2012, Experiment 1) tested whether the effect of the individuation instructions on the CRE is moderated by participants' levels of self-reported CR contact (as assessed through Hancock & Rhodes', 2008, interracial contact questionnaire). Consistent with the CIM (Hugenberg et al., 2010), they obtained a significant interaction between instructions and self-reported CR contact on the CRE, such that the greater the CR contact, the smaller the CRE, but only in individuation instruction condition.[4] Tullis et al. (2014, Experiment 3) investigated the influence of the individuation versus control instructions on the CRE as a function of whether participants can control the amount of time they allocate to each face (self-paced learning) or not (fixed-paced learning). The authors did not obtain any effect of the individuation instructions on the CRE in the self-paced learning condition nor in the fixed-paced learning condition. Tüttenberg and Wiese (2021, Experiment 1) used a two-by-two design similar to that of Hugenberg et al. (2007, Experiments 1A and 1B) but also measured participants' event-related brain potentials (ERPs) while learning and recognizing the faces. Regarding memory sensitivity, they did not obtain a statistically significant interaction between instructions and face race. Finally, Pica et al. (2015) and Wan et al. (2015) examined the effect of the individuation instructions on White versus Black participants and on White versus Asian participants, respectively. None of the two studies found any impact of individuation instructions on the CRE, both for White and non-White participants. Thus, only Young et al. (2010, Experiment 1) and Young and Hugenberg (2012, Experiment 1) found an effect of the individuation instruction on the CRE.

Another factor that may help explain why some replication studies succeeded and some failed to find an effect of individuation motivation on the CRE is the cultural context in which these studies took place. Namely, two (i.e., Young et al., 2010; Young & Hugenberg, 2012) of the six studies that were conducted in the United States successfully replicated the findings of Hugenberg et al.'s (2007) study, which also took place in the United States, while all four studies performed outside of the United States (i.e., in Australia and the United Kingdom) failed to do so. Thus, there appears to be a correlation between replication success and cultural setting, suggesting that the effectiveness of individuation instructions may be culture-dependent.

The possible effects of cultural setting may be better understood through the lenses of the CIM (Hugenberg et al., 2010) since it is the primary explanatory model for the role of individuation motivation in the CRE. The association between the effectiveness of individuation instructions and cultural background can be thought of as reflecting the influence of race as a social category. It is worth noting that the countries outside of the United States where replications were conducted score lower on racial diversity, as defined through ethnic fractionalization, or the probability that two randomly sampled individuals belong to different ethnic backgrounds (see Fearon, 2003).

The lower racial diversity of said countries may translate into differences at the level of both the motivation to individuate and expertise with CR individuals, due to reduced contact with CR individuals. On the one hand, low contact with CR individuals may reduce the relevance of the racial category, as interactions with said social category are not expected to emerge often, thus lowering the motivation to individuate

[2] The search strategy started with an analysis of the papers citing Hugenberg et al. (2007), Hugenberg et al. (2010), or Young et al. (2012). Papers identified as replications through this method had their references screened similarly. Additional keyword searches performed on Google Scholar revealed the remainder of the replications. Searches on Connected Papers (connectedpapers.com) and Scite (scite.ai) yielded no further data.

[3] The authors claim in the abstract of the paper that "…the normally robust other-race effect was absent when participants were instructed to individuate other-race faces…". However, such a claim is based on one-sample *t*-tests conducted within each condition (for details, see p. 239 of Rhodes et al., 2009).

[4] Surprisingly, the authors did not report whether they obtained a two-way interaction between instructions and face race. Instead, they report the results of one-sample *t*-tests comparing CRE difference scores against zero within each instruction condition, and the results of a t-test comparing the CRE difference scores between conditions (for details, see p. 3 of Young and Hugenberg (2012).

Table 1
List of studies that have examined the role of individuation instructions on the CRE (Hugenberg et al., 2007).

Study	Sample	Design	Results (*d'*)	Differences to the original article	Cultural setting
Rhodes et al. (2009, Experiment 2)	128 White participants	2 (face race: SR, CR) x 3 (instruction condition: control, race categorization, individuation), with instructions as a between-subjects variable.	No significant interaction between face race and instruction condition.	Different face stimuli; different number of stimuli; different duration of stimuli.	Australia
Young et al. (2010, Experiment 1)	44 White participants	2 (face race: SR, CR) x 3 (instruction condition: no instruction, pre-encoding, post-encoding), with instructions as a between-subjects variable.	Significant interaction between face race and instruction condition.	Different duration of stimuli.	USA
Young & Hugenberg (2012, Experiment 1)	60 White participants	2 (face race: SR, CR) x 2 (instruction condition: control, individuation), with instructions as a between-subjects variable.	Significant 3-way interaction between face race, instruction condition, and participant's self-reported CR contact (measured continuously).	Different duration of stimuli.	USA
Bornstein et al. (2013, Experiment 2)	83 White, 99 Hispanic, 182 participants	2 (face race: SR [White and Hispanics], CR [Black]) x 3 (instruction condition: control, CRE-encoding, CRE-retrieval), with instructions as a between-subjects variable.	No significant interaction between face race and instruction condition.	Different face stimuli.	USA
Tullis et al. (2014, Experiment 3)	160 White participants	2 (face race: SR, CR) x 2 (instruction condition: control, individuation) x 2 (pacing: self, fixed), with instructions and pacing as between-subjects variables.	No significant interaction between face race and instruction condition in any of the pacing conditions.	Different face stimuli; different; number of stimuli; different duration of stimuli; use of a recognition confidence scale.	USA
Pica et al. (2015, Experiment 1)	55 White, 23 Black, 78 participants	2 (face race: SR, CR) x 2 (instruction condition: control, individuation), with instructions as a between-subjects variable.	No significant interaction between face race and instruction condition for any of the participant races.	Different duration of stimuli.	USA
Pica et al. (2015, Experiment 2)	50 White, 28 Black, 78 participants	2 (face race: SR, CR) x 2 (instruction condition: control, individuation), with instructions as a between-subjects variable.	No significant interaction between face race and instruction condition for White participants.	Different face stimuli; different duration of stimuli.	USA
Wan et al. (2015, Experiment 4)	198 White, 181 Asian, 379 participants	2 (face race: SR, CR) x 2 (instruction condition: control, motivation instruction) x 2 (participant race: Western-raised White group, Eastern-raised Asian group), with instructions and participant race as between-subjects variables.	No significant interaction between face race and instruction condition for any of the participant races.	Different duration of stimuli; no distractor task.	Australia
Jerovich (2017, Experiment 1)	45 White, 25 Asian, 70 participants	3 (face race: Asian, White, Black) x 2 (instruction condition: control, individuation) x 2 (participant race: White, Asian), with instructions and participant race as between-subjects variables.	No significant interaction between face race and instruction condition for any of the participant races.	Different stimuli number and blocked design	Australia
Tüttenberg & Wiese (2021, Experiment 1)	36 White participants	2 (face race: SR, CR) x 2 (instruction condition: control, individuation), with instructions as a between-subjects variable.	No significant interaction between face race and instruction condition.	Different face stimuli; different number of stimuli; non-self-paced recognition; measurement of ERPs during learning and recognition	UK

CR individuals. On the other hand, an absence of opportunities to interact with CR individuals will translate into a lack of expertise to individuate CR faces, hindering potential individuation attempts.

Yet another aspect that may be worth reflecting on is the statistical power of some of the replication studies, which is hard to evaluate given the great variability in sample sizes between studies (ranging from 36 to 379 participants) and, more importantly, given that most of these studies do not provide a rationale for their sample sizes. Rhodes et al. (2009) mention that their sample size is large, though such comparison is drawn with their previous study; Pica et al. (2015) suggest that the low statistical power of their two experiments may explain the failure to detect a significant interaction between instructions and face race, but do not provide any support for this statement; Wan et al. (2015) advocate that their results cannot be attributed to lack of power due to their large sample size, but do not state the grounds in which they base such claim. Only Bornstein et al. (2013) refer to preliminary power analysis for their sample size, based on the effect sizes reported by Hugenberg et al.'s (2007). Although it is a common practice to base sample size estimations on published effect sizes, it is now widely accepted that such practice is problematic and often leads to underpowered studies as published effect size estimates are typically inflated due to publication bias (see da Silva Frost & Ledgerwood, 2020).

Besides the aspects mentioned before, there are a number of differences in the face stimuli and procedures used in each study that constitute another source of variability, though these are arguably a minor factor in explaining the disparity of results (Table 1 summarizes differences from the original article).

The present section highlighted several dimensions that vary across replications, which may help account for the reported conflicting findings. In light of these conflicting results, and given the high impact and theoretical relevance of Hugenberg et al. 's (2007) study, we believe that a pre-registered replication is an important contribution for a better understanding of the obtained results. Replications are especially valuable in the context of conflicting findings since their goal is to polish theories, providing new evidence for previous understanding (Nosek & Errington, 2020).

4. Overview of the current research

In light of the discussed replication studies, the current research aimed to address some of the highlighted issues in two experiments. Experiment 1A in the current study was a close replication of

Experiments 1A and 1B (which have the same experimental design) from Hugenberg et al. (2007). Namely, it was conducted in the United States with White American participants in a laboratory setting. To test for possible effects of cultural setting on the effectiveness of the individuation instructions, we also conducted a second laboratory experiment - Experiment 1B - in Portugal. Portugal is rated as a country with relatively low ethnic diversity (Fearon, 2003), especially compared with the United States. If indeed the individuation instructions have a greater likelihood of reducing the CRE in a cultural setting with greater racial diversity, as explained before, then we should obtain a significant interaction between instructions and face race only in Experiment 1A. At the end of both experiments, we measured participants' self-reported contact with White and Black individuals (Hancock & Rhodes, 2008). By including this measure, we can examine whether the two samples differ in their self-reported CR contact and whether CR contact interacts with instructions and face race, as predicted by the CIM (Hugenberg et al., 2010) and demonstrated by Young and Hugenberg (2012).

Besides the cultural setting in Experiment 1B, there are two other differences between the present and the original experiments that should be noted. The first pertains to the inclusion of attention checks and a self-report measure of effort to evaluate participants' investment in the task in both experiments. Because these measures occur after recognition performance is measured, they should not affect the results. Another dimension that differentiates both of our experiments from the original study is the use of face stimuli from a high-resolution face database (i.e., Chicago Face Database version 2.0.3; Ma, Correll, & Wittenbrink, 2015). Again, it is not clear to us that this alteration should have any effect on the results.

To guarantee that we have sufficient statistical power to detect an interaction between instructions and face race, we calculated our sample size based on a method that is not dependent on published effect sizes and therefore it is immune to publication bias. Namely, we used a simulation-based approach to perform power analyses for different interaction patterns between our variables (see Lakens & Caldwell, 2021), as explained in the Method section.

