**This is the preformatted version of the following article:**

Töredi, D.; Mansour, J. K.; Jones, S. E.; Skelton, F.; McIntyre, A. Creating a Cross-Race Effect Inventory to Postdict Eyewitness Accuracy [In Press]. *Law and Human Behaviour*

**This article will be published in final form at:** https://psycnet.apa.org/PsycARTICLES/journal/lhb

**Creating a Cross-Race Effect Inventory to Postdict Eyewitness Accuracy**

Dilhan Töredi1, Jamal K. Mansour1, 2, Sian E. Jones1, Faye Skelton3, and Alex McIntyre3

1 Department of Psychology, Sociology, & Education, Queen Margaret University

2 Department of Psychology, University of Lethbridge

3 School of Applied Sciences, Edinburgh Napier University

**Author Notes**

Dilhan Töredi A green circle with white letters

Description automatically generated <https://orcid.org/0000-0001-8420-1245>

Jamal K. Mansour A green circle with white letters

Description automatically generated <https://orcid.org/0000-0001-7162-8493>

Sian E. Jones A green circle with white letters

Description automatically generated <https://orcid.org/0000-0002-2399-1017>

Faye Skelton A green circle with white letters

Description automatically generated <https://orcid.org/0000-0003-4792-4238>

Alex McIntyre A green circle with white letters

Description automatically generated <https://orcid.org/0000-0002-8470-4886>

This study was conducted as a part of the first authors’ PhD studies at Queen Margaret University.

Dilhan Töredi is now at the Department of Psychology, John Jay College of Criminal Justice.

The authors of this manuscript do not have conflicts of interest to disclose.

Correspondence concerning this article should be addressed to Dilhan Töredi via email: [dtoredi@jjay.cuny.edu](http://dtoredi@jjay.cuny.edu).

This study is preregistered ([osf.io/j9nzc](https://doi.org/10.17605/OSF.IO/J9NZC)), and the raw data and analysis scripts are shared ([osf.io/nx5mf/](https://osf.io/nx5mf/?view_only=)) on the Open Science Framework.

**Abstract**

**Objective:** The Cross-Race Effect (CRE) is a reliable and robust phenomenon, whereby individuals better recognize faces that belong to their race compared to another race. Our goal was to produce items for a self-report Inventory (i.e., CRE-I) that brings together known predictors of the CRE to improve the postdiction of cross-race eyewitness accuracy.

**Hypotheses:** We expected a CRE for White and Asian participants. We anticipated that developed CRE-I subscales would correlate positively with extant (some modified) scales and predict accuracy.

**Method:** Participants completed four trials (two White targets and two Asian targets). For each trial, they watched a mock crime video, performed a distractor task, made a sequential lineup decision (target-present or target-absent), and indicated confidence in their lineup decision. After all trials, participants completed the potential items for the CRE-I.

**Results:** We replicated prior findings of a CRE for White participants but did not find a CRE for Asian participants. Exploratory factor analysis produced internally reliable scales for the CRE-I to be used with White eyewitnesses: general face recognition ability, race-specific face recognition ability, racial attitudes, quantity of contact, quality of contact, motivated individuation, and cognitive disregard. Responses to several scales predicted identification accuracy. In particular, three CRE-I scales predicted identification accuracy beyond the predictiveness of confidence: race-specific face recognition ability, racial attitudes towards White people, and motivated individuation of White people.

**Conclusions:** Variables suggested separately by the perceptual expertise hypothesis and the social cognitive hypothesis predicted identification accuracy, providing support for integrative models of the CRE. The CRE-I contributes to the CRE literature both in terms of theory—by showing which factors among many may best relate to recognition—and practice—by improving evaluations of eyewitness reliability.

*Keywords:* the cross-race effect, motivation, measurement, interracial contact, confidence

**Public Significance Statement**

We explored how people recognize faces from different races and developed items for a practical tool, the Cross-Race Effect Inventory (CRE-I), which could be used to measure this ability after future validation. The CRE-I incorporates theory-driven, user-friendly scales that enhance our understanding of individual differences influencing race-based differences in face recognition. The development of items for the CRE-I is a first step towards providing a valuable tool to the criminal justice system. Once validated, the CRE-I could aid in assessing an eyewitness's reliability, particularly in cases involving interracial crimes, and help prevent miscarriages of justice arising from misidentifications.

**Creating a Cross-Race Effect Inventory to Postdict Eyewitness Accuracy**

In 1979, a White woman, B.N., was raped by a Black man in her store (Innocence Project, 2021). Later, Malcolm Alexander, a Black man, was sentenced to life imprisonment after B.N. identified him from a lineup. Twenty-eight years later, DNA evidence established Alexander's innocence, and he was exonerated. Like Alexander, 3478 people in the US have been wrongly convicted (National Registry of Exonerations, 2024). Notably, 27% (957) of these wrongful convictions involved inaccurate eyewitness identifications and a considerable proportion also involved people of different races. To date, 31% of murder exonerations and 51% of sexual assault exonerations have involved a White victim identifying a Black defendant.

Sometimes eyewitnesses make errors and one reason for these errors is the cross-race effect (CRE, i.e., cross-race bias, other-race bias, own-race bias, or other-race effect). Eyewitnesses tend to better remember individuals with whom they share (vs. do not share) common features, such as race, age, and gender (Young et al., 2012). The CRE is defined as being better at recognizing same-race faces compared to faces of another race (also referred to as cross-race)[[1]](#footnote-1). This includes both being more likely to identify a guilty suspect correctly and being less likely to identify an innocent suspect incorrectly (Meissner & Brigham, 2001).

Meta-analyses (e.g., Lee & Penrod, 2022; Meissner & Brigham, 2001) affirm the robustness and reliability of the CRE. Additionally, numerous field studies using naturalistic scenarios support its existence (e.g., Platz & Hosch, 1988; Wright et al., 2001). Relevant to the current study, the CRE has been found for Asian and White participants for White and Asian faces (e.g., Ge et al., 2009; Hancock & Rhodes, 2008; Walker & Tanaka, 2003).

The CRE has real-world impact—from everyday social interactions between people (McKone et al., 2023) to failures of eyewitness recognition that may result in wrongful convictions (Scheck et al., 2003; Thompson, 2007; Wells & Olson, 2001). Reliable measurement of the CRE is a crucial step. We assessed the CRE with two participant race groups (White, Asian) and used psychometrics to develop a self-report inventory.

**Theories of the CRE**

Research investigating the CRE has resulted in several key theoretical positions (for a detailed review, see Young et al., 2012): the *perceptual expertise hypothesis* (Meissner & Brigham, 2001), the *social-cognitive hypothesis* (Anthony et al.,1992; Levin, 2000), and *integrative models* (Hugenberg et al., 2010; Meissner, Brigham, & Butz, 2005; Sporer, 2001).

The perceptual expertise hypothesis suggests that differences in face recognition performance stem from discrepancies in encoding strategies acquired through differential contact with same-race and cross-race individuals (e.g., Meissner & Brigham, 2001; Meissner et al., 2005; Young et al., 2012). Due to greater exposure, people accumulate expertise in recognizing same-race faces. In contrast, due to limited exposure, people accumulate lower levels of expertise in recognizing cross-race faces (MacLin & Malpass, 2001). As a result, face-encoding strategies are only optimized for processing same-race faces. When these strategies are applied to cross-race faces, they are ineffective, resulting in the CRE. Evidence supportive (e.g., Tanaka et al., 2004; Hills & Lewis, 2006) and unsupportive (e.g., Hugenberg et al., 2007; Walker & Hewstone, 2006, 2008) exists.

In contrast, the social-cognitive hypothesis (Levin, 2000) posits that the CRE arises from differences in how socially relevant information is processed when it concerns ingroup (one's own group) versus outgroup members (those belonging to a different group; Anthony et al., 1992). People tend to categorize outgroup members, like cross-race faces, by social group features such as race (categorization) while identifying ingroup members, like same-race faces, by unique facial features (individuation). Because categorization occurs at the expense of individuation, cross-race faces are encoded superficially and processed less effectively than same-race faces. This impairs one’s ability to distinguish cross-race faces, resulting in the CRE (Levin, 2000). Again, evidence supportive (e.g., Bernstein et al., 2007; Shriver et al., 2008) and unsupportive (e.g., Rhodes et al., 2009, 2010) exists.

Integrative models combine components from the perceptual expertise hypothesis and the social-cognitive hypothesis (see Young et al., 2012). The categorization-individuation model (CIM; Hugenberg et al., 2010) is a highly tested integrative model relevant to this study. The CIM suggests that while same-race faces are individuated, cross-race faces are categorized at the expense of individuation, leading to differences in the processing of same-race and cross-race faces, and thus the CRE. It further proposes that the CRE can be mitigated by increased contact with a racial group because people will accumulate expertise at recognizing faces, which will increase attendance to individuating features, and therefore improve recognition. However, for this to happen, the person must have the intention to differentiate faces accurately (i.e., motivated individuation, elaborated later). Supportive (e.g., Hugenberg et al., 2007) and unsupportive evidence (e.g., Cruz et al., 2023) exists.

Additionally, it is important to consider *the contact hypothesis* (Ng & Lindsay, 1994), which integrates a variety of hypotheses about the CRE and which is compelling in part because of its high face validity. This hypothesis suggests that limited contact with people of different races contributes to the CRE, highlighting the importance of increased experience with other races to reduce the CRE (Ng & Lindsay, 1994). It differs from other hypotheses of the CRE but also exists alongside them because it posits a *prerequisite* for effective facial perception rather than explaining why facial perception or recognition may be ineffective. Meta-analyses of the CRE support the contact hypothesis as explaining a small but significant amount of variance in the CRE (Meissner & Brigham, 2001; Singh et al., 2022).

Determining the relative primacy of these theoretical positions presents challenges, as they often yield similar predictions (Sporer, 2001; Young et al., 2012). However, together, they propose several measurable variables as contributors to the CRE. A comprehensive understanding of the CRE necessitates an exploration of these variables. By doing so, insights can be gained into the mechanisms underlying proficient face recognition and the ability to perceive individuals from different race groups (Sporer, 2001; Valentine, 1991), which can be used to develop reliable instruments for informing criminal justice practices.

**Variables Influencing the CRE**

A review of the CRE literature for variables reliably related to the CRE revealed six measurable dimensions: general face recognition ability (Rhodes et al., 2014), race-specific face recognition ability (Hourihan et al., 2012), racial biases and attitudes (Walker & Hewstone, 2008), the quantity and quality of interracial contact (Singh et al., 2022), cognitive disregard (Marsh et al., 2016), and motivated individuation (Hugenberg et al., 2007). Various scales, including the Racial Attitudes scale (Brigham, 1993) and the Positive and Negative Contact Scale (Hayward et al., 2017), have been used to predict some of these dimensions. However, extant scales are quite old and/or *ad hoc* within studies without an appreciation of the principles of scale development and focused only on White and Black faces. Furthermore, some of these dimensions (e.g., cognitive disregard) have not been explored via self-report. We examined whether self-reports along these theoretically-driven dimensions enhance the ability to postdict eyewitness accuracy in cross-race—and same-race—circumstances. Our goal was to combine reliable predictors of the CRE within a self-report inventory, the Cross-Race Effect Inventory (i.e., CRE-I), to improve postdiction of lineup accuracy.

**General Face Recognition Ability**

Face recognition involves complex processes influenced by perceptual, cognitive, and social factors. Although not explicitly proposed by any CRE hypothesis, face recognition ability is important for distinguishing and categorizing individuals, and therefore the CRE. Perceptual experience, cognitive processing, and attentional biases can impact general face recognition ability (Bruce & Young, 1986; Caldara & Fusaroli, 2016; Eastwood et al., 2001). Thus, it stands to reason that they will also affect the magnitude and occurrence of the CRE.

Yet, general face recognition ability (e.g., the Glasgow Face Matching Test, GFMT; Burton, White & McNeill, 2010) often does not predict the CRE (e.g., Correll, Ma, & Davis, 2021). Further, high-performers do as well with same-race and cross-race faces (Robertson et al., 2020, Study 1) and are similar to low-performers (Bate et al., 2019). In contrast, Slone, Brigham, and Meissner (2010) found a relationship between general face recognition ability and same-race (White) but not cross-race (Black) face recognition, while Rhodes et al. (2014) found a positive correlation between general face recognition ability and the CRE for White participants recognizing Asian faces. Hence, research is scarce and contradictory.

Self-report scales of general face recognition ability are often used to study prosopagnosia and super-recognition (e.g., Bobak et al., 2017; Gray, Bird, & Cook, 2017; Shah et al., 2015). They have not yet been utilized to study the CRE. By incorporating a general face recognition ability scale into the CRE-I, we not only assessed whether this ability predicts the CRE but also controlled for individual differences in face processing, thereby enhancing the possibility that our inventory will be valid and reliable.

**Race-Specific Face Recognition Ability**

Incorporating a scale for race-specific face recognition allows assessment of how race influences face recognition. Currently, there are no race-specific face recognition ability self-report scales. Several studies adopted pre-identification ratings of memory and judgements of learning as predictors of the CRE. For instance, Smith, Stinson, and Prosser (2004) reported that pre-identification ratings of memory clarity were higher for same-race than other-race faces, hinting at differences in memory judgments for faces of different races. Furthermore, Hourihan et al. (2012) found that people’s self-reports of the probability that they would recognize a face better predicted their ability to recognize same-race than cross-race faces.

The perceptual expertise hypothesis highlights differences in face perception between same- and cross-race faces as a contributor to the CRE. Specifically, the configural-featural processing model (O’Toole et al., 1995) posits that cross-race faces are processed featurally (facial features recognized independently), whereas same-race faces are processed configurally (faces recognized as unified wholes; Fallshore & Schooler, 1995; Jacques & Rossion, 2006; Tanaka et al., 2004). Because featural processing is less effective than configural processing for distinguishing between individuals (e.g., Fallshore & Schooler, 1995; Rhodes et al., 1989; Tanaka et al., 2004), the CRE occurs. Accordingly, we included items that were feature-specific (e.g., *I recognize someone is Asian/White by their eye shape*.) and configural (e.g., *I can tell someone is Asian/White by their skin color.)*

**Quantity and Quality of Interracial Contact**

Individuals usually have less contact with people of other races than people of their own race, either due to choice or geographic segregation (Reardon et al., 2008). The contact hypothesis suggests that the differences in contact contribute to the CRE and that increased contact with other races will reduce the CRE (Ng & Lindsay, 1994; Singh et al., 2022). Several studies support the contact hypothesis (e.g., Slone et al., 2000; Walker & Tanaka, 2003). Recently, across six studies, multiple participant race crossings (e.g., White and Black, White and Middle Eastern), and 30 measures of contact (e.g., GPS tracking, self-reports), Stelter et al. (2023) found weak but significant correlations between contact and the CRE. This finding is in line with meta-analyses that found increased interracial contact was associated with a small decrease in the CRE, explaining approximately 2% of the variability in the CRE across samples (Meissner & Brigham, 2001; Singh et al., 2022).

The relationship between type of contact and the CRE varies. Both quantity and quality of contact appear to predict the CRE (e.g., Stelter et al., 2023; Walker & Hewstone, 2006). For instance, Stelter et al. found that both higher self-reported quantity and quality (positive, negative) of contact were associated with a lower CRE (cf. lower contact and a higher CRE). However, other findings indicate quality plays a greater role than the quantity (e.g., Hancock & Rhodes, 2008; Lavrakas et al., 1976). For instance, Lavrakas et al. found that White participants who lived in racially homogeneous neighborhoods (low quantity of contact) with Black friends (high quality of contact) were better at recognizing Black faces than those who lived in racially heterogeneous neighborhoods (high quantity of contact) but did not have Black friends (low quality of contact). Other studies indicate contact is unrelated to the CRE (e.g., McKone et al., 2019). Thus, evidence for the contact hypothesis is mixed.

Recently, Singh et al. (2022) highlighted the lack of rigorous validation for many scales used to measure the contact hypothesis. They argued that scales are often developed *ad hoc* within studies testing the effects of contact rather than in studies with consideration of best practices for scale development. In certain scales, contact was evaluated with a single item, while in others, it was evaluated using multiple items perhaps intended to measure one or more constructs (e.g., quantity of contact, quality of contact, childhood contact, recent contact; Hancock & Rhodes, 2008; Slone et al., 2000; Walker & Hewstone, 2006). Variations in how contact is measured may contribute to contradictory findings. We aimed to produce a reliable measure of interracial contact that addresses these criticisms by including items measuring the quantity and quality of interracial contact.

**Racial Biases and Attitudes**

Biases about and attitudes towards different races may affect recognition (Wells & Olson, 2001). Early cross-race studies found that lower (cf. higher) prejudice was associated with better performance (e.g., Allport & Kramer, 1946; Lindzey & Rogolsky, 1950). Recent studies show that Black faces are recognized more accurately by White observers when labelled with high-status professions (Shriver & Hugenberg, 2010) while White faces are recognized less accurately when described as “poor Whites” (Shriver et al., 2008). These findings suggest racial biases/attitudes and the CRE are related. Race signals from a face may trigger racial attitudes and prejudices which indirectly affect recognition ability via encoding.

Multiple studies have tested whether individuals with positive racial attitudes towards different races are better at recognition using self-report measures (Racial Attitudes Scale, Sedlacek & Brooks, 1970; Multifactor Racial Attitude Inventory; Brigham, Woodmansee, & Cook, 1976; Racial Attitudes scale; Brigham, 1993; Attitudes to Asians Scale, Walker, 1994). These studies found no relationship between racial attitudes and cross-race face recognition (e.g., Brigham & Barkowitz, 1978; Lavrakas et al., 1976; Slone et al., 2000). For example, Ferguson et al. (2001) found no association between racial attitudes and the recognition performance of White and Asian participants for their respective cross-race face groups.

People are often reluctant to report or are unaware of their racial biases. Thus, several studies have measured implicit racial biases and attitudes—unconscious biases that influence attitudes and behaviors. For instance, Walker and Hewstone (2008) found that White and South Asian participants with more negative implicit racial attitudes about cross-race faces showed a stronger CRE. Using eye tracking, Anzures et al. (2022) found that while there was no significant relationship between implicit racial bias and the CRE; greater implicit racial bias was associated with more scrutiny of same-race internal facial features. However, Jackiw (2007) found no such relationship for Caucasian and First Nation samples using implicit (The IAT; Greenwald et al., 1998) and explicit (The EAT; Stephan & Stephan, 1993) measures.