The present replication also stands out from previous ones because it is a pre-registered replication. As mentioned by Nosek et al. (2019), pre-registration strengthens whatever conclusions follow from data collection by reducing self-serving flexibility in data analysis and publication bias, and therefore is a powerful tool to improve research credibility.

The pre-registration protocol, materials, data, and analysis script are publicly available in the Open Science Framework (OSF) platform (https://osf.io/f3vcm/?view_only=5af44520f7b14debbccf1c8bb03c0945). The Ethics Committee of the authors' institutions approved this research. The two experiments reported bellow were conducted simultaneously from October 2021 to May 2022.

5. Experiment 1A: United States sample

Experiment 1A attempted to closely replicate Experiments 1A and 1B of Hugenberg et al. (2007), which showed that instructing White American participants to individuate CR faces (i.e., Black) eliminates the CRE.

5.1. Method

5.1.1. Participants

To adequately power our experiments to detect a significant ($p < .05$) ordinal interaction, we used a simulation-based tool that allows the realization of power analysis for factorial ANOVA designs based on predicted patterns of means - the *superpower* R package (Lakens & Caldwell, 2021). Because the shape of interaction effects influences power calculations (e.g., Giner-Sorolla, 2018), we simulated power for three different interaction patterns. These simulations revealed that a sample size of 300 participants (150 per condition) gives us (a) > 99% power to detect an interaction effect ($\eta_p^2 = 0.12$) in which the CRE is completely suppressed in the individuation condition; (b) 99% power to detect an interaction effect ($\eta_p^2 = 0.06$) in which the CRE is reduced by 75% in the individuation condition; and (c) 89% power to detect an interaction effect ($\eta_p^2 = 0.04$) in which the CRE is reduced by 50% in the individuation condition (see Appendix A for a detailed explanation of these simulations). This method was used to determine the sample of both experiments.

We planned to oversample, with an *N* of 400, because we anticipated data loss due to the exclusion of non-White participants and the application of the remaining exclusion criteria. We ultimately collected data from 520 undergraduate students who participated in return for course credits. Following our pre-registered exclusion criteria, we excluded (a) 6 participants who did not complete the entire experiment, (b) 131 who reported belonging to a racial group other than White,[5] (c) 12 that said they did not adhere to the instructions during the completion of the tasks, and (d) 8 participants who failed all three attentional checks trials described in the Procedure (i.e., do not press the "Enter" key within the 10 s). The Experiment 1A pre-registration is available here: https://osf.io/nmy79. Additionally, we also excluded 53 participants whose average sensitivity scores (*d'*) were minor or equal to zero, indicating that these participants were unable to discriminate the signal (seen faces) from the noise (unseen faces). Because we only thought about this criterion after the preregistering the experiments, they were not added to the pre-registration protocols. However, we do believe this is an essential exclusion criterion as it allows for the removal of those participants who responded at chance level ($d' = 0$) and those who seem to have mistaken the new faces by the old faces ($d' < 0$).[6] This left us with a final sample of 310 White participants (204 females, 97 males, 4 transgender, 5 belonging to other genders; $M_{age} = 18.90$, $SD_{age} = 1.28$). A sensitivity power analysis using the R package *pwr* (Champely, 2018) revealed that this sample size gave us 94% statistical power (alpha = 0.05) to detect an effect size of $\eta_p^2 = 0.04$.

5.1.2. Materials

The materials consisted of face photographs of White and Black faces selected from the Chicago Face Database (CFD; version 2.0.3; Ma et al., 2015). This database includes high-quality, standardized photographs of male and female faces from different racial groups, pre-tested on several subjective (e.g., trustworthiness, attractiveness) and objective (e.g., nose width, lip thickness) dimensions.

From the full set of photographs available (see http://www.chicagofaces.org/), all male faces belonging to both races used in Hugenberg et al. (2007) – White and Black – were considered. From this initial group of faces, those that were not easily recognizable as male and White/Black – defined as male and target-race identification rates lower than 0.9 on the CFD Norming Data v. 2.0.3 document – were discarded. 70 White faces and 61 Black faces remained. To round the number of stimuli and make them even, 1 Black and 10 White photographs were randomly excluded. The final set of photographs was composed of 60 White and 60 Black faces with a neutral expression and direct gaze. These photographs were then cropped to display only the targets' face and hair, gray-scaled (as in Hugenberg et al., 2007), and resized to 382 × 330 pixels. Each participant was presented with a set of 80 randomly selected photographs from our list (i.e., 20 White and 20 Black at encoding and an additional 20 White and 20 Black in the recognition phase). These numbers (i.e., 80 faces total) were taken from the methods of the original study. Sampling stimulus faces from a larger pool increases the chances that any results obtained in the present experiments

[5] 7 Black, 40 Asian, 53 mixed race, and 31 other races.

[6] Inclusion of these participants does not meaningfully change the results.

generalize across a broader range of faces.

5.1.3. Procedure

Prior to initiating the experiment, participants were presented with an informed consent screen explaining that they could quit the experiment at any time and that their data was entirely anonymous. Additionally, participants were informed that they should read the instructions carefully and adhere to them during the completion of the tasks. Upon acceptance of these requirements, participants were told that they would be taking part in a face recognition experiment. Though we did not control for visual angle (Hugenberg et al. do not report any information about the visual angle themselves), participants' expected viewing distance (i.e., 80 cm from the screen) and corresponding stimuli dimensions would approximately equate to a visual angle of 6.2° during participation (https://elvers.us/perception/visualAngle/).

Following Hugenberg et al. (2007), participants were randomly allocated to one of two conditions — control and individuation. Participants in the control condition were simply told to learn each face for a later recognition test, while participants in the individuation condition were warned of the CRE and instructed to pay close attention to what differentiates individual CR faces (e.g., Black faces) from one another in order to avoid this effect. The individuation instructions used in the present experiments were identical to those used by Hugenberg et al. (2007). See page 7 of this paper (or the project's page at OSF) for the verbatim instructions used.

In the experiment's learning phase, participants saw 40 Gy-scale target faces (20 Black and 20 White) presented in a random order against a white background for 3 s each. As explained before, these faces were randomly selected from our list of 120 faces. The faces were presented sequentially in the middle of the screen, with an inter-trial interval of 1 s. After the 10th, 20th, and 30th learning trials, participants were presented with an attentional check trial for 10 s. These trials consisted of a face (not presented in the recognition test) shown in the middle of the screen with a sentence underneath. This sentence instructed participants to press the "Enter" key before the face disappears from the screen (i.e., 10 s).

After the learning phase, participants completed a 5-min distractor task (i.e., recalling as many capital cities as they could) meant to clear their working memory. In the subsequent recognition phase, they saw 40 target faces (i.e., faces shown during the learning phase) randomly intermixed with 40 (20 Black and 20 White) lure faces (i.e., faces not shown in the learning phase). Participants' task was to decide whether each face was presented in the learning phase or not by selecting either a "Yes" (key 1) or "No" (key 2) response option, respectively. Each face was displayed on the screen until participants responded, after which the next face was immediately presented until all 80 faces were judged.

Following the recognition phase, participants completed an adapted version of a CR contact scale developed by Hancock and Rhodes (2008). As the original version measures self-reported interaction with White and Chinese race groups, the scale was adapted to fit our CR target – Black individuals - instead (see Appendix B). For each statement (e.g., "I know lots of Black people"), participants indicated their agreement on a scale ranging from 1 (very strongly disagree) to 6 (very strongly agree).

Finally, participants answered three socio-demographic questions (age, gender, and racial group) and the following control question: "Did you read the instructions carefully and adhere to them during the completion of the tasks? (Yes/No)". For the participants who responded "Yes" to the previous question, we assessed the degree of effort they put into the task using two items taken from Wan et al. (2015). Specifically, they were asked to indicate on a 7-point scale, "How much special effort did you put into telling apart the faces of the White people you saw?" followed by "How much special effort did you put into telling apart the faces of the Black people you saw?"

5.2. Results and discussion

5.2.1. Analytic strategy

Following our pre-registration protocol, below we present the same statistical analyses reported by Hugenberg et al. (2007; Experiments 1A and 1b) and Young and Hugenberg (2012; Experiment 1)[7] to facilitate the comparison between the results obtained by these authors and ours. In addition, we complemented these analyses with Bayes factors as these allow for quantifying the amount of evidence for the presence or absence of a given effect, thus providing a more nuanced interpretation of the results (Jeffreys, 1961; Wagenmakers, 2007).[8] Specifically, we report the inclusion Bayes factor (B_{incl}) which quantifies the change from prior inclusion odds (i.e., the probability that a predictor is included in the model before seeing the data) to posterior inclusion odds (i.e., the probability that a predictor is included in the model after seeing the data) and can be interpreted as the evidence for or against including a predictor (van den Bergh et al., 2020). Following Lee and Wagenmakers's (2014) classification scheme for facilitating the interpretation of Bayes factors, a BF_{incl} between 1 and 3 is conventionally interpreted as anecdotal evidence *in favor* of an effect's inclusion, a BF_{incl} between 3 and 10 as moderate evidence, and a $BF_{incl} > 10$ as strong evidence. Conversely, a BF_{incl} between 1 and 0.33 is interpreted as anecdotal evidence *against* an effect's inclusion, a BF_{incl} between 0.33 and 0.1 as moderate evidence, and $BF < 0.1$ as strong evidence. For conciseness, we only report inclusion Bayes factors for the analyses that are directly relevant for assessing the success or failure of the current replication attempt. That is, for the analyses examining the effect of individuation instructions on the CRE (*d'*), and the analyses examining the interactive effect of individuation instructions and CR contact on the CRE. Inclusion Bayes factors for the remaining analyses reported here can be found in the supplementary material accompanying this paper and here: https://osf.io/f3vcm/?view_only= 5af44520f7b14debbccf1c8bb03c0945.

The data obtained in the two experiments were also submitted to a series of generalized mixed-model analyses treating both participants and face stimuli as random factors, which yielded results consistent with the ones described below and thus are only reported in the supplementary material.

5.2.2. Effects of individuation instructions on the CRE

Following Hugenberg et al. (2007), the main analysis was conducted on participants' *d'* scores, a signal detection measure of sensitivity (see Green & Swets, 1966; Macmillan & Creelman, 1991). Sensitivity (*d'*) refers to one's ability to discriminate between target faces (old) and lure faces (new), with higher *d'* scores representing higher sensitivity. For each participant, we first calculated hits (H) and false alarms (FA) for SR and CR faces and then used these to calculate their respective *d'* scores ($d' = [z(H) - z(FA)]$).[9] Table 2 displays the means and standard deviations for all these measures as a function of face race and instruction condition.