Therefore, there is mixed evidence that higher levels of explicit or implicit racial prejudice are associated with a larger CRE. This literature is especially scarce for Asian and White faces (cf. Anzures et al., 2022; Ferguson et al., 2001). Thus, we investigated the role of self-reported, consciously held biases towards Asian and White individuals. In our racial biases/attitudes scale, we included biases that may be stereotypical of that race as well as potential racial attitudes towards particular races. We measured self-reported biases because we hope to develop (and in future work validate) a self-report inventory and because self-reported biases are easier to obtain than other measures (e.g., the IAT). Because these items were mixed with all other items in the CRE- I, we hoped they would be less salient than in prior work and, therefore, would effectively measure bias. Moreover, existing racial bias and attitude scales concerning the CRE may be outdated and have often been developed *ad-hoc* within empirical studies. The items created for this scale offer a psychometrically sound and recent measure of racial attitudes.

**Cognitive Disregard**

The social-cognitive hypothesis, specifically the cognitive-disregard model, posits differences in the selective allocation of attention as the cause of the CRE (Rodin, 1987). Individuals are expected to disregard faces that are unimportant to their social purposes, meaning those faces are processed just enough for categorization (e.g., Devine, 1989; MacLin & Malpass, 2001, 2003). Specifically, they are processed for general characteristics (e.g., race) rather than for facial characteristics (e.g., blue eyes) unique to them as individuals. Consistent with this view, studies have found recognition of unique facial characteristics (Chance & Goldstein, 1996) as well as longer and more frequent looking at same-race compared to cross-race faces (Goldinger et al., 2009). Furthermore, priming monoracial bicultural Latino-Americans with a specific social identity, such as Latino or American, altered the CRE for Latino and White faces (Marsh et al., 2016). The authors argued that priming influences the significance of target faces and, therefore, which faces receive more/less cognitive regard. By incorporating a cognitive disregard scale in the CRE-I, for the first time, we examined whether self-reported cognitive disregard predicts the CRE.

**Motivated individuation**

According to the CIM, a lack of motivated individuation of cross-race faces, along with a lack of expertise with them, causes the CRE (Hugenberg et al., 2007). Like the cognitive disregard model, the CIM posits that social categories signal perceivers whether faces should be individuated or processed superficially (Young et al., 2012). This signaling can be driven by subjectively important ingroup/outgroup distinctions, such as race and other cues to social significance (Bernstein et al., 2007; Rule et al., 2007; Shriver et al., 2008).

In addition to the cognitive disregard proposition, the motivated individuation argument suggests that cognitive disregard for other races can be mitigated by enhancing motivation to individuate them. Hugenberg et al. (2007) found that warning participants about the CRE and instructing them to attend to differences among faces of racial outgroup members eliminated the CRE. Similarly, recognition of cross-race targets improved when they were identified as having economically high-status roles (cf. low-status roles; Shriver & Hugenberg, 2010) and when participants had to depend on another person to achieve a goal, which is known to increase individuation without affecting group membership (i.e., outcome dependency; Baldwin et al., 2013). However, Tracy et al. (2023) recently tested the relationship between motivation and CRE. To measure motivation, they asked participants questions about the perceived importance of recognizing same- and cross-race faces, anticipated contact with same- and cross-race faces, and willingness to develop cross-race friendships. Only anticipated contact with other-race individuals predicted the CRE. Hence, the connection between motivation and the CRE is complex and lacks a clear conclusion.

Regardless of these mixed results, the CRE may, in part, reflect a lack of motivation to process and recall the faces of individuals from different races (Hugenberg et al., 2010). However, there is currently no reliable scale for measuring motivated individuation. Therefore, we included items aimed at measuring motivated individuation in the CRE-I.

**Overview of Current Study and Hypotheses**

We reasoned that the CRE may reflect specific dimensions: general face recognition ability, race-specific face recognition ability, quantity of interracial contact, quality of interracial contact, racial biases/attitudes, cognitive disregard, and motivation to individuate. We presented items pertinent to these dimensions to participants. To validate our anticipated scales, we included items from existing scales modified for White and Asian participants (cf. White and Black participants): the Positive and Negative Contact Scale (Hayward et al., 2017) and The Racial Attitudes Scale (Brigham, 1993). Additionally, participants completed a reliable and valid general face recognition scale, the Stirling Face Recognition Ability Scale (Bobak et al., 2017) for comparison with our general face recognition ability scale.

We used a classic lineup paradigm to assess the CRE. Participants completed trials in which they watched a mock crime video (with Asian or White targets) and, after a distractor task, made a lineup decision and rated their confidence in that lineup decision. With this experimental design, we also aimed to address gaps in the literature. First, we fully crossed participant and target race. This is optimal for comprehensively understanding the underlying mechanisms of the CRE among different race groups (Lee & Penrod, 2022) but has rarely occurred in lineup studies with Asian and White participants. Second, we used sequential lineups. The CRE literature has primarily utilized the old-new paradigm (cf. lineups; Lee & Penrod, 2022). Although some studies have used lineups, particularly simultaneous lineups (cf. Wright et al., 2001; Wylie et al., 2015), there has been limited research in this area (e.g., Meissner et al., 2005; Steblay et al., 2011). Moreover, sequential lineups are common in some regions, including the UK, parts of the US, and Canada (Fitzgerald et al., 2021). Given that decision-making is affected by lineup procedure, further research in this area is needed, particularly with respect to the use of sequential lineups. Third, most studies on the CRE have used Black and White faces, with limited attention given to other races, such as Asian faces (cf. Estudillo, 2021; Gross, 2009; Smith et al., 2001), which we examined.

We tested three hypotheses. First, we predicted that both races of participants (Whites, Asians) would be more accurate (i.e., make more correct identifications and rejections) with targets of their own compared to different races (i.e., a CRE). Second, we predicted that the developed scales (i.e., the CRE-I) would correlate positively with extant (some modified) scales measuring similar dimensions. Third, we predicted that our scales would postdict eyewitness accuracy, particularly cross-race lineup decisions. Finally, we explored whether the predictive CRE-I scales of identification accuracy explained variance in identification accuracy beyond what was explained by confidence alone.

**Method**

# We preregistered ([osf.io/j9nzc](https://osf.io/j9nzc/?view_only=99efe7be897343b1a4587207166c07f6)) our hypotheses, method, sample sizes, exclusion criteria, and analytic plan on the Open Science Framework (OSF), where we also share our data and analysis scripts: [osf.io/nx5mf/](https://osf.io/nx5mf/?view_only=fefb92953d024351bbf81b65b9e6ceb9). We obtained ethical approval for this study from Queen Margaret University Research Ethics Committee **Division of Psychology, Sociology and Education and Edinburgh Napier University** School of Applied Sciences Research Integrity Committee Chair**.**

**Participants**

Our final sample comprised participants self-identifying as Asian (*n* = 203) and White (*n* = 202) recruited via Mechanical Turk/CloudResearch (MT/CR) and university participant pools. Participants recruited from university participant pools received course credit for participating, while those recruited via MT/CR (CloudResearch Approved sample for White participants and demographically constrained CloudResearch sample for Asian participants) received monetary compensation of $2.50. We recruited most Asian participants via MT/CR (*n* = 188), whereas we recruited approximately 6% of the White participants via MT/CR (*n* = 28).

***Sample Size Determination***

We estimated the required sample size using a simulation conducted in R (R Studio Team, 2022). The simulation considered sample sizes from 150-1000, in 50 participant increments and reported how often a significant CRE (i.e., *p* < .05) would occur in a multilevel logistic regression with target race (same-race, cross-race) as a predictor of correct identifications and estimating a random effect for participant. The simulation assumed an odds ratio of 1.4 for correct identifications, based on Meissner and Brigham (2001). The simulation produced 100 samples for each sample size. It indicated that 200 participants provided more than 80% power to detect the effect. Because this study assessed the CRE for both Asian and White participants, we needed 200 Asian and 200 White participants.

***Data Inclusion and Exclusion***

Individuals could participate in this online study provided they were 18 or older, using a laptop or a desktop computer, and spoke and understood English. For data collection from the university participant pools, 17-year-old participants were allowed to ensure inclusivity per the National policy of persons 16 or older being considered adults for such a purpose. Participants were required to complete the study on a laptop or desktop computer (cf. mobile device) to ensure they had a clear enough view of the targets to allow participants to identify them from a lineup. Only participants who self-identified as Asian or White were recruited via MT/CR. Participants of any ethnicity were recruited via university participant pools. We excluded participants from the final analysis if they did not self-identify as Asian or White.

Initially, 461 participants consented and completed the survey in full. We excluded 56 participants from the analyses. These included 9 who did not consent for their data to be included, 14 who failed all three attention checks (participants had to get at least one correct in order to be included), 17 who had technical problems that affected their view of the mock-crime video or lineups (e.g., the video/image did not appear), and 1 who stated they cheated by taking a picture of the target. Because we were interested in participants who self-identified as Asian or White, we also excluded 6 participants who self-identified as Black, 2 participants self-identified as Hispanic/Latinx, and 6 participants who preferred not to answer the race question. Although we did not pre-register to do so, we excluded one participant because they self-reported as having “beginner” English level. We reasoned that beginner-level English would mean this individual’s responses to this inventory would not be valid because it is plausible that they would not have appreciated nuanced differences between questions intended to capture similar but different dimensions.