Participants' *d'* scores for SR and CR faces were then submitted to a

[7] Due to a misinterpretation of the analytic strategy followed by Young and Hugenberg (2012), we have pre-registered that we would calculate a contact difference score by subtracting participants' mean contact with Black people from their mean contact with White people. However, the authors did not compute such difference score, using only participants' mean contact with Black people in their regression analysis. Thus, we decided to follow Young and Hugenberg's (2012) strategy and use only participants' mean contact with Black people in all analyses involving contact. We want to note, though, that repeating these analyses with the pre-registered contact difference score yields almost identical results.

[8] All computations used the standard priors considered by the *BayesFactor* R package.

[9] To avoid infinite values of *d'*, hit and false-alarm proportions of 0 and 1 were converted to $1/(2N)$ and $1-1/(2N)$, respectively, where N is the number of signal or noise trials (Macmillan & Creelman, 1991).

Table 2
Hits, False Alarms, and Response Criterion scores for SR and CR faces as a function of Instruction condition (Experiment 1A).

	Control Condition		Individuation Condition	
	SR faces	CR faces	SR faces	CR faces
Hits	0.52 (0.16)	0.56 (0.16)	0.50 (0.16)	0.56 (0.16)
False Alarms	0.23 (0.14)	0.30 (0.15)	0.20 (0.13)	0.29 (0.15)
Sensitivity (*d'*)	0.92 (0.54)	0.74 (0.45)	0.95 (0.57)	0.76 (0.49)
Criterion (*c*)	0.40 (0.42)	0.19 (0.38)	0.48 (0.38)	0.23 (0.39)

Note. Standard deviations are presented within brackets.

mixed ANOVA with face race (SR vs. CR) as a within-subjects factor and instruction condition (control vs. individuation) as a between-subjects factor. Following Steiger's (2004) suggestion, we report 90% confidence intervals for partial eta-squared values (see also Lakens, 2013). If instructing participants to individuate CR faces indeed eliminates or reduces the CRE, then we should obtain an ordinal interaction between instruction condition and face race, such that participants in the control condition would display the standard CRE while participants in the individuation instruction condition would either not show the CRE or an attenuated CRE.

The results of the ANOVA revealed the expected main effect of face race, $F(1, 308) = 30.80$, $p < .001$, $\eta_p^2 = 0.09$, 90% CI [0.05, 0.14], showing that participants recognized SR faces more accurately than CR faces, and no main effect of instruction, $F(1, 308) = 0.26$, $p = .613$, $\eta_p^2 < 0.001$, 90% CI [0.00, 0.01], suggesting that overall recognition performance was similar in the two instruction conditions. More importantly, there was no significant interaction between instructions and face race, $F(1, 308) = 0.02$, $p = .878$, $\eta_p^2 < 0.001$, 90% CI [0.00, 0.01]. A complementary Bayesian ANOVA showed a BF_{incl} of 0.07 for the interaction, indicating strong evidence *against* the inclusion of the interaction effect. This analysis also yielded a BF_{incl} of 0.11 for the main effect of instructions, indicating moderate evidence against the effect of instructions, and a BF_{incl} of 1.17e+5 for the main effect of race, denoting strong evidence *in favor* of a race effect. These results demonstrate that instructing participants to attend to individuated features of CR faces did not lead to a recognition improvement of these faces, thus failing to replicate the findings of Hugenberg et al. (2007). Fig. 1 presents the mean sensitivity scores for SR and CR faces as a function of instructions for the two experiments.

To better understand the sensitivity (*d'*) results described above, we conducted separate ANOVAs on the proportion of hits (H) and false alarms (FA), and on participants' response criterion ($c = -0.5[z(H) + z(FA)]$; Macmillan & Creelman, 1991). The response criterion reflects participants' threshold for selecting one type of response over the other, irrespective of the actual status of an item ("old" or "new"). Negative values of *c* indicate a more liberal response criterion (i.e., a bias towards characterizing faces as "old"), while positive values of *c* reflect a more conservative criterion (i.e., a bias towards characterizing faces as "new"). To keep this section concise, below we only report the significant effects obtained (see the supplementary material for the complete effects tables).

We obtained a higher proportion of hits for CR than for SR faces, $F(1, 308) = 27.52$, $p < .001$, $\eta_p^2 = 0.08$, 90% CI [0.04, 0.13]. There was also a higher proportion of false alarms for CR than SR faces, $F(1, 308) = 95.91$, $p < .001$, $\eta_p^2 = 0.24$, 90% CI [0.17, 0.30], as expected. This pattern of results indicates that the obtained CRE in recognition performance (*d'*) was entirely driven by the greater proportion of false alarms for CR than SR faces, which is inconsistent with the "mirror effect" pattern (i.e., lower proportion of hits and higher proportion of false alarms for CR faces when compared with SR faces) reported in Meissner and Brigham's (2001) meta-analysis. Finally, for the response criterion (*c*), and as expected (see Meissner & Brigham, 2001), we found that participants exhibited a more liberal response criterion when responding to CR faces than to SR faces, $F(1, 308) = 91.38$, $p < .001$, $\eta_p^2 = 0.23$, 90% CI [0.16, 0.29].

5.2.3. Interactive effects of individuation instructions and CR contact on the CRE

Following Young & Hugenberg (2012; Experiment 1), we examined whether individuation instructions and CR contact are simultaneously implicated in eliminating the CRE. First, we averaged the eight items referring to contact with Black individuals (Chronbach's alpha = 0.86) and the seven items referring to contact with White individuals (Chronbach's alpha = 0.83). Not surprisingly, our final sample of only White participants reported significantly greater contact with White people ($M = 5.10$, $SD = 0.62$, minimum = 3.14, maximum = 6.00) than Black people ($M = 3.20$, $SD = 0.77$, minimum = 1.00, maximum = 5.38; $t(309) = -30.25$, $p < .001$). Next, we calculated a CRE index by subtracting participants' CR *d'* scores from their SR *d'* scores, with higher scores indicating a stronger CRE. Finally, we conducted a linear regression analysis with the instruction condition (individuation condition = − 1, control condition = 1) as a between-subjects factor and self-reported CR (cross-race) contact (mean-centered) as a continuous predictor of the newly created CRE index along with their interaction. Unlike Young & Hugenberg (2012; Experiment 1), we did not obtain a significant interaction between instructions and contact, $b = -0.017$, $SE = 0.04$, $t(306) = -0.39$, $p = .700$, indicating that greater levels of CR contact did not facilitate the effect of individuation instructions on the CRE (see Fig. 2). Not surprisingly given the small relationship between contact and the CRE (see Singh et al., 2021), we did not find a significant effect for contact, $b = -0.052$, $SE = 0.04$, $t(306) = -1.18$, $p = .237$. Lastly, the effect of instructions was also not significant, $b = -0.007$, $SE = 0.03$, $t(306) = -0.19$, $p = .847$. A Bayesian regression revealed strong evidence against the inclusion of both the main effect of instructions ($BF_{incl} = 0.09$) and the face race x instruction condition interaction ($BF_{incl} = 0.02$), as well as moderate evidence against the inclusion of the main effects of contact ($BF_{incl} = 0.18$).

5.2.4. Interactive effects of individuation instructions, contact, and self-reported effort on the CRE

In this last section, we performed a number of *non-pre-registered exploratory analyses* involving participants' ratings of how much effort they put into telling apart SR faces and CR faces during the learning phase (see the experiment's Procedure section) as we believe these may be informative about the key idea underlying the CIM (Hugenberg et al., 2010) — that perceivers have equal ability to process CR and SR faces but do not invest effort when processing CR faces, thus giving rise to the CRE. For the analyses described next, we report all significant and non-significant effects directly relevant to the questions addressed. Statistical information about the remaining effects and Bayes Factors values for all effects are available in the supplementary material.

In the first analysis, we examined whether instructing participants to individuate CR faces increased the amount of effort applied to learning CR faces relative to the amount of effort applied to these faces in the absence of such instructions. The results of a mixed ANOVA, with face race as a within-subjects factor and instruction condition as a between-subjects factor, showed a main effect of face race indicating that participants applied more effort to CR faces than to SR faces, $F(1, 308) = 90.58$, $p < .001$, $\eta_p^2 = 0.23$, 90% CI [0.16, 0.29]. This main effect was however qualified by an interaction between effort and instruction condition, $F(1, 308) = 18.59$, $p < .001$, $\eta_p^2 = 0.06$, 90% CI [0.02, 0.10]. A contrast analysis revealed that while the effort applied to SR faces was roughly the same in the two instruction conditions ($M_{individuation} = 3.81$, $SD = 1.59$ versus $M_{control} = 3.90$, $SD = 1.76$; $t(355) = 0.46$, $p = .649$), there was indeed more effort put into the processing of CR faces in the individuation ($M_{individuation} = 4.51$. $SD = 1.62$) relative to the control condition ($M_{control} = 4.16$. $SD = 1.67$), although this difference did not reach the conventional level of significance; $t(355) = -1.85$, $p = .064$. Importantly, participants reported applying more effort to CR than SR faces in both the individuation, $t(308) = 9.53$, $p < .001$, and control

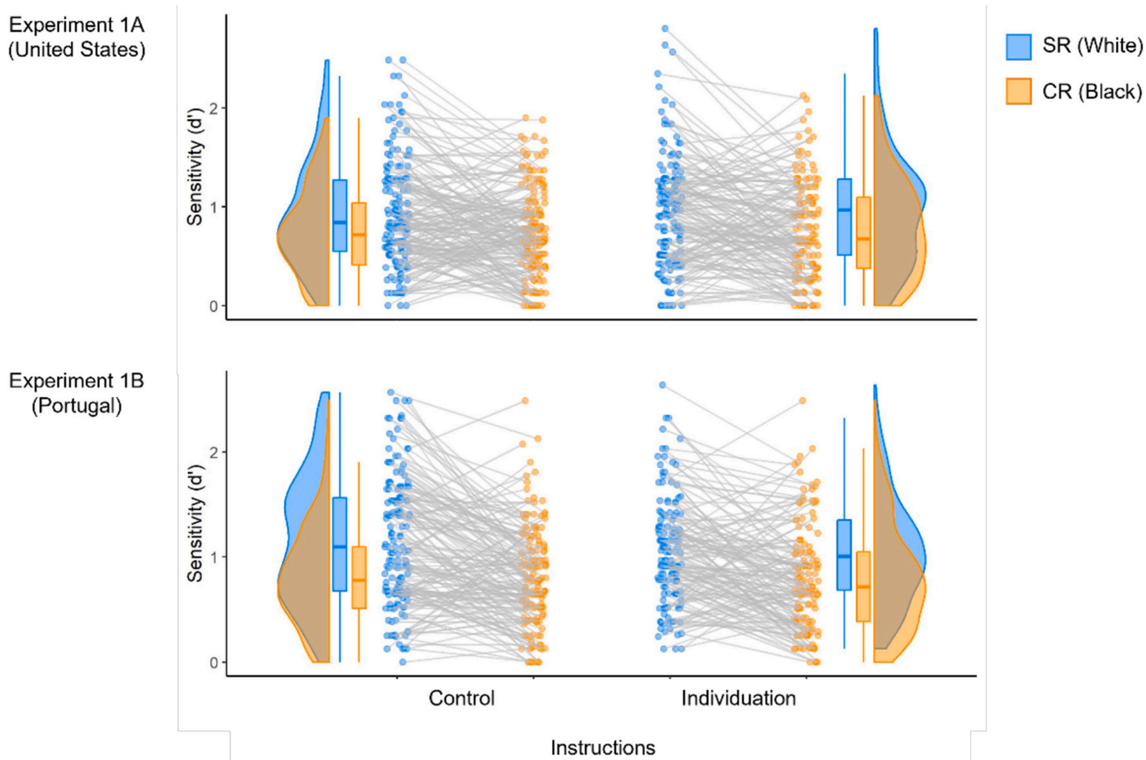

Fig. 1. Raincloud plots representing mean sensitivity (d') as a function of face race and instruction condition for both Experiment 1A (United States) and Experiment 1B (Portugal). Error bars represent 95% confidence intervals. Dots represent participants' mean sensitivity scores.

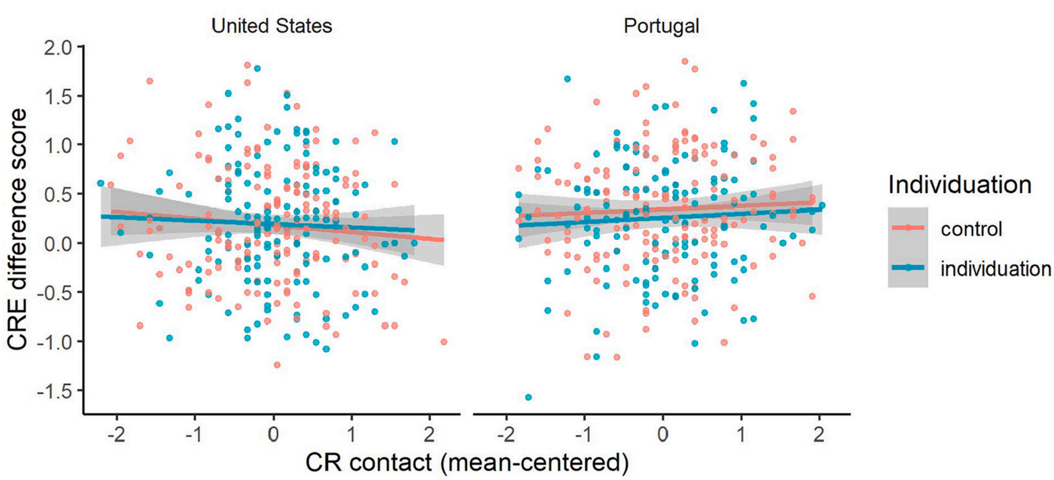

Fig. 2. Effect of CR contact and instruction condition on CRE mean difference score for both Experiment 1A (United States) and Experiment 1B (Portugal). The shaded regions show 95% confidence intervals. Dots represent participants' mean difference scores.

conditions, $t(308) = 3.78$, $p < .001$. This difference in self-reported effort between CR and SR in the control condition seems inconsistent with the categorization-individuation model (Hugenberg et al., 2010), as the model proposes that the CRE is in part driven by the lack of effort typically applied to CR faces (for similar findings, see Wan et al., 2015).

The second analysis tested whether the amount of self-reported effort applied to CR faces during learning is positively associated with recognition performance for these faces and whether this association is stronger for participants in the individuation condition than for participants in the control condition. To examine these questions, we ran a linear regression analysis with the *d'* scores for CR faces as the dependent variable, with instruction condition (individuation condition $= -1$, control condition $= 1$) as a between-subjects factor, effort to CR (Black) faces as a continuous predictor (mean-centered), and their interaction. This analysis revealed a positive effect of effort on sensitivity trending towards significance, $b = 0.028$, $SE = 0.02$, $t(306) = 1.69$, $p = .093$, but no significant interaction between instruction condition and effort applied to CR faces, $b = -0.001$, $SE = 0.02$, $t(306) = -0.02$,

$p = .985$.[10] There was also no effect of instruction condition, $b = -0.0003$, $SE = 0.03$, $t(306) = 0.17$, $p = .862$.

In the third and final analysis, we examined whether individuation instructions, contact with CR individuals, and self-reported effort applied to CR faces interact to predict the CRE.[11] A linear regression analysis with CR contact and effort applied to CR faces (both mean-centered) as continuous predictors and instruction condition as a between-subjects factor, and all possible interactions, revealed a non-significant interaction between these three variables on the CRE ($b = 0.012$, $SE = 0.03$, $t(302) = 0.46$, $p = .645$). We did obtain a significant interaction between contact and effort ($b = 0.066$, $SE = 0.03$, $t(302) = 2.54$, $p = .012$), which we only describe in the supplementary material.

6. Experiment 1B: Portugal sample

Experiment 1B aimed at replicating the effect of individuation instructions on the CRE (Hugenberg et al., 2007) using a sample of Portuguese participants.

6.1. Method

6.1.1. Participants

The sample size for this experiment was determined following the rationale explained before. As can be seen in our pre-registration protocol for Experiment 1B (https://osf.io/sydhf), we set our target sample to 330 participants. However, because of the high number of participants' exclusions observed in Experiment 1A,[12] we decided to continue recruiting participants until no more participants were available. We ended up with a sample of 361 participants who participants in return of a 10€ gift voucher ($n = 252$) or partial course credit ($n = 109$).[13] Per our pre-registered exclusion criteria, we excluded (a) 6 participants who did not complete the entire experiment, (b) 29 who reported belonging to a racial group other than White,[14] (c) 2 that said they did not adhere to the instructions during the completion of the tasks, and (d) 7 participants who failed all three attentional checks trials (i.e., did not press the "Enter" key within 10 s). As in Experiment 1A, we excluded an additional 24 participants whose average *d'* scores were minor or equal to zero, indicating a complete inability to discriminate seen from unseen faces.[15] Our final sample included 293 White participants (223 females, 68 males, 1 transgender, 1 belonging to another gender; $M_{age} = 24.06$, $SD_{age} = 6.59$). Although the final sample size did not reach the desired number, a sensitivity power analysis using the *pwr* R package (Champely, 2018) showed that we had 93% statistical power to detect an effect size of $\eta_p^2 = 0.04$.

6.1.2. Materials and procedure

Experiment 1B was conducted in Portuguese. The face stimuli and procedures were identical to Experiment 1A.

6.2. Results and discussion

6.2.1. Analytic strategy

In line with pre-registration protocol, and the previous experiment, we present the same statistical analyses reported by Hugenberg and colleagues (Hugenberg et al., 2007; Young & Hugenberg, 2012) and complement these analyses with inclusion Bayes factors (BF_{incl}).

6.2.2. Effects of individuation instructions on the CRE

A 2 (face race: SR vs. CR) X 2 (instruction condition: control vs. individuation) on participants sensitivity scores (*d'*) yielded the expected main effect of face race, $F(1, 291) = 76.85$, $p < .001$, $\eta_p^2 = 0.21$, 90% CI [0.14, 0.28], with higher sensitivity for SR than CR faces. Once again, there was no significant interaction between face race and instruction condition, $F(1, 291) = 1.62$, $p = .203$, $\eta_p^2 = 0.006$, 90% CI [0.00, 0.03]. See Fig. 1. The main effect of instruction condition was also not significant, $F(1, 291) = 1.05$, $p = .307$, $\eta_p^2 = 0.004$, 90% CI [0.00, 0.02]. Consistent with these results, a Bayesian ANOVA revealed strong evidence *in favor* of the inclusion of the face race main effect ($BF_{incl} = 4.20 + 13$), and moderate evidence *against* the inclusion of either the interaction between race and instruction condition ($B_{incl} = 0.21$) or the main effect of instruction condition ($B_{incl} = 0.19$). In line with Experiment 1A, these results showed that instructing participants to individualize CR faces did not increase their sensitivity for these faces (i.e., better discrimination between target and lure faces).

Following Experiment 1A, we ran separate ANOVAs on the proportion of hits (H) and false alarms (FA), and on participants' response criterion (*c*). Table 3 presents the means and standard deviations for all measures as a function of face race and instruction condition. To keep this section as concise as possible, we only report the significant effects that emerged from the analyses (for the complete effects tables, see the supplementary materials).

Regarding the proportion of hits, there was a significant interaction between face race and instruction condition, $F(1, 291) = 7.15$, $p = .008$, $\eta_p^2 = 0.02$, 90% CI [0.00, 0.06], showing a similar proportion of hits for CR and SR faces in the control condition, $t(291) = 1.08$, $p = .283$, but a higher proportion of hits for CR than SR faces in the individuation instruction condition, $t(291) = -2.66$, $p = .008$. Planned contrasts also showed an increase in the proportion of hits for CR faces in the individuation condition compared to the control condition, $t(494) = -2.52$, $p = .012$, while for SR faces the proportion of hits was roughly the same in the two conditions, $t(494) = 0.35$, $p = .724$.

For the proportion of false alarms, we obtained a main effect of face race, $F(1, 291) = 100.44$, $p < .001$, $\eta_p^2 = 0.26$, 90% CI [0.19, 0.32], indicating more false alarms for CR than SR faces. The main effect of instruction condition trended towards significance, $F(1, 291) = 3.59$, $p = .059$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.04], suggesting a greater proportion of false alarms for both CR and SR faces in the individuation condition than in the control condition. In line with the previous experiment, these results demonstrate that the CRE obtained in the present experiment was only caused by the higher proportion of false alarms for CR than SR faces (see Meissner & Brigham, 2001).

Table 3
Hits, False Alarms, and Response Criterion scores for SR and CR faces as a function of Instruction condition (Experiment 1B).

	Control Condition		Individuation Condition	
	SR faces	CR faces	SR faces	CR faces
Hits	0.56 (0.16)	0.55 (0.17)	0.56 (0.15)	0.59 (0.15)
False Alarms	0.20 (0.14)	0.27 (0.15)	0.21 (0.13)	0.32 (0.15)
Sensitivity *(d')*	1.15 (0.61)	0.81 (0.48)	1.05 (0.51)	0.80 (0.50)
Criterion (*c*)	0.40 (0.40)	0.27 (0.43)	0.37 (0.37)	0.14 (0.38)

Note. Standard deviations are presented within brackets.

[10] The effect of effort on sensitivity becomes significant ($b = 0.078$, $SE = 0.03$, $t(306) = 2.43$, $p = .016$) when we use an effort difference score (CR effort – SR effort) as the continuous predictor, but not the interaction between effort and individuation instructions ($b = -0.023$, $SE = 0.03$, $t(306) = -0.73$, $p = .468$). A similar regression analysis with the CRE index as the dependent variable yielded no significant results.