***Participant Demographics***

Asian participants ranged in age from 18 to 69 years (*M* = 34, *SD* = 9.69; 4 preferred not to answer). Participants self-identified as male (*n* = 100), female (*n* = 99), non-binary/third gender (*n* =2) or preferred not to answer (*n* = 2). Participants self-reported their English level as native English speakers (*n* =147), fluent (*n* = 52), advanced (*n* = 3), or intermediate (*n* = 1). Furthermore, 110 of the Asian participants reported that they were originally from the USA, while 62 reported being from a range of other countries[[2]](#footnote-2) (32 preferred not to answer). Similarly, 187 responded that they currently resided in the USA, with 2 in Pakistan and 1 in Kazakhstan (13 preferred not to answer). Finally, 149 reported living longest in the USA, and 34 in other countries[[3]](#footnote-3) (21 preferred not to answer).

White participants ranged in age from 17 to 78 (*M* = 24.02, *SD* = 10.58; 3 preferred not to answer). Thirty-seven self-identified as male, 162 as female, and 3 as non-binary/third gender. Most participants had a native level of English (*n* = 165), 35 were fluent, and 2 were at an advanced level. Furthermore, 141 of the participants reported that they were originally from the UK, 31 from the USA, while 22 reported being from a range of other countries[[4]](#footnote-4) (8 preferred not to answer). Similarly, 165 indicated they currently resided in the UK, 29 in the USA, and 1 in Egypt (8 preferred not to answer). Finally, 149 reported living longest in the UK, 28 in the USA, and 16 in other countries[[5]](#footnote-5) (7 preferred not to answer).

**Design**

The study was a 2 (Target Presence: target-present, target-absent; between-subjects) x 2 (Target Race; Asian, White; within-subjects) x 2 (Participant Race: Asian, White; between-subjects) x 2 (Target: a, b; within-subjects) mixed design. Target Presence refers to whether a lineup shown to participants contained the target (i.e., the person shown in a watched mock-crime video). Participant Race refers to whether a participant was Asian or White. Target Race refers to whether a mock-crime video depicted an Asian or White actor. Finally, Target refers to the fact that the mock-crime videos featured different actors to ensure the findings were not unique to individual actors (Wells & Windschitl, 1999). The order of the video/lineup trials was randomized and participants were randomly assigned to conditions. Each participant completed four trials: they saw two Asian and two White targets and for each target, they received a target-present or target-absent lineup (counterbalanced).

**Materials**

***Mock-crime Videos***

Four mock-crime videos showing an Asian or White male stealing money from a purse were used. All actors were in their early 20s, wearing their own clothes. These videos came from a set of videos and lineups developed for research, though the specific videos used varied (e.g., Beaudry et al., 2013; Mansour et al., 2009). Each video was approximately 30 seconds long. The display of the videos was clear and of good quality for viewing. Prior to seeing any videos, participants were informed that they would watch four videos with audio and answer questions about them.

***Intervening Task***

Participants completed a 30-second intervening task between each mock-crime video and lineup to disrupt working memory (i.e., the temporary retention of visual stimuli before further processing). The intervening task involved viewing a Where’s Waldo[[6]](#footnote-6) image that depicts a beach. Participants were instructed to answer as many questions as they could about this image during the allotted time. The questions were designed to be impossible to complete in the allotted time. For example, the questions included: “*How many open umbrellas are there?”* and *“Does the man in the red-and-white bathing suit have a moustache?”.*

***Lineups***

We used six-person target-present and target-absent photographic sequential lineups. White lineup stimuli were from the same set as those used by Mansour et al. (2009), though the specific lineups used varied. Asian lineup stimuli were developed using the match-to-description technique (Koehnken, Malpass, & Wogalter, 1996; Luus & Wells, 1991; Navon, 1992) with the help of ten self-identifying Asian volunteers. For detailed information about the construction of our Asian lineups, please see the supplementary materials.

For each target, the target-present lineup contained the target and five fillers, while the target-absent lineup contained six fillers. No fillers were repeated across lineups. The positions of the target and fillers were randomized across participants. There were no designated innocent suspects. Lineup members were presented sequentially in two laps. Participants were asked to make a lineup decision after seeing both laps. The order lineup members shown on each lap was the same for a participant, but the position of the suspect and fillers was randomised across participants. Before each lineup, participants read the instructions:

The video you saw showed a mock-crime. We would now like to show you a lineup (an array of photographs of faces) for the "criminal" you saw in the video. Please view each member carefully. Please note that the "criminal" you saw may or may not be in the lineup. You should remember that it is just as important to clear innocent persons from suspicion as to identify the guilty. After you have seen the lineup, you will be asked for the number of the "criminal" or to indicate that they were not present. If you see the "criminal", please make sure you keep the number of that photo in your mind.

The resultant Tredoux’s *E*s (Tredoux, 1998) for the White target-absent lineups were 3.48 95% CI [2.44 to 6.02] and 4.11 95% CI [3.26 to 5.56] based on the White participants’ decisions. For the Asian target-absent lineups, the Tredoux’s *E*s were 1.96 95% CI [1.54 to 2.69] and 2.49 95% CI [1.95 to 3.45] based on the Asian participants’ decisions. Higher values of Tredoux’s *E* (closer to nominal size) indicate greater fairness.

**Measures**

***Lineup Responses***

We coded target-present lineup decisions as suspect identifications (correct identifications), filler identifications, or incorrect rejections. We coded target-absent lineup decisions as filler identifications or correct rejections. Lineup accuracy was operationalized in three ways: identification accuracy (correct identifications versus all other identifications), target-present lineup accuracy (correct identifications versus all other decisions from target-present lineups), and target-absent lineup accuracy (correct rejections versus filler identifications from target-absent lineups).

***Eyewitness Confidence***

Following the lineup decision, we asked the participants to indicate their confidence in their decision on a 0-100% (0 = *Not at all confident*, 100 = *Extremely confident*) scale.

***The CRE Inventory (CRE-I)***

The CRE Inventory (CRE-I) is intended to predict the extent to which people are likely to show a CRE based on theoretically suggested and empirically established individual differences dimensions of the CRE. We utilized 20 potential scales (including the amended extant scales) to measure seven dimensions with items rated on a 6-point Likert-type scale (1 = *Strongly disagree*, 6 = *Strongly agree*) and with the option to respond, *“I do not know*.*”* We followed DeVellis’ (2017) guidelines for the initial development of a psychometric scale. All items were reviewed by two CRE experts for relevance, representativeness, specificity, clarity, and concision. We adjusted the items based on the experts’ feedback before presenting them to participants.

With two exceptions, each anticipated scale had two potential item sets: one worded for White and one worded for Asian faces. The potential items for the general face recognition ability scale were not race-specific, so only one version was used. There were three versions of the motivation to individuate scale as we created items asking about motivation to individuate that were race-specific but also that were race-neutral.

Thus, we produced items for 14 potential scales for the CRE-I. The potential scales were general face perception ability (19 items, e.g., *I notice when someone has distinctive eyes*), race-specific face perception ability (20 items on each race-specific scale, e.g., *I can see the differences between two different White/Asian people*), quantity of interracial contact (21 items on each race-specific scale, e.g., *There are White/Asian people I see frequently*), quality of interracial contact (42 items on the White scale, 41 items on the Asian scale; e.g., *I enjoy the company of my White/Asian friends*), racial attitudes/biases (40 items on the White scale, 35 items on the Asian scale; e.g., *White/Asian people are friendly*), cognitive disregard (8 items one each race-specific scale and 1 race-neutral scale item,[[7]](#footnote-7) e.g., *I see White/Asian faces a lot in my daily life*), and motivated individuation (17 items each race-specific scale and 6 race-neutral scale items, e.g., *It is important for me to distinguish between faces; I deliberately pay close attention to tell the difference between White/Asian faces*). Differences in the number of potential items for each racial group reflect the challenges of identifying items that reflect specific racial groups (e.g., racial biases towards Asian people and White people differ in both context and quantity). The total number of items across these dimensions was 316.

As noted above, we tailored scales developed for White and Black faces for this research. These *existing amended scales* were the Positive and Negative Contact scale (Hayward et al., 2017) and The Racial Attitudes scale (Brigham, 1993), the former of which comprises two subscales. We produced separate versions for White and Asian faces. Thus, the existing amended scales were the Positive Contact scale (14 items for each race-specific scale, e.g., *I feel free to express myself when I encounter White/Asian people.*), the Negative Contact scale (32 items for each race-specific scale, e.g., *White/Asian people have been condescending towards me.*), and the Racial Attitudes scale (6 race-neutral scale items,[[8]](#footnote-8) 8 items for each race, scale, e.g., *I would rather not have White/ Asian people live in the same apartment building I live in.*). The total number of items was 114.

We also gave the participants the 20-item Stirling Face Recognition Scale (SFRS; Bobak, Mileva & Hancock, 2017; e.g., *I never forget a face*) to assess individual differences in face recognition ability, explore whether the SFRS predicts the CRE, and to determine how the SFRS correlated with the general face perception ability scale we developed.

The 450 items for the potential scales, existing amended scales, and the SFRS were presented in random order. This was done to minimize the potential influence of cognitive fatigue and demand characteristics on participants' responses to specific items. Scale scores were calculated by averaging the responses to the items in the respective scales. *“I do not know”* responses were not included in the averages.

***Demographics and Data Quality Checks***

 Participants were asked about the device they used to complete the study (e.g., mobile, laptop, iPad, desktop computer) and were unable to start the study if they reported using a device other than a laptop or a desktop computer. We asked three attention-check questions about the mock-crime videos (e.g., “*What was the color of the purse from which the "criminal" stole the money?”*). We used the answers to the attention check questions to determine whether participants paid sufficient attention to the mock-crime video to be included in the analyses. We also asked participants six demographic questions (age, sex, English level, country of origin, country of residence, and country they lived longest in; they had an option of not responding).

We asked four multiple-choice questions to assess data quality. We asked participants whether they experienced technical difficulties (*Did you experience any technical difficulties with viewing the images [e.g., could not hear or see one or both]?*). If they responded yes, they were asked to expand on their answer so that we could exclude the data of participants whose technical difficulties meant their data was invalid. We also asked participants whether they cheated on the task in any way (*When doing this study, did you cheat in any way?*)*.* If they responded in the affirmative, participants were encouraged to give further details, and we used the details to determine whether to exclude them (e.g., if they took a screenshot of the video, which they looked at while making the lineup decision). Furthermore, we asked participants whether they hadcompleted an eyewitness study before (*“Have you ever done a study on eyewitness memory before?”*) and whether they had seen similar mock-crime videos before (*“Have you ever seen a mock-crime video like the one we showed you before?”*). We did not use answers to these questions as exclusion criteria but rather to inform the lab about the naivety of our samples. For all questions described in this paragraph, we assured participants that their responses would not influence whether they received credit/payment.  

**Procedure**

Participants completed four lineup paradigm trials. On each trial, they watched a mock-crime video, worked on the Where’s Waldo task, made their lineup decision, and indicated their confidence in their lineup decision. Next, participants completed the potential CRE-I items, existing amended scale items, and SFRS items. Following this, participants answered the demographic questions, attention check questions, and data quality questions. Finally, participants were debriefed and reimbursed/credited, as appropriate.

**Results**

We conducted all data coding and analyses in R (R Studio Team, 2020). Beyond the base built-in R functions, we used the *Tidyverse* *repository* (Wickham et al., 2019) and functions within the following packages: *ggplot (*Wickham et al., 2019*), data.table* (Dowle & Srinivasan, 2023), *magrittr* (Bache & Wickham,2020), *lme4* (Bates, 2015), *boot* (Canty & Ripley, 2022), *janitor* (Firke, 2023), *sjmisc* (Lüdecke, 2021), and *psych* (Revelle, 2021). For all analyses, we regarded *p*-values less than .050 as statistically significant.

**Lineup Decisions**

We predicted that White participants would be more accurate at identifying and rejecting White faces compared to Asian faces in a lineup, while Asian participants would be more accurate at identifying and rejecting Asian faces compared to White faces in a lineup. To test this, we employed multilevel binomial logistic regression with participant as a random effect and Target Race as a fixed effect to predict lineup decision accuracy (separate models for identification accuracy, target-present lineup accuracy, and target-absent lineup accuracy). We conducted separate analyses for White and Asian participants.

Visualizations of lineup decisions, resultant lineup fairness for target-present and target-absent lineups, and lineup member selections (Tables 1 to 3 in supplementary materials) showed that both White and Asian participants chose Filler 1 in the Asian Lineup 1 target-absent lineup more so than any other filler, skewing results towards that filler. This was not the case for target-present lineups, indicating that the lineup was not biased in a way that caused the target to stand out; however, because these results are atypical, lineups for that Asian target were removed from subsequent analyses. We present analyses using the full set of lineups (per our preregistration) in the supplementary materials.

For White participants, there was a CRE for identification accuracy (correct vs. incorrect identifications), *z* = 2.21, *p* = .02, and target-absent accuracy, *z* = 3.07, *p* = .002, but not for target-present accuracy, *z* = 0.14, *p* = .88. For Asian participants, there was no CRE for identification accuracy, *z* = 1.12, *p* = .26, or target-present accuracy, *z* = 1.34, *p* = .17, but there was an inverted CRE for target-absent accuracy, *z* = 2.41, *p* = .01. Asian participants made more correct rejections of White lineups than Asian lineups. We attribute this finding to the low resultant lineup size of the Asian lineup (Tredoux’s *E* = 2.49).

**Table 1**

*Lineup Decisions (percentages)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Target-present lineups | | | Target-absent lineups | | |
|  | Correct Identifications | False Identifications | Incorrect rejections | | Correct Rejections | False Identifications |
| White participants, White targets | 86.63 | 7.92 | 5.44 | | 59.20 | 40.79 |
| White participants, Asian targets | 86.45 | 10.41 | 3.12 | | 40.56 | 59.49 |
| Asian participants, White targets | 82.75 | 9.65 | 7.58 | | 59.35 | 40.64 |
| Asian participants, Asian targets | 89.55 | 4.47 | 5.97 | | 43.37 | 56.62 |

Next, we used multilevel probit logistic regression to predict lineup decision accuracy with Target Presence, Target Race, and their interaction as fixed effects and participant as a random effect. Probit regression can be used to test whether a variable (in this case, Target Race) influences discriminability (i.e., the ability to differentiate guilty from innocent suspects) and response criterion (i.e., the evidence threshold that determines whether an eyewitness identifies someone; DeCarlo, 1998; Wright & London, 2009). Correct identifications and filler identifications from target-absent lineups were coded as 1, and all other lineup decisions were coded as 0. Separate analyses were conducted for White and Asian participants. For White participants, the interaction was significant, indicating that discriminability was lower for Asian (*d’*= 0.88) than White targets (*d’*= 1.39), *z* = 1.97, *p* = .04. For Asian participants, discriminability was similar for Asian (*d*’= 1. 10) and White targets (*d*’= 1. 19), *z* = 0.30, *p* = .75. In order to estimate discriminability in a forensically relevant way, we also computed discriminability using the average target-absent filler identification rate (i.e., by dividing the total target-absent filler identification rate by the resultant Tredoux’s *E;* Quigley-McBride & Wells, 2021). We calculated discriminability for each target based on the lineup decisions for those targets and used Tredoux’s *E* for that target’s target-absent lineup from the respective Participant Race. When discriminability for White targets was computed for each target separately, White participants (*d’* =1.88; *d’* =3.03) and Asian participants (*d’* =1.57; *d’* =2.76) had similar discriminability across the targets. When these discriminability scores are averaged, White participants had higher discriminabilityfor White targets (*d’* =2.45) than the Asian target (*d’* = 1.79). Similarly, Asian participants had higher discriminability for the White targets (*d’* =2.16) than the Asian target (*d’* =1.98).

Finally, running the multilevel probit logistic regression models without the interaction term tested for a difference in response criterion. The response criterion for Asian targets (*c* = - 0.12) was more liberal than for White targets (*c* = 0.17), *z* = 9.54, *p* < .001. Likewise, for Asian participants, the response criterion for Asian targets (*c* = - 0.14) was more liberal than for White targets (*c* = 0.22), *z* = 7.68, *p* < .001.

In summary, there was a CRE on identification accuracy, target-absent accuracy, and discriminability for White participants but not on target-present accuracy. There was no CRE on identification accuracy, target-present accuracy, or discriminability for Asian participants but there was an inverted CRE on target-absent accuracy.

**Creation of the Cross-Race Effect Inventory**

Our primary objective was to use internally reliable and valid scales to improve predictions of the likelihood that an eyewitness is accurate. Since the Asian participants did not exhibit the expected CRE, we focused on developing the CRE-I for White eyewitnesses to White and Asian targets. Following DeVellis (2017), we conducted an exploratory factor analysis (EFA) on the presented items for each of the 20 scales[[9]](#footnote-9). Once the final factor structure for each scale was established, we calculated Cronbach’s alpha (*α*) for each scale to assess reliability (> .7 is considered acceptable). We assessed the convergent validity of each factor on each scale comprising the CRE-I in relation to the existing-amended scales. Furthermore, we examined the predictive abilities of the factors on the CRE-I.

***EFA***

We first checked if our data were appropriate for factor analysis. For each scale, we conducted Bartlett's test of sphericity (to determine whether there are significant correlations between items) and the KMO Test (a measure of the proportion of variance in the items that can be explained by underlying factors) and checked the determinant of the correlation matrix (a measure of how much variance is shared amongst items). Passing all three tests indicates that factor analysis is appropriate (DeVellis, 2017). Next, we checked the correlation matrix for each scale’s items for high factor loadings (~ .9), but none were found so no items were excluded on this basis. We also excluded the items that correlated at a low rate or not at all with other items (< .3). The first step of the factor analysis is to determine how many factors underly the items. For each scale, we inspected the scree plot and used Kaiser’s criterion (i.e., >1, DeVellis, 2017) to determine the optimal number of factors and considered both structures when they suggested different numbers of factors.

We conducted an EFA using the maximum likelihood approach and direct oblimin rotation on each of the 20 scales comprising the initial 430 items. We aimed for an optimal balance of brevity and reliability; therefore, for each scale, .60 was the minimum loading for retaining an item. If there was an inadequate number of items loaded on the factors, we lowered this cut-off by .05 until either there were more than two items loading on each factor or until a cut-off of .32 was reached (Tabachnik & Fidel, 2013 suggest this value as a minimum cut-off). Therefore, the decision to retain or exclude items based on factor loadings was based on a combination of statistical and substantive reasoning (see Flowchart 1 in the supplementary materials for more detail).

For each scale, we iteratively dropped items with low loadings or considerable cross-loadings (Samuels, 2016) and re-ran the factor analysis until a clear structure emerged. Table 1 (for developed scales) and 2 (for existing-amended scales) in the Appendix report the factor loading cut-off used, the number of items retained for each factor, the range of item loadings, the variance accounted for by each factor, Cronbach’s alpha, and item reliability. In summary, we derived a one or two-factor structure for all scales. We named the factors of two-factor structure scales after the content of the items that constituted those factors. For instance, the Motivation to Individuate-Asian scale and the Motivation to Individuate-White scale resulted in a two-factor structure and, for both scales, the factors were labelled *Attention* (motivation to pay attention) and *Disinterest* (lacking motivation to pay attention).