[11] A similar analysis with sensitivity for CR faces as the dependent variable yields no significant results.

[12] Although the two experiments started running almost simultaneously, the data collection for Experiment 1A was completed first, which allowed us to verify how many participants we would need to exclude and use this knowledge to inform the data collection for Experiment 1B.

[13] Of the 293 participants that constitute our final sample, 205 received a 10 € gift card and 88 received course credit. Being compensated with money or course credit did not influence any of the results reported here.

[14] 7 Black, 16 mixed race, and 5 other races.

[15] Inclusion of these participants does not meaningfully change the results.

As expected, given the pattern of results obtained for hits and false alarms, the analysis of participants' *c* scores showed a more liberal response criterion for CR faces than to SR faces, $F(1, 291) = 55.94$, $p < .001$, $\eta_p^2 = 0.16$, 90% CI [0.10, 0.23]. There was also a significant main effect of instructions, $F(1, 291) = 4.30$, $p = .039$, $\eta_p^2 = 0.02$, 90% CI [0.00, 0.05], showing that participants in the individuation condition exhibited a more liberal criterion than those in the control condition, and an interaction between the two factors, $F(1, 291) = 4.22$, $p = .041$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.05], indicating that the difference in response criterion between CR and SR was more pronounced in the individuation instruction condition, $t(291) = 6.55$, $p < .001$, than in the control condition, $t(291) = 3.95$, $p < .001$. Moreover, relative to the participants in the control condition, participants in the individuation condition applied a more liberal response criterion to CR faces, $t(491) = 2.85$, $p = .005$, but not to SR faces, $t(491) = 0.66$, $p = .512$.

Taken together, the results of these auxiliary analyses indicate that the individuation instructions influenced the response criterion that participants used to respond to CR faces, making it more liberal (i.e., more biased towards answering "old"). Because Hugenberg et al. (2007) did not present these analyses, we do not know the relative contribution of hits and false alarms to the effect of individuation instructions on the CRE, nor whether a change in response criterion accompanied this effect.

6.2.3. Interactive effects of individuation instructions and CR contact on the CRE

To test whether individuation instructions and CR contact jointly influence the CRE, we followed the same procedures outlined for Experiment 1A. Again, we started by averaging the items referring to contact with Black people (Chronbach's alpha = 0.80) and the items referring to contact with White people (Chronbach's alpha = 0.83). As expected, our final sample of White Portuguese participants reported more contact with White people ($M = 5.33$, $SD = 0.68$, minimum = 1.00, maximum = 6.00) than with Black people ($M = 2.84$, $SD = 0.85$, minimum = 1.00, maximum = 4.88; $t(292) = -35.37$, $p < .001$). Next, we conducted a regression analysis with the instruction condition (individuation condition = − 1, control condition = 1) as a between-subjects factor, CR contact (mean-centered) as a continuous predictor, and their interaction. The CRE index was the dependent variable ($d'_{SR} - d'_{CR}$). Congruent with Experiment 1A, we did not obtain a significant interaction between instructions and contact, $b = -0.004$, $SE = 0.04$, $t(289) = -0.09$, $p = .926$. See Fig. 2. There was also no effect of contact, $b = 0.038$, $SE = 0.04$, $t(289) = 0.95$, $p = .343$, nor an effect of instruction condition, $b = 0.043$, $SE = 0.03$, $t(289) = 1.25$, $p = .212$. A Bayesian regression revealed strong evidence against the inclusion the interaction effect ($BF_{incl} = 0.02$), and moderate evidence against the inclusion of both main effects (contact: $BF_{incl} = 0.14$; individuation instructions: $BF_{incl} = 0.19$).

6.2.4. Interactive effects of individuation instructions, contact, and self-reported effort on the CRE

Following Experiment 1A, we conducted three *non-pre-registered exploratory analyses* involving participants self-reported effort ratings. To keep this section as brief as possible, we only report the effects (significant or non-significant) that are directly relevant to the examined questions. The complete effects tables for these analyses and complementary Bayesian analyses are presented in the supplementary material.

Firstly, we examined whether the amount of effort applied to the processing of CR and SR faces varied as a function of individuation instructions. The results of a mixed ANOVA yielded a main effect of face race, $F(1, 291) = 76.35$, $p < .001$, $\eta_p^2 = 0.21$, 90% CI [0.14, 0.27], showing that participants applied more effort to CR faces than to SR faces. In line with our previous results, there was an interaction between face race and instruction condition, $F(1, 291) = 10.22$, $p = .002$, $\eta_p^2 = 0.03$, 90% CI [0.01, 0.07]. Decomposing this interaction showed that the level of effort applied to the processing of CR faces was higher to in the individuation condition than in the control condition ($M_{individuation} = 5.38$. $SD = 1.22$ versus $M_{control} = 4.99$, $SD = 1.39$; $t(415) = -2.34$, $p = .020$), while the effort applied to the processing of SR was essentially the same in both conditions, ($M_{individuation} = 4.53$. $SD = 1.49$ versus $M_{control} = 4.59$, $SD = 1.55$; $t(415) = 0.39$, $p = .698$). In addition, participants reported applying more effort to CR than SR faces in the individuation condition, $t(291) = -8.21$, $p < .001$, but also in the control condition, $t(291) = -4.04$, $p < .001$.

Secondly, we examined whether self-reported effort applied to CR faces interacts with instruction condition to predict recognition sensitivity for CR faces. To do so, we ran a regression analysis with participants *d'* scores for CR faces as the dependent variable, with instruction condition (individuation condition = − 1, control condition = 1) as a between-subjects factor, effort applied to CR faces as a continuous predictor (mean-centered), and their interaction. This analysis revealed an interaction between instruction condition and effort trending towards significance, $b = -0.037$, $SE = 0.02$, $t(289) = -1.65$, $p = .101$, such that, in the individuation instruction condition, more self-reported effort applied to CR faces was associated to a reduction in the CRE. In contrast, in the control condition, more effort seems to be associated with an increased CRE.

Thirdly, we tested whether individuation instructions, contact with CR individuals, and effort applied to CR faces interact to predict the CRE.[16] A regression analysis with CR contact and effort applied to CR faces (both mean-centered) as continuous predictors, instruction condition as a between-subjects factor, and all possible interactions yields a significant interaction between instruction condition and effort, $b = 0.059$, $SE = 0.03$, $t(285) = 2.17$, $p = .031$, showing that, in the individuation instruction condition, greater levels of self-reported effort applied to CR faces were associated with a decrease in CRE, while no such association existed for participants in the control condition. Although this may seem a theoretically relevant interaction effect, it should be interpreted with caution, given that a complementary Bayesian regression analysis showed moderate evidence against this effect ($BF_{incl} = 0.15$) and the same did not emerge in Experiment 1A. Thus, we will not further elaborate on this finding.

6.2.5. Examining the moderating role cultural setting

As highlighted in the Introduction, Hugenberg and colleagues' research (Hugenberg et al., 2007; Young & Hugenberg, 2012) was conducted in the United States, a country with greater racial diversity than Portugal (Fearon, 2003). Thus, assuming that a cultural setting with more racial diversity affords more opportunities for CR contact and therefore tends to increase perceptual expertise for CR faces, it seems plausible that our United States participants would more easily overcome the CRE when instructed to do so compared to our Portuguese participants. Before testing this exploratory hypothesis, we first examined whether the two samples differ in their mean levels of self-reported CR contact. As expected, U.S. participants indeed reported greater contact with CR individuals ($M = 3.20$, $SD = 0.77$) than Portuguese participants ($M = 2.84$, $SD = 0.85$; $t(601) = 5.45$, $p < .001$). To examine the possible moderating role of the cultural setting on the effectiveness of the individuation instructions, we then submitted participants' ($N = 603$) *d'* scores to a mixed ANOVA with experiment (1A vs. 1B) and instruction condition as between factors and face race as a within-subjects factor. Such analysis yielded two significant effects, the predicted main effect of face race, $F(1, 599) = 102.67$, $p < .001$, $\eta_p^2 = 0.15$, 90% CI [0.11, 0.19], $B_{incl} = 6.98e+13$, and an interaction between face race and

[16] Considering the sensitivity for CR faces instead of the CRE index as the dependent variable did not show any significant results.

experiment, $F(1, 599) = 5.47$, $p = .020$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.03], $B_{incl} = 1.43$. Follow-up contrasts revealed that while sensitivity for CR faces did not significantly differ between the two samples, $t(1058) = -1.19$, $p = .235$, sensitivity for SR faces was actually significantly better in the Portugal sample than in the United States sample, $t(1058) = 3.83$, $p < .001$.[17] Consistent with the results of the individual experiments, the interaction between instruction condition and face race was not significant, $F(1, 599) = 0.64$, $p = .425$, $\eta_p^2 = 0.001$, 90% CI [0.00, 0.01], $B_{incl} = 0.02$, and, more importantly for the present analysis, there was also no significant interaction between experiment, instruction condition, and face race, $F(1, 599) = 1.03$, $p = .311$, $\eta_p^2 = 0.002$, 90% CI [0.00, 0.01], $B_{incl} = 0.002$. We address the possible reasons why the cultural setting did not interact with individuation instructions in the General Discussion.[18]

7. General discussion

We conducted two pre-registered replications in two different cultural settings (i.e., the United States and Portugal) to clarify the role of individuation instructions on the CRE (Hugenberg et al., 2007). Specifically, we examined the following two hypotheses: (a) that instructing White participants to individuate CR faces (i.e., Black faces) enhances recognition performance for these faces, thus reducing or eliminating the CRD (Hugenberg et al., 2007; Experiments 1A and 1B); and (b) that the effectiveness of these instructions depends on participants' individuation experience with CR faces, such that participants with greater experience would more easily recognize CR faces (Young & Hugenberg, 2012). Altogether, the results of the individual experiments as well as that of a mega-analysis combining the data from both experiments did not support any of the hypotheses. In fact, Bayesian analyses revealed moderate to strong evidence *against* the two hypotheses.

Additionally, we also examined a third *exploratory* hypothesis concerning the potential moderating role of the cultural setting. Namely, we tested whether the participants from a more racially diverse culture (i.e., the United States) would more easily individuate CR faces when instructed to do so than the participants from a less diverse culture (i.e., Portugal). We found that, although the United States participants reported greater contact with CR faces than the Portuguese participants, as expected, their recognition performance for CR faces did not improve when instructed to individuate these faces. In other words, the cultural setting did not interact with individuation instructions and face race.

7.1. How can individuation instructions enhance CR face recognition

As mentioned in the Introduction, the basis for reducing the CRE via individuation instructions rests on the assumptions underlying the CIM (Hugenberg et al., 2010). Building on well-established social cognition models (Brewer, 1988; Fiske & Neuberg, 1990), the CIM proposes that SR faces and CR faces are typically processed and encoded in two qualitatively distinct ways. While SR faces are encoded based on identity-diagnostic facial information, CR faces are encoded based on category-diagnostic facial information. Because CR faces are encoded based on information shared by all category members, perceivers have greater difficulty differentiating previously seen CR faces from novel CR faces.