Streiner (1994) suggested that factors explaining 50% or more variance are sufficiently valid. Of our 20 scales, eight met this criterion: General Face Recognition Ability, Race-Specific Face Recognition Ability-Asian, Racial Attitudes-Asian, Quantity of Interracial Contact-Asian, Quality of Interracial Contact-Asian, as well as the existing amended scales Negative Racial Contact (White, Asian) and Racial Attitudes-Asian. Thus, these eight scales had adequate factor loadings for their items, which indicates a strong association between the items and factors. Streiner’s suggestion is only a rule of thumb, therefore, we note that five scales just missed the 50% criterion (i.e., explained 46-49 % variance): Race-Specific Face Recognition Ability-White, Racial Attitudes- White, Quantity of Interracial Contact-White, Quality of Interracial Contact-White, as well as the existing amended scales Positive Contact-Asian and Racial Attitudes-White.

Most of the scales had good internal reliability (*α*s > .7, *Range* = .70 to .97) except the Cognitive Disregard scales (*α* = .51 & *α* = .55 for the White and Asian scales respectively), the Motivated Individuation-Neutral scale (*α* = .58), and the Motivated Individuation-White scale (*α* = .60). Notably, the Quantity of Contact-Asian, Quality of Contact-Asian scale, Racial Attitudes-Asian scale, and Negative Contact-Asian scales had excellent reliability (*α*s > .9).

The final CRE-I consists of 14 scales and 100 items,[[10]](#footnote-10) to be used with White eyewitnesses. Table 2 below (developed scales) and Table 6 in the Appendix (existing amended scales) contain the final version and descriptive statistics for the items. The scales could be used individually, together, or together with the refined existing-amended scales.

| **Table 2**  *Means, Standard Deviations, and Internal Consistency Estimates for Factors and Items in the Final CRE-I* | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Factors & Items | | *M* (*SD*) | | Item-Total Correlation | | *ɑ* if item deleted | |
| General Face Recognition Ability Scale (*ɑ* = .84 95% CI [.81, .88]) | | 3.82 (1.11) | |  | |  | |
|  | I generally forget faces. (R) | | 3.73 (1.44) | | .76 | | .70 | |
|  | I don’t trust my ability to remember a face.(R) | | 3.72 (1.33) | | .71 | | .66 | |
|  | I often meet people who greet me, but I fail to recognize their faces.(R) | | 3.75 (1.41) | | .74 | | .68 | |
|  | I can rarely remember a face.( R) | | 4.12 (1.32) | | .71 | | .65 | |
| Race-Specific Face Recognition Ability-Asian Scale (*ɑ* = .71 95% CI [.65, .77]) | |  | |  | |  | |
| Race-Specific Face Recognition Ability-Asian, *Difficulty* (*ɑ* = .76 95% CI [.70, .82]) | | 4.54 (1.03) | |  | |  | |
|  | All Asian people look alike. (R) | | 4.62 (1.33) | | .79 | | .67 | |
|  | All Asian people have the same eye shape. (R) | | 4.61 (1.27) | | .78 | | .66 | |
|  | When I meet an Asian person, they often remind me of someone. (R) | | 4.26 (1.19) | | .49 | | .44 | |
| Race-Specific Face Recognition Ability-Asian, *Noticing* (*ɑ* = .75 95% CI [.69, .81]) | | 4.18 (0.88) | |  | |  | |
|  | I am good at recognizing Asian faces. | | 3.81 (1.15) | | .65 | | .56 | |
|  | I can see the differences between two different Asian people. | | 4.38 (1.01) | | .73 | | .63 | |
|  | I can see the differences between two similar-looking Asian people. | | 4.26 (1.05) | | .62 | | .53 | |
|  | Race-Specific Face Recognition Ability-White Scale (*ɑ* = .70 95% CI [.64, .77]) | |  | |  | |  | |
| Race-Specific Face Recognition Ability-White, *Difficulty* (*ɑ* = .77 95% CI [.71, .82]) | | 4.64 (0.91) | |  | |  | |
|  | All White people look alike. (R) | | 4.53 (1.16) | | .67 | | .58 | |
|  | All White people have the same eye shape. (R) | | 4.72 (1.08) | | .65 | | .57 | |
|  | After meeting a White person, the only thing I remember is their eyes. (R) | | 4.65 (1.08) | | .73 | | .64 | |
|  | Race-Specific Face Recognition Ability-White, *Noticing* (*ɑ* = .69 95% CI [.62, .76]) | | 4.46 (0.86) | |  | |  | |
|  | I am good at recognizing White faces. | | 4.37 (1.06) | | .68 | | .57 | |
|  | I could describe the facial features of a White face I have seen. | | 4.35 (1.12) | | .65 | | .54 | |
|  | I can see the differences between two different White people. | | 4.71 (0.90) | | .49 | | .42 | |
| Racial Attitudes-Asian Scale (*ɑ* = .88 95% CI [.85, .90]) | | 4.72 (0.92) | |  | |  | |
| Asian people speak English poorly. (R) | | 4.80 (1.17) | | .80 | | .75 | |
| Asian people are nerds. (R) | | 4.63 (1.19) | | .72 | | .67 | |
| Asian people are bad drivers. (R) | | 4.78 (1.16) | | .81 | | .74 | |
| Asian languages usually sound like gibberish. (R) | | 4.43 (1.32) | | .60 | | .56 | |
| Asian people are not very social. (R) | | 4.45 (1.25) | | .74 | | .69 | |
| I find Asian people boring. (R) | | 4.93 (1.09) | | .73 | | .68 | |
| Racial Attitudes-White Scale (*ɑ* = .88 95% CI [.85, .90]) | | 4.25 (0.90) | |  | |  | |
|  | White people are bad drivers. (R) | | 4.08 (1.28) | | .72 | | .67 | |
|  | I find White people boring. (R) | | 4.28 (1.25) | | .64 | | .60 | |
|  | White people are irresponsible. (R) | | 3.86 (1.23) | | .78 | | .69 | |
|  | White parents don’t discipline their children well. (R) | | 3.89 (1.23) | | .65 | | .61 | |
|  | White people have a weak work ethic. (R) | | 4.27 (1.24) | | .63 | | .59 | |
|  | White people can’t dance. (R) | | 4.22 (1.19) | | .65 | | .59 | |
|  | White people can’t cook. (R) | | 4.60 (1.21) | | .73 | | .67 | |
|  | White people don’t work hard on their relationships. (R) | | 4.54 (1.22) | | .74 | | .68 | |
| Quantity of Contact-Asian Scale (*ɑ* = .90 95% CI [.87, .92]) | |  | |  | |  | |
| Quantity of Contact-Asian, *Interaction* (*ɑ* = .89 95% CI [.86, .91]) | | 3.26 (1.12) | |  | |  | |
|  | In my life, I have rarely interacted with Asian people. (R) | | 3.25 (1.39) | | .59 | | .55 | |
|  | I live or have lived in an area where I interact with Asian people. | | 3.52 (1.58) | | .68 | | .65 | |
|  | I interact/interacted with Asian people in school. | | 4.03 (1.42) | | .61 | | .58 | |
|  | I know lots of Asian people. | | 2.61 (1.34) | | .81 | | .76 | |
|  | There are Asian people I see frequently. | | 3.32 (1.46) | | .75 | | .71 | |
|  | I see many Asian people during my daily life (e.g., at school, sports, activities, shops, living area, etc.) | | 3.30 (1.46) | | .77 | | .71 | |
|  | I encounter Asian people frequently. | | 3.27 (1.37) | | .70 | | .65 | |
|  | I have more than one Asian friend. | | 3.04 (1.66) | | .70 | | .65 | |
| Quantity of Contact-Asian, *Presence* (*ɑ* = .84 95% CI [.80, .87]) | | 2.05 (1.04) | |  | |  | |
|  | I know fewer White people than Asian people. | | 1.90 (1.16) | | .81 | | .73 | |
|  | I interact with more Asian than White people during my daily life. | | 2.15 (1.33) | | .75 | | .68 | |
|  | Most of my friends are Asian. | | 2.06 (1.19) | | .81 | | .72 | |
|  | I see more Asian than White faces during my daily life. | | 2.12 (1.41) | | .59 | | .55 | |
|  |  | |  | |  | |  | |
|  |  | |  | |  | |  | |
| Factors & Items | | *M* (*SD*) | | Item-Total Correlation | | *ɑ* if item deleted | |
| Quantity of Contact-White Scale (*ɑ* = .88 95% CI [.86, .90]) | | 5.24 (0.68) | |  | |  | |
| I know lots of White people. | | 5.33 (0.83) | | .62 | | .57 | |
| There are White people I see frequently. | | 5.24 (0.86) | | .73 | | .68 | |
| I encounter White people frequently. | | 5.37 (0.78) | | .70 | | .66 | |
| I have worked with White people. | | 5.18 (0.96) | | .65 | | .59 | |
| I have a lot of White friends. | | 5.06 (1.05) | | .64 | | .60 | |
| I live or have lived in an area where I interact with White people. | | 5.30 (0.90) | | .66 | | .62 | |
| I interact/interacted with White people in school. | | 5.28 (1.02) | | .69 | | .64 | |
| Most of my friends are White. | | 5.08 (1.05) | | .65 | | .61 | |
| I see more White faces than Asian faces in my daily life. | | 5.29 (0.88) | | .70 | | .66 | |
| Quality of Contact-Asian Scale (*ɑ* = .90 95% CI [.88, .92]) | |  | |  | |  | |
| Quality of Contact-Asian Scale, *Value* (*ɑ* = .93 95% CI [.92, .95]) | | 2.68 (1.16) | |  | |  | |
| There are Asian people I know very well. | | 3.40 (1.68) | | .74 | | .72 | |
| I often spend time with Asian people. | | 2.73 (1.33) | | .83 | | .80 | |
| I spend a lot of my free time doing things with Asian people. | | 2.59 (1.36) | | .80 | | .78 | |
| I often visit Asian people. | | 2.61 (1.35) | | .80 | | .77 | |
| Asian people often visit me. | | 2.42 (1.32) | | .85 | | .81 | |
| I socialize a lot with Asian people. | | 2.76 (1.42) | | .83 | | .80 | |
| My closest friends are Asian. | | 2.26 (1.20) | | .79 | | .76 | |
| I have intimate friendships with Asian people. | | 2.91 (1.56) | | .75 | | .73 | |
| Quality of Contact-Asian Scale, *Reliability* (*ɑ* = .90 95% CI [.87, .92]) | | 4.51 (0.82) | |  | |  | |
| My Asian friends help me when I need it. | | 4.56 (1.00) | | .76 | | .71 | |
| My Asian friends are willing to give me advice. | | 4.48 (0.97) | | .76 | | .72 | |
| My Asian friends support my decisions. | | 4.53 (0.99) | | .82 | | .77 | |
| My Asian friends make me feel better when I am feeling sad. | | 4.36 (1.07) | | .77 | | .73 | |
| I enjoy the company of my Asian friends. | | 4.70 (1.02) | | .74 | | .70 | |
| I would like to spend more time with my Asian friends. | | 4.54 (1.02) | | .70 | | .66 | |
| Quality of Contact-White Scale (*ɑ* = .78 95% CI [.72, .82]) | |  | |  | |  | |
| Quality of Contact-White Scale, *Value* (*ɑ* = .77 95% CI [.70, .82]) | | 5.13 (0.75) | |  | |  | |
| I socialize a lot with White people. | | 5.11 (0.89) | | .76 | | .65 | |
| I interact with White people during recreational periods. | | 4.89 (0.97) | | .76 | | .66 | |
| There are White people I know very well. | | 5.36 (0.87) | | .54 | | .48 | |
| Quality of Contact-White Scale, *Reliability* (*ɑ* = .70 95% CI [.62, .76]) | | 4.66 (0.79) | |  | |  | |
| My White friends are there when I need them. | | 4.82 (0.94) | | .64 | | .53 | |
| My White friends help me with my studies. | | 4.55 (1.02) | | .55 | | .47 | |
| I like receiving guidance from White friends. | | 4.60 (0.94) | | .64 | | .53 | |
| Motivated individuation-General Scale (*ɑ* = .59 95% CI [.46, .67]) | | 4.61 (0.77) | |  | |  | |
| It is important to understand how people differ. | | 5.01 (0.92) | | .57 | | .42 | |
| It is important for me to distinguish between faces. | | 4.54 (0.97) | | .58 | | .43 | |
| I am motivated to recognize people of another race. | | 4.26 (1.16) | | .40 | | .31 | |
| Motivated individuation-Asian Scale (*ɑ* = .71 95% CI [.65, .77]) | |  | |  | |  | |
| Motivated individuation- Asian, *Attention* (*ɑ* = .70 95% CI [.63, .76]) | | 4.29 (0.89) | |  | |  | |
| It is not important for me to recognize an Asian person. (R) | | 4.37 (1.36) | | .69 | | .57 | |
| It is not important to distinguish between Asian faces. (R) | | 4.24 (1.50) | | .52 | | .44 | |
| I do not pay attention to the facial features of Asians. (R) | | 3.99 (1.33) | | .49 | | .40 | |
| I do not care about Asian people. (R) | | 5.22 (1.10) | | .50 | | .40 | |
| I do not try to differentiate the facial features of Chinese and Japanese people. (R) | | 3.69 (1.36) | | .42 | | .35 | |
| I am not motivated to study the facial features of Asians. (R) | | 3.74 (1.31) | | .46 | | .38 | |
| Motivated individuation- Asian, *Disinterest* (*ɑ* = .64 95% CI [.53, .71]) | | 4.10 (1.06) | |  | |  | |
| I would like to be good at distinguishing between Asian faces. | | 4.54 (1.09) | | .63 | | .49 | |
| It is important for me to distinguish between Asian faces. | | 4.23 (1.26) | | .63 | | .48 | |
| I deliberately pay close attention to telling the difference between Asian faces. | | 3.61 (1.33) | | .43 | | .35 | |
| Motivated individuation-White Scale (*ɑ* = .60 95% CI [.52, .69]) | |  | |  | |  | |
| Motivated individuation-White, *Attention* (*ɑ* = .63 95% CI [.55, .70]) | | 4.00 (0.86) | |  | |  | |
|  | I would like to be good at distinguishing between White faces. | | 4.51 (0.98) | | .41 | | .31 | |
|  | It is important for me to distinguish between White faces. | | 4.13 (1.16) | | .53 | | .41 | |
|  | It is important in my job to recognize different White faces. | | 3.71 (1.48) | | .48 | | .38 | |
|  | I deliberately pay close attention to telling the difference between White faces. | | 3.60 (1.28) | | .45 | | .36 | |
|  | It is important for me to name my White acquaintances. | | 4.24 (1.21) | | .51 | | .40 | |
|  | I care to know where a White person I meet is from. | | 4.06 (1.31) | | .37 | | .29 | |
| Motivated individuation-White, *Disinterest* (*ɑ* = .57 95% CI [.46, .66]) | | 4.16 (1.06) | |  | |  | |
|  | I do not care about White people. (R) | | 4.69 (1.30) | | .57 | | .43 | |
|  | I do not pay attention to the facial features of Whites. (R) | | 3.86 (1.35) | | .48 | | .36 | |
|  | It is not important for me to recognize a White person. (R) | | 3.97 (1.36) | | .46 | | .35 | |
| Factors & Items | | *M* (*SD*) | | Item-Total Correlation | | *ɑ* if item deleted | |
| Cognitive Disregard-Asian Scale (*ɑ* = .55 95% CI [.45, .65]) | | 4.08 (0.83) | |  | |  | |
|  | After meeting an Asian person, I forget where they are from. | | 4.15 (1.23) | | .57 | | .42 | |
|  | I never spot Asian faces when I walk on the street. | | 4.16 (1.17) | | .42 | | .31 | |
|  | If I saw an Asian person, I would not pay attention to their features. | | 3.70 (1.21) | | .38 | | .27 | |
|  | I find it challenging to understand Asian people. | | 4.21 (1.20) | | .44 | | .31 | |
|  | Cognitive Disregard-White Scale (*ɑ* = .52 95% CI [.40, .62]) | | 4.62 (0.69) | |  | |  | |
|  | I never spot White faces when I walk on the street. (R) | | 4.54 (1.02) | | .45 | | .32 | |
|  | I see White faces a lot in my daily life. | | 5.31 (0.82) | | .38 | | .26 | |
|  | It is easy for me to remember the facial features of a White person I have just met. | | 4.11 (1.23) | | .36 | | .27 | |
|  | I care about telling the difference between faces. | | 4.45 (1.12) | | .50 | | .36 | |
| *Note*. *ɑ* = Cronbach’s alpha. *M* = Mean scale scores. Higher values indicated greater agreement with the statement. R *=* Reverse coded items. | | | | | | | |