One central tenet of the model is that perceivers have the ability to redirect selective attention to the identity-diagnostic characteristics of CR faces, but they are not motivated to do so, perhaps because the identities of CR faces are not as personally or situationally relevant as are those of SR faces. In other words, perceivers have equal ability to individuate SR and CR faces but not equal motivation. Therefore, tests of the model have focused on manipulating motivation to individuate CR faces. Hugenberg et al.'s (2007) individuation instructions manipulation warns participants of the existence of CRE and asks them to pay close attention to what differentiates CR faces from one another. However, the present research failed to replicate these instructions' effect on the CRE (see Table 1 for other failed replications), which raises several important theoretical and methodological questions that we will discuss next.

Globally, the failure to reduce the CRE via individuation instructions may reflect a mismatch between the CIM's assumptions and the experimental manipulation or a problem with the model, itself. That is, even if the CIM is correct, individuation instructions may fail to motivate individuals to individuate CR faces (i.e., the issue lies within the implementational/methodological level). It is also possible that, regardless of the instructions' efficacy, the explanatory account for how motivation enhances CR face recognition is incorrect or incomplete (i.e., the issue lies at the theoretical/conceptual level).

In either case, it is crucial to consider how individuation instructions act to mitigate the CRE. Individuation instructions are supposed to diminish the gap between SR and CR recognition by providing participants awareness of the CRE, prompting them to rely on different processing strategies for CR faces. Thus, individuation instructions are believed to make participants go beyond their initial processing, reliant merely on categorization strategies, and put in additional effort to individuate each CR face as such, as they would naturally do for SR faces.

Even if the mechanisms put forward for better recognition of CR faces when individuals are motivated are true, the role of individuation instructions in shaping these is ambiguous. On the one hand, this ambiguity can be referred to as inherently implementational. The fact that individuation instructions are not directive indicates that participants are not necessarily guided in what they should do in order to improve CR face recognition. Though the CIM (and hybrid theories in general; see Young & Hugenberg, 2012) postulate that expertise and motivation are necessary but not sufficient conditions for adequate individuation, the success of individuation instructions in reducing the CRE rests on the assumption that individuals always hold the skill required to individuate CR faces, and only need the motivation to do so.

However, even if individuation instructions indeed do motivate participants to individuate CR faces, it does not necessarily follow that these instructions enable better recognition for CR faces as participants still might not have the tools to do so. The CIM argues for the existence of individuation cues that individuals pay more attention to when motivated to do so; additionally, the theory argues that participants are aware of what constitutes an appropriate individuation cue – or that participants know whether each feature is relevant or not when trying to individuate a face –, though it does not substantiate this claim. Even if motivated to individuate CR faces, individuals might fail to do so because they are using suboptimal strategies to reach this goal, such as focusing on features that, contrary to their beliefs, do not help discriminate across faces.

Thus, the fact that the CRE is not reduced under individuation instructions may reflect the lack of ability of participants to properly individuate CR faces even when motivated to do so. Our results lend provisional support for such an account, considering that participants were more motivated to individuate CR faces in the individuation (vs. control) condition, judging from an increased self-reported effort for

[17] This effect may be explained by the significantly greater level of SR contact reported by the Portuguese participants ($M_{Portugal} = 5.33$, $SD_{Portugal} = 0.68$ versus $M_{United\ States} = 5.10$, $SD_{United\ States} = 0.62$; $t(601) = 4.37$, $p < .001$).

[18] In the supplementary material, we report several other analyses involving the cultural setting, including a regression analysis with CR contact and individuation instructions as the other predictors of the CRE (difference score) that also yielded a non-significant three-way interaction, $b = -0.027$, $SE = 0.12$, $t(595) = -0.22$, $p = .824$.

these faces[19]; this suggests that the failure to improve CR face recognition is not centered on the motivational dimension of the explanatory framework, but may be related to inadequate strategy selection to improve performance for CR stimuli.

This possibility is reinforced by the ongoing debate on what constitutes a diagnostic feature (i.e., those that are central to correctly identifying face identity) for face recognition. The perceived faceness of non-face objects is dependent on the presence and/or salience of elements resembling either eyes or a mouth, thus suggesting that these are the critical elements subserving face detection (Omer, Sapir, Hatuka, & Yovel, 2019). Towler, Keshwa, Ton, Kemp, and White (2021) report that training methods focused on specific features – ears and facial marks – are highly effective in improving recognition, suggesting that some facial features might be more useful in driving accurate recognition judgments. Other features such as lip thickness and eye shape have been found to hold high perceptual sensitivity (i.e., these are dimensions in which it is easy to distinguish faces from one another) in White faces (Abudarham & Yovel, 2016). These judgments may rest on more than attending to features alone, considering the relationships between several features as well (e.g., such as those established between the eyes and the mouth, Schyns, Bonnar, & Gosselin, 2002).

The difference in importance of facial features for accurate recognition may translate into specific attention allocation patterns depending on the target race. For example, Goldinger, He, and Papesh (2009) identified feature gaze differences such that there were more fixations of the eyes for SR than for CR faces (for both White and Asian participants), while Ellis, Deregowski, and Shepherd (1975) identified asymmetries in facial descriptions, such that eye attributes are more highlighted for SR faces, whereas the lower half of the face is more mentioned for CR faces. Experimentally manipulating fixation location reveals that SR and CR face recognition is increased in White participants following high and low fixations, respectively (Hills & Lewis, 2006; Hills & Lewis, 2011, Hills & Pake, 2013, Study 2; but see Palma, Quarenta, Carvalho, Ma, & Correll, 2022 and Wittwer, Tredoux, Py, & Paubel, 2019 for failures to replicate these results).

However, recent findings suggest that feature diagnosticity is not contingent on the target face race. Namely, Correll, Ma, Kenny, Palma, and A. (2022) show that features (e.g., nose length and width, chin length, lip thickness) are as highly diagnostic for differentiating Black and White faces. In any case, participants were more capable of taking highly diagnostic information into account when providing judgments for SR (vs. CR) faces, which suggests that, though the features attended to are roughly the same, increased contact might still contribute to higher discriminative power within social categories.

Furthermore, individuation instructions per se do not enable one to directly assess whether there is a shift in features attended to for CR faces as a function of instructions (i.e., towards those naturally attended to in SR faces). Alternatively, increased reliance on idiosyncratic strategies may arise. Peterson and Eckstein (2013) found observer-specific optimal points of fixation in unconstrained face identification tasks, emerging from the moment the saccadic movement initiates (Leonards & Scott-Samuel, 2005). What is more, these idiosyncratic strategies have been shown to emerge recurrently and to be stable over time (Mehoudar, Arizpe, Baker, & Yovel, 2014). These individual-level strategies in face processing might contribute to variability in face recognition accuracy. Individuation instructions might promote reliance on these idiosyncratic strategies, since the instructions do not enforce the use of a specific strategy or cue; by definition, idiosyncratic strategies will tend to differ across individuals, and this uniqueness would reflect itself on the effectiveness of the strategies adopted, overshadowing the consistent improvement of CR face recognition expected if instructions promoted transversal optimal strategy usage. These limitations could be circumvented by controlling for individual differences (i.e., idiosyncratic strategies) in a blocked design in which individuals perform a recognition task both with and without individuation instructions, as such design would enable the identification of an individual baseline for both SR and CR face recognition.

Regardless of the ongoing debate on feature diagnosticity explored above, it is worth mentioning that previous accounts arguing for increasing motivation to individuate CR faces via instructions do not detail how individuals come to identify and/or focus on these relevant features. Kawakami et al. (2014) shows that individuation instructions reduce the CRE by increasing eye fixations in CR faces, suggesting that the preference for eyes in SR faces relative to CR faces predicts CRE. These findings have, nonetheless, been contested on two grounds. On the one hand, Stelter, Rommel, and Degner (2021) fail to find support for the idea that this bias towards fixating the SR faces' eyes contributes to a reduction of the CRE (i.e., as the preference for fixating eyes of SR vs. CR faces obtained was unrelated to recognition performance). On the other hand, Correll and Hudson (2020) show that Kawakami et al.'s (2014) analyses are misguided in the sense that the interaction argued for by the authors (i.e., increased attention allocation to the eyes of CR faces) actually reflects a general tendency to process the incentivized face more (i.e., increased attention allocation towards all CR face features), meaning that the increase in performance can be merely due to the tendency to devote more attention to all CR facial features (vs. the eyes, specifically).

Another possible explanation for the absence of enhanced recognition of CR faces in the individuation condition may be related to the nature of the mental representations of these faces. Namely, recent neuro-scientific evidence has shown that SR (White) and CR faces are differently represented in the brain, with more detailed individuated representations for the former than the latter (Reggev, Brodie, Cikara, & Mitchell, 2020). The less detailed representations of CR faces may be due to less experience or familiarity, as other recent work suggests that lower familiarity with faces is associated with impoverished mental representations of these faces (Zhan, Garrod, van Rijsbergen, & Schyns, 2019). Thus, besides not promoting the selection of relevant facial features, the individuation instructions may also not be promoting the creation of more detailed representations of CR faces.

Overall, our results suggest that the role of individuation instructions is still unclear and that multiple alternative explanations for the absence of their impact may exist, both at the theoretical and implementational levels. In this vein, and in order to achieve a complete understanding of the CRE, it is crucial to obtain a greater understanding of what constitutes effective face discrimination processes and how motivation could harness these to increase accuracy in CR face recognition. Only then will we be capable of appropriately evaluating the success of individuation instructions in reducing the CRE from both theoretical and implementational standpoints.

7.2. The role of CR contact and motivation to individuat

Another key prediction of the CIM is that the effectiveness of individuation motives in shifting selective attention towards identity-diagnostic characteristics of CR faces is determined by perceivers' individuation experience or expertise with CR faces, such that perceivers with more experience will more easily attend to and use those identity-diagnostic facial characteristics to recognize CR faces (Young & Hugenberg, 2012). In the current experiments we also examined this prediction and failed to find support for it (for similar findings, see Tullis

[19] Our results on self-reported effort are inconsistent with the CIM in that we found higher reported effort for CR faces even in the control condition, while the model claims that lower performance stems from reduced effort dedicated to CR faces. Though contradicting the CIM, the general additional effort reported for CR faces has been previously reported (i.e., Tullis et al., 2014; Wan et al., 2015), and argues for an experience-based account of the CRE on the grounds that increased effort for CR faces in control conditions represents explicit recognition of higher task difficulty due to low contact with the social category.

et al., 2014). In fact, we did not even obtain a significant relationship between CR contact and the CRE.

In understanding these results, one should take into account nuanced considerations on the role of contact in the CRE. In fact, contact only explains about 2% of variation in the CRE (for a recent meta-analysis on contact and the CRE, see Singh et al., 2021), and has been regarded as a highly specific contributing factor to the CRE, with contact quality being more relevant than quantity, especially during a critical period in development (i.e., 6 to 12 years-old). As per the stated above, we may have failed to find a decrease in the CRE associated with increased CR contact due to either its modest impact in the CRE, which would make it hard to detect consistently across studies, or due to focusing on a general definition of contact (i.e., disregarding the quality vs. quantity distinction and ignoring when said contact took place), though previous positive results also failed to endorse these fine-grained considerations on the contact construct (e.g., Wan et al., 2015; Young & Hugenberg, 2012).

Another possible explanation for our findings could be tied to sample characteristics. If our samples were to not vary substantially in reported contact with CR individuals, we could fail to identify the association between contact and the CRE even if it does exist, as the restriction of variance in a predictor compromises power. Though restriction of variance could always be an issue, it is worth noting that our samples did show meaningful variation in self-reported CR in both Experiments 1A ($SD = 0.77$, minimum $= 1.00$, maximum $= 5.38$, $IQR = 0.875$) and 1B ($SD = 0.85$, minimum $= 1.00$, maximum $= 4.88$, $IQR = 1.125$; in a 7-point Likert-type scale).

In any case, even if reported contact with CR individuals is low and/or not diverse for both cultures, we know from objective measures that ethnic diversity differs in our two samples. This is not, however a sufficient condition for increased contact with CR individuals, as participants in an ethnically diverse context may not interact with CR individuals (e.g., Anicich, Jachimowicz, Osborne, & Phillips, 2021). The case for the increased contact with CR individuals in our American (vs. Portuguese) sample can be argued for on the basis of the exclusions based on ethnicity (Experiment 1A – United States: $n = 131$ vs. Experiment 1B – Portugal: $n = 6$). Considering that our study argues for differing levels of contact at both the individual- and group-levels, in both theoretical and data-informed ways, and that still no differences in the CRE emerged across the two samples is thus robust evidence that this factor fails to moderate the CRE.

Finally, it is worth noting that contact is not a sufficient condition for better CR identification, as the idea that expertise contributes to improved discrimination of features does not equate to claiming that individuals will become more adept at discriminating diagnostic features. Better feature discrimination will only equate to more accurate face discrimination if individuals know which features to pay attention to; put differently, the fact that contact contributes to expertise and more fine-grained feature identification is not necessarily associated with developing awareness of which distinct facial features are useful for face individuation. Though the CIM claims that increased experience or expertise with CR faces would lead to decreased CRE under motivation conditions, this should only hold true if the cues attended to are relevant for correct face individuation.

7.3. Individuation motivation across cultures

Contextual sensitivity (i.e., the extent to which a topic is sensitive to contextual variability) has been used as an explanation for not revising original findings after replications fail to converge with them, but it can also constitute a drive for model testing, when models' mechanisms are somehow related to or modulated by these contextual factors (Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016). In the present studies, we used the informativeness of culture as a contextual dimension to test for the socially embedded nature of the CRE as argued for in hybrid accounts of the CRE (e.g., Hugenberg et al., 2010; Wan et al., 2015).

Hybrid explanations for the CRE weigh not only the role of expertise (thoroughly discussed in the previous subsection), but also the influence of ethnicity as a (potentially relevant) social category, which allows for additional tests of the framework's assumptions. In this vein, cultural differences across samples should be associated with differences in the extent to which the CRE emerges. Differing levels of ethnic diversity should promote differences at both the perceptual and motivational facets of hybrid explanations. On the one hand, diversity leads to more contact with CR faces, which is associated with increased expertise in individuating CR individuals; on the other hand, CR individuals are more relevant as a social category the more they are expected to interact with the perceiver, prompting motivation to individuate CR faces. Culture should therefore be an important modulating factor of the CRE, not only because it can motivate individuals to individuate CR faces, but also because it can enhance the accuracy of these judgments.

Wan et al. (2015) focused on culture as a means to control for differences in social status between SR and CR groups, while accounting for differing levels of contact with CR faces (i.e., by including SR participants raised in CR cultures – Western-raised Asians). They found that cultural differences in upbringing were relevant for the CRE, such that the effect did not emerge when participants were raised in CR contexts (i.e., with extensive contact with CR faces). In our experiments, we failed to replicate the elimination of the CRE as a function of cultural background. It might be the case that our cultural settings are not distinct enough to impact the CRE, and Wan et al. (2015) used starkly distinct cultures (i.e., Whites and Asians) while considering both endpoints of plausible contact with CR faces (i.e., participants raised in SR- vs. CR-dominant cultures). The absence of social-motivational components in the CRE found in Wan et al. (2015) has been taken as evidence in favor of expertise accounts, and the authors have argued that differences in the CRE relative to the usual White-Black comparison (i.e., as considered in our Experiments) might have a social-motivational basis centered on social status imbalance between SR and CR stimuli category. In any case, all these findings argue against one of the assumptions of the CIM, namely the idea that motivation to individuate would be more important for individuals high in contact with CR faces (i.e., as these would be the ones capable of mobilizing that expertise due to their motivation).

8. Conclusion

In an impactful paper, Hugenberg et al. (2007) showed that instructing participants to individuate CR faces eliminated the CRE by improving participants' recognition of these faces. In the present paper, we attempted to replicate such an effect but failed to find evidence that individuation instructions mitigate the CRE. The CRE emerged regardless of whether the sample came from a country high (i.e., United States) or low (i.e., Portugal) in racial diversity. Moreover, we also failed to replicate the finding that individuation instructions are particularly effective for the participants with more contact with CR individuals (Young & Hugenberg, 2012). Together, our results call into question the robustness of individuation instructions manipulation as a means to eliminate (or even attenuate) the CRE.

Author note

The first two authors share first authorship as they contributed equally to this project. The order of the first two authors was based on reverse seniority, while the order of the remaining authors reflects their contribution to this project. We have no conflicts of interest to disclose. This research was partially supported by Fundação para a Ciência e a Tecnologia, I.P. (DL57/2016/CP[1439]/CT[3]) and by a Pre-registered Research Grant of the European Association of Social Psychology awarded to T. A. Palma.

Disclosure statement

Measures, manipulations, and exclusions are reported in-text, as well as sample size determination method. Supplementary material is available at: https://osf.io/f3vcm/?view_only=5af44520f7b14debbccf1c8bb03c0945

Open practices statement

Data collection was pre-registered on OSF; materials and methodological approach were determined a priori and are accessible via the following link, in which data and analysis code is also available: https://osf.io/f3vcm/?view_only=5af44520f7b14debbccf1c8bb03c0945. Deviations from the pre-registered plan are identified in-text.

CRediT authorship contribution statement

Francisco Cruz: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Tomás A. Palma:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Emil Bansemer:** Software, Validation, Formal analysis, Investigation, Data curation, Writing – review & editing, Supervision. **Joshua Correll:** Software, Validation, Formal analysis, Investigation, Data curation, Writing – review & editing, Supervision, Project administration. **Sara Fonseca:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Patrícia Gonçalves:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Ana Sofia Santos:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors have no conflict of interests to declare.

Data availability

Materials, data, and analysis script are publicly available in the Open Science Framework (OSF) platform (see links in the paper).

Acknowledgements

We thank the reviewers for their insightful comments and suggestions.

Appendix A

To compute the power simulations reported in this paper, we relied on a set of recent unpublished data from our lab. These data were obtained from three experiments (total $N = 327$) using White Portuguese participants and face stimuli selected from the same face database (Ma et al., 2015) as the one we intend to use in the present research. In these experiments, participants saw 60 faces (30 SR and 30 CR) and made recognition predictions (i.e., judgments of learning) for each of them during a study phase. Following a brief filler task, participants completed an old-new recognition test composed of the same 60 old faces and 60 (30 SR and 30 CR) lure faces (i.e., new).

A combined analysis of the data (*d'*) of these experiments yielded significant main effect of race, $F(1, 324) = 120.15$, $p < .001$, $\eta_p^2 = 0.27$, such participants correctly recognized more SR faces ($M = 1.96$, $SD = 0.74$) than CR faces ($M = 1.62$, $SD = 0.71$). For the simulations, we used the obtained SR and CR average *d'* scores (1.96 and 1.62), the correlation between these scores ($r = 0.73$), and the standard deviation ($SD = 0.63$) associated with the sample average *d'* score (1.79). See the Power Simulation folder at the OSF (https://osf.io/f3vcm/?view_only=5af44520f7b14debbccf1c8bb03c0945) for the simulations R script and results.

Appendix B. Racial contact questionnaire

For the following questionnaire, we would like you to indicate how well the following statements represent the type of interactions you have with Black and White people. Please indicate the extent to which each statement represents your interactions by crossing out the number which best represents your opinion.

Very strongly disagree	Strongly disagree	Disagree	Agree	Strongly agree	Very strongly agree
1	2	3	4	5	6
(1) I know lots of Black people ………………………………………...					1 2 3 4 5 6
(2) I interact with White people during recreational periods …….....					1 2 3 4 5 6
(3) I live, or have lived in an area where I interact with White people					1 2 3 4 5 6
(4) I live, or have lived in an area where I interact with Black people ...					1 2 3 4 5 6
(5) I interact with Black people during recreational periods …………...					1 2 3 4 5 6
(6) I interact with White people on a daily basis …………………..					1 2 3 4 5 6
(7) I socialize a lot with White people …………………………….					1 2 3 4 5 6
(8) I went to a high school where I interacted with Black students …...					1 2 3 4 5 6
(9) I socialize a lot with Black people ………………………………….					1 2 3 4 5 6
(10) I know lots of White people …………………………………..					1 2 3 4 5 6
(11) I generally only interact with Black people ……………………….					1 2 3 4 5 6
(12) I interact with Black people on a daily basis ……………………...					1 2 3 4 5 6
(13) I went to a high school where I interacted with White students					1 2 3 4 5 6
(14) I generally only interact with White people …………………					1 2 3 4 5 6
(15) I have lived in an African country where the predominant race is Black ………………………………………………………….....					1 2 3 4 5 6

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2022.104423.