***Convergent Validity***

We investigated whether the CRE-I scales correlated with relevant existing amended scales. We considered *r* = .50 as the threshold for adequate convergent validity. For detailed inferential results, please see Table 3 in the Appendix. The General Face Recognition Ability scale and the SFRS (Bobak et al., 2017) were positively correlated*,* *r* = .78, *p* < .001. Most combinations of the developed contact scales and existing amended contact scales (Hayward et al., 2017) were positively and significantly correlated, .14 < *r* < .68, *p*s ≤ .04. Several combinations were not significant: for instance, the correlation between the Positive Contact-Asian scale's *Amicable Experiences* factor and the Quantity of Contact-Asian scale's *Presence* factor (*p* = .85). Four combinations were negatively and significantly correlated; however, these were all weak correlations (*r*s < -.24). For instance, the Positive Contact-Asian scale's *Supportive Bonds* factor and the Quantity of Contact-Asian scale's *Presence* factor, *r =* -.17, *p =* .01, were negatively correlated. Finally, the developed and existing amended Racial Attitudes scales (Brigham, 1993) were positively and significantly correlated, .49 < *r* < .67, all *p*s < .001.

**Predictive Role of Individual Differences in the CREI**

A key reason for developing the CRE-I was to help improve the criminal justice system’s ability to predict whether an identification is likely to be accurate. This research does not address the CRE-I's predictive validity as it is not ideal to test predictive validity using the same sample from which the scale was developed. However, the individual differences observed in responses to the CRE-I scales provide valuable theoretical insights.

For the following analyses, we produced composite scores for each factor of each scale in the CRE-I. All models included Target Race as it accounted for significant variance in identification accuracy and the CRE-I was developed to explain variance beyond Target Race (i.e., when Target Race varies). The key results are summarized below (see Table 4 in the Appendix reports for full inferential statistics).