References

Abudarham, N., & Yovel, G. (2016). Reverse engineering the face space: Discovering the critical features for face identification. *Journal of Vision, 16*, 40. https://doi.org/10.1167/16.3.40

Anicich, E. M., Jachimowicz, J. M., Osborne, M. R., & Phillips, L. T. (2021). Structuring local environments to avoid racial diversity: Anxiety drives Whites' geographical and institutional self-segregation preferences. *Journal of Experimental Social Psychology, 95*, Article 104117. https://doi.org/10.1016/j.jesp.2021.104117

Baldwin, M., Keefer, L. A., Gravelin, C. R., & Biernat, M. (2013). Perceived importance of cross-race targets facilitates recall: Support for a motivated account of face memory. *Group Processes & Intergroup Relations, 16*, 505–515. https://doi.org/10.1177/1368430212460893

van den Bergh, D., Van Doorn, J., Marsman, M., Draws, T., Van Kesteren, E. J., Derks, K., … Wagenmakers, E. J. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *LAnnee psychologique, 120*, 73–96. https://doi.org/10.3917/anpsy1.201.0073

Bornstein, B. H., Laub, C. E., Meissner, C. A., & Susa, K. J. (2013). The cross-race effect: Resistant to instructions. *Journal of Criminology, 2013*, 1–6. https://doi.org/10.1155/2013/745836

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull, & R. S. Wyer, Jr. (Eds.), *A dual process model of impression formation* (pp. 1–36). Hillsdale, NJ: Erlbaum.

Brigham, J. C., & Malpass, R. S. (1985). The role of experience and contact in the recognition of faces of own and other race persons. *Journal of Social Issues, 41*, 139–155. https://doi.org/10.1111/j.1540-4560.1985.tb01133.x

Champely, S. (2018). pwr: Basic functions for power analysis. R package version 1.3–0. https://CRAN.R-project.org/package=pwr.

Correll, J., & Hudson, S. M. (2020). An error in the analysis of "an eye for the I". *Journal of Personality and Social Psychology, 119*, 1030–1036. https://doi.org/10.1037/pspa0000200

Correll, J., Ma, D. S., Kenny, D. A., Palma, T., & A.. (2022). *Examining the contribution of physical cues for same and cross-race face individuation*. Under review.

Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Descriptions of white and black faces by white and black subjects (1). *International Journal of Psychology, 10*, 119–123. https://doi.org/10.1080/00207597508247325

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth, 8*, 195–222. https://doi.org/10.1023/A:1024419522867

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. https://doi.org/10.1016/S0065-2601(08)60317-2

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training "greeble" experts: A framework for studying expert object recognition processes. *Vision Research, 38*, 2401–2428. https://doi.org/10.1016/S0042-6989(97)00442-2

Giner-Sorolla, R. (2018). Powering your interaction. Approaching significance—A methodology blog for social psychology. Retrieved from https://approachingblog.wordpress.com/2018/01/24/powering-your-interaction-2/.

Goffaux, V., & Rossion, B. (2006). Faces are "spatial"—Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1023–1039. https://doi.org/10.1037/0096-1523.32.4.1023

Goldinger, S. D., He, Y., & Papesh, M. H. (2009). Deficits in cross-race face learning: Insights from eye movements and pupillometry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1105–1122. https://doi.org/10.1037/a0016548

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *British Journal of Psychology, 99*, 45–56. https://doi.org/10.1348/000712607X199981

Hills, P. J., & Lewis, M. B. (2006). Short article: Reducing the own-race bias in face recognition by shifting attention. *Quarterly Journal of Experimental Psychology, 59*, 996–1002. https://doi.org/10.1080/17470210600654750

Hills, P. J., & Lewis, M. B. (2011). Reducing the own-race bias in face recognition by attentional shift using fixation crosses preceding the lower half of a face. *Visual Cognition, 19*, 313–339. https://doi.org/10.1080/13506285.2010.528250

Hills, P. J., & Pake, J. M. (2013). Eye-tracking the own-race bias in face recognition: Revealing the perceptual and socio-cognitive mechanisms. *Cognition, 129*, 586–597. https://doi.org/10.1016/j.cognition.2013.08.012

Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of Experimental Social Psychology, 43*, 334–340. https://doi.org/10.1016/j.jesp.2006.02.010

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the cross-race recognition deficit. *Psychological Review, 117*, 1168–1187. https://doi.org/10.1037/a0020463

Jeffreys, H. (1961). *Theory of probability*. Oxford, England: Oxford University Press.

Jerovich, A. (2017). Implicit attitudes and the other race effect (Doctoral dissertation, Western Sydney University, Sydney, Australia). Retrieved from https://researchdirect.westernsydney.edu.au/islandora/object/uws:43067/datastream/PDF/view.

Kawakami, K., Williams, A., Sidhu, D., Choma, B., Rodriguez-Bailon, R., Canadas, E., et al. (2014). An eye for the I: Preferential attention to the eyes of ingroup members. *Journal of Personality and Social Psychology, 107*, 1–20. https://doi.org/10.1037/a0036838

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00863

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science, 4*(1). https://doi.org/10.1177/2515245920951503, 2515245920951503.

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.

Leonards, U., & Scott-Samuel, N. E. (2005). Idiosyncratic initiation of saccadic face exploration in humans. *Vision Research, 45*, 2677–2684. https://doi.org/10.1016/j.visres.2005.03.009

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*, 89–94. https://doi.org/10.1038/82947

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General, 129*(4), 559–574. https://doi.org/10.1037/0096-3445.129.4.559

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*, 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Macmillan, N., & Creelman, C. (1991). *Detection theory: A user's guide*. Cambridge, UK: Cambridge University Press.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology, 13*, 330.

Mehoudar, E., Arizpe, J., Baker, C. I., & Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision, 14*, 6. https://doi.org/10.1167/14.7.6

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.

Meissner, C. A., Brigham, J. C., & Butz, D. A. (2005). Memory for own- and other-race faces: A dual-process approach. *Applied Cognitive Psychology, 19*, 545–567. https://doi.org/10.1002/acp.1097

Michel, C., Caldara, R., & Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition, 14*, 55–73. https://doi.org/10.1080/13506280500158761

Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., … Vazire, S. (2019). Pre-registration is hard, and worthwhile. *Trends in Cognitive Sciences, 23*, 815–818. https://doi.org/10.1016/j.tics.2019.07.009

Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLoS Biology, 18*, Article e3000691. https://doi.org/10.1371/journal.pbio.3000691

Omer, Y., Sapir, R., Hatuka, Y., & Yovel, G. (2019). What is a face? Critical features for face detection. *Perception, 48*, 437–446. https://doi.org/10.1177/0301006619838734

Palma, T. A., Quarenta, J., Carvalho, M., Ma, D. S., & Correll, J. (2022). *Can the own-race bias be reduced by shifting attention to the lower half of faces? A pre-registered replication of Hills & Lewis (2011).*. Manuscript in preparation.

Peterson, M. F., & Eckstein, M. P. (2013). Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychological Science, 24*, 1216–1225. https://doi.org/10.1177/0956797612471684

Pica, E., Warren, A. R., Ross, D. F., & Kehn, A. (2015). Choosing your words and pictures wisely: When do individuation instructions reduce the cross-race effect? *Applied Cognitive Psychology, 29*, 360–368. https://doi.org/10.1002/acp.3112

Reggev, N., Brodie, K., Cikara, M., & Mitchell, J. P. (2020). Human face-selective cortex does not distinguish between members of a racial outgroup. *eNeuro, 7*. ENEURO.0431-19.2020 https://10.1523/ENEURO.0431-19.2020.

Rhodes, G., Locke, V., Ewing, L., & Evangelista, E. (2009). Race coding and the other-race effect in face recognition. *Perception, 38*, 232–241. https://doi.org/10.1068/p6110

Rodin, M. J. (1987). Who is memorable to whom: A study of cognitive disregard. *Social Cognition, 5*, 144–165. https://doi.org/10.1521/soco.1987.5.2.144

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science, 13*(5), 402–409. https://doi.org/10.1111/1467-9280.00472

da Silva Frost, A., & Ledgerwood, A. (2020). Calibrate your confidence in research findings: A tutorial on improving research methods and practices. *Journal of Pacific Rim Psychology, 14*, Article e14. https://doi.org/10.1017/prp.2020.7

Singh, B., Mellinger, C., Earls, H. A., Tran, J., Bardsley, B., & Correll, J. (2021). Does cross-race contact improve cross-race face perception? A meta-analysis of the cross-race deficit and contact. *Personality and Social Psychology Bulletin, 48*(6), 865–887. https://doi.org/10.1177/01461672211024463

Steiger, J. H. (2004). Beyond the F test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychological Methods, 9*(2), 164–182. https://doi.org/10.1037/1082-989X.9.2.164

Stelter, M., Rommel, M., & Degner, J. (2021). (eye-) tracking the other-race effect: Comparison of eye movements during encoding and recognition of ingroup faces with proximal and distant outgroup faces. *Social Cognition, 39*, 366–395. https://doi.org/10.1521/soco.2021.39.3.366

Towler, A., Keshwa, M., Ton, B., Kemp, R. I., & White, D. (2021). Diagnostic feature training improves face matching accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(8), 1288–1298. https://doi.org/10.1037/xlm0000972

Tullis, J. G., Benjamin, A. S., & Liu, X. (2014). Self-pacing study of faces of different races: Metacognitive control over study does not eliminate the cross-race recognition effect. *Memory & Cognition, 42*, 863–875. https://doi.org/10.3758/s13421-014-0409-y

Tüttenberg, S. C., & Wiese, H. (2021). Recognising other-race faces is more effortful: The effect of individuation instructions on encoding-related ERP Dm effects. *Biological Psychology, 158*, Article 107992. https://doi.org/10.1016/j.biopsycho.2020.107992

Valentine, T. (2001). Face-space models of face recognition. In M. J. Wenger, & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges. Scientific psychology series* (pp. 83–113). Mahwah, NJ: Erlbaum.

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences, 113*, 6454–6459. https://doi.org/10.1073/pnas.1521897113

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review, 14*(5), 779–804. https://doi.org/10.3758/BF03194105

Wan, L., Crookes, K., Reynolds, K. J., Irons, J. L., & McKone, E. (2015). A cultural setting where the other-race effect on face recognition has no social–motivational component and derives entirely from lifetime perceptual experience. *Cognition, 144*, 91–115. https://doi.org/10.1016/j.cognition.2015.07.011

Wittwer, T., Tredoux, C. G., Py, J., & Paubel, P. V. (2019). Training participants to focus on critical facial features does not decrease own-group bias. *Frontiers in Psychology, 10*, 2081. https://doi.org/10.3389/fpsyg.2019.02081

Young, S. G., Bernstein, M. J., & Hugenberg, K. (2010). When do own-group biases in face recognition occur? Encoding versus post-encoding. *Social Cognition, 28*, 240–250. https://doi.org/10.1521/soco.2010.28.2.240

Young, S. G., & Hugenberg, K. (2012). Individuation motivation and face experience can operate jointly to produce the own-race bias. *Social Psychological and Personality Science, 3*, 80–87. https://doi.org/10.1177/1948550611409759

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the cross-race effect. *Personality and Social Psychology Review, 16*, 116–142. https://doi.org/10.1177/1088868311418987

Zhan, J., Garrod, O. G., van Rijsbergen, N., & Schyns, P. G. (2019). Modelling face memory reveals task-generalizable representations. *Nature Human Behaviour, 3*, 817–826. https://doi.org/10.1038/s41562-019-0625-3