Several scales predicted identification accuracy: General Face Recognition Ability, Race-Specific Face Recognition Ability-Asian (*Difficulty* factor), Race-Specific Face Recognition Ability-White (*Difficulty* factor), Racial Attitudes-White, Quantity of Contact-Asian scale (*Presence* factor), Motivated Individuation-White (*Disinterest* factor) and Cognitive Disregard-Asian (*p*s ≤ .03). Two of the six existing amended scales were predictive: Racial Attitudes-Asian and Racial Attitudes-White (*p*s ≤ .04). All relationships were positive, except that higher cognitive disregard was associated with lower accuracy. Finally, the SFRS (Bobak, Mileva & Hancock, 2017) was not predictive, *z* = 1.28, *p* = .19.

To test whether the predictive scales predicted overlapping variance, a stepwise binomial multilevel logistic regression was conducted including fixed effects of General Face Recognition Ability, Race-Specific Face Recognition Ability-Asian (*Difficulty* factor), Race-Specific Face Recognition Ability-White (*Difficulty* factor), Racial Attitudes-White, Quantity of Contact-Asian scale (*Presence* factor), Motivated Individuation-White (*Disinterest* factor) and Cognitive Disregard-Asian. The stepwise model started with all seven predictors and utilised backwards selection. The final model only contained the Racial Attitudes-White scale, *z* = 3.29, *p* < .001. This indicates that only the Racial Attitudes-White scale explained unique variance in identification accuracy regardless of whether the target was White or Asian. Researchers may choose to modify the scales (except for the Racial Attitudes scale which focuses on biases toward Asians) for different racial groups (e.g., changing "Asian" to "Black") to test the generalizability of the scales for predicting the CRE.

***Combining Predictors of Accuracy***

Confidence, a robust postdictor of accuracy under pristine conditions (e.g., Wixted & Wells, 2017), predicted identification accuracy for White participants, *z* = 8.14, *p* < .001[[11]](#footnote-11). Confidence had significant small correlations with various developed scales (0.11 < *r*s < 0.16; *p*s < .03)—Race-Specific Face Recognition Ability-Asian (both factors), Motivated Individuation-Asian (*Disinterest* factor), Motivated Individuation-White (*Attention* factor)—and existing-amended scales (0.11 < *r*s < 0.13; *p*s < .03)—Positive Contact-Asian (*Amicable Experiences* factor) and Positive Contact-White. The inferential results are presented in Table 4 in the supplementary materials. Next, we examined whether the CRE-I’s scales improved predictions of accuracy beyond what confidence provides. We expected that scales predictive of accuracy would explain separate variance from confidence in identification accuracy. To test this, identification accuracy was predicted using multilevel logistic regression models with confidence and CRE-I scales as fixed effects and estimating participant as a random effect. The results of these models are summarized here but see Table 5 in the Appendix for the full inferential results. Confidence was always a reliable predictor of accuracy (all *p*s < .001). Several scales were predictive in the models with confidence (*p*s < .03): Race-Specific Face Recognition Ability-Asian (*Difficulty* factor), Racial Attitudes-White, Motivated Individuation-White (*Disinterest* factor), and existing-amended Racial Attitudes-White.

To test whether these three predictive scales predicted overlapping variance, we conducted a stepwise binomial multilevel logistic regression including fixed effects of confidence, Race-Specific Face Recognition Ability-Asian (*Difficulty* factor), Racial Attitudes-White, and Motivated Individuation-White (*Disinterest* factor) which utilized backwards selection. The final model contained the Race-Specific Face Recognition Ability-Asian scale (*Difficulty* factor), *z* = 1.48, *p* = .13, the Racial Attitudes-White scale, *z* = 2.44, *p* = .01, and confidence, *z* = 5.68, *p* < .001.

**Discussion**

We found a CRE for White but not Asian participants. Importantly, the CRE-I was produced with 14 scales and 100 items, pending further validation. Most of the inventory's scales had good internal reliability and convergent validity. Individual differences in responses to half of the scales predicted the identification performance of White participants. Individual differences in three scales explained variance in White participants' identification accuracy beyond what was explained by confidence alone.

**A Minority Effect**

The lack of a CRE for Asian participants may be due to their daily contact with White individuals. Out of 203 Asian participants, 110 reported being originally from the USA, 187 reported currently residing in the USA, and 149 reported having lived longest in the USA. Given their experience in the USA and the fact that the predominant race in most parts of the USA is White, it is perhaps unsurprising that Asian participants did not show a CRE.

Other researchers have also found a CRE for the majority race but not for the numerical minority race (e.g., Chiroro & Valentine, 1995; Havard et al., 2017, 2023). Relevant to this study, Zhou et al. (2019) found that while White participants showed a traditional CRE for Asian faces (c.f. White), Asian participants recognized White and Asian faces equally well. Similarly, Stelter et al. (2023) found across multiple studies a traditional CRE for White participants with Black faces, while Black participants predominantly living in the United States did not exhibit a CRE for White faces. Finally, the most recent meta-analysis found a smaller CRE amongst non-White than White participants (Lee & Penrod, 2022; Figure 4). The literature has predominantly involved samples with limited racial diversity, which may account for why the minority effect—no CRE for numerical minority groups—is rarely reported. Furthermore, researchers may have difficulty publishing failures to find a CRE; therefore, a file drawer effect may also be playing a role. Regardless, these findings highlight the need to incorporate diverse populations and report fully about them.

**Sequential Lineups**

Another finding of interest is that we found a CRE using sequential lineups. Although the CRE has been replicated, the effect has been shown almost exclusively with old/new face recognition paradigms (Lee & Penrod, 2022) and simultaneous lineups (Smith et al., 2001; cf. Wright et al., 2001; Wylie et al., 2015). This is important because sequential lineups are used in some regions, such as the UK, parts of the US, and Canada (Fitzgerald et al., 2021). As such, it is important to examine whether effects on old/new face recognition paradigms and simultaneous lineups are also found with sequential lineups to ensure policymakers and practitioners can make informed decisions about eyewitness identification procedures.

**The CRE-I**

A key goal of our study was to produce reliable items to predict eyewitness identification accuracy in cross-race scenarios. Most of the CRE-I scales had good internal reliability—except the Motivated Individuation and Cognitive Disregard scales—and convergent validity—except the Racial Attitudes-White scale (among those tested). Five of the CRE-I scales (i.e., General Face Recognition Ability, Race-Specific Face Recognition Ability-Asian, Racial Attitudes-Asian, Quantity of Interracial Contact-Asian, Quality of Interracial Contact-Asian) met Streiner’s (1994) rule of thumb for internal validity (i.e., 50% total variance minimum). The remaining nine scales did not meet Streiner’s (1994) suggestion (i.e., Motivated Individuation-Asian, Motivated Individuation-White, Cognitive Disregard-Asian, Cognitive Disregard-White, Race-Specific Face Recognition Ability-White, Racial Attitudes-White, Quantity of Interracial Contact-White, Quality of Interracial Contact-White), though four were above 46%. These results show that the CRE-I has strong internal reliability and validity. Nevertheless, additional validation is needed prior its use in practice.

Importantly, we also tested whether individual differences in responses to CRE-I scales predicted identification accuracy, and individual differences in 7 of the 14 developed scales were predictive. Some of the seven scales measured dimensions only suggested by the perceptual expertise hypothesis (i.e., race-specific face recognition ability) or only the social-cognitive hypothesis (i.e., racial attitudes, motivated individuation, cognitive disregard). Some scales reflect dimensions suggested by both theoretical positions and the contact hypothesis. For example, the quantity of interracial contact predicted identification accuracy, irrespective of target race. Moreover, motivated individuation, a variable suggested by an integrative model (i.e., CIM), predicted identification accuracy. Finally, three of the seven developed scales predicted identification accuracy beyond the predictive ability of confidence, indicating they can improve upon what we learn by asking an eyewitness how confident they are in their identification. These findings underscore the multifaceted nature of face recognition and highlight the importance of considering both perceptual expertise and social-cognitive factors when explaining the differences in face recognition performance across diverse racial groups. Notably, only the Racial Attitudes scale was retained in the stepwise regression model with confidence as a significant predictor of accuracy. This suggests stronger support for social-cognitive hypotheses than perceptual-expertise hypotheses, emphasizing the role of social cues in recognition accuracy.

The SFRS (Bobak et al., 2017) did not predict identification accuracy regardless of target race. The SFRS has been shown to predict face recognition performance (Bobak et al., 2017) and correlates well with other widely used and validated face recognition scales, such as the Cambridge-Face Memory Test (CFMT, Duchaine & Nakayama, 2006). The SFRS's limited predictive ability may be due to chance or that the SFRS is predictive only in some recognition contexts. Considering that individual differences in general and face-race-specific face recognition ability scales, as measured by the CRE-I, were predictive of identification accuracy, it is also possible that the SFRS measures something that applies generally to face recognition but not well to the eyewitness scenario.

Importantly, this study provides the first evidence that self-reported (explicit) racial biases and attitudes towards a race group predict identification accuracy. Although racial attitudes have been explored, research has primarily measured implicit racial biases (e.g., Walker & Hewstone, 2008), and the results are contradictory (Meissner & Brigham, 2001). The disparity between prior findings and our findings could be attributed to the nature of the items themselves or how they were presented. The current items may have more accurately captured the relationship between racial biases and the CRE than previous scales, which were either developed *ad-hoc* or originated over 30 years ago. This finding deserves more empirical attention as the CRE has been considered in relation to racial biases and attitudes only a handful of times (e.g., Anzures et al., 2022). However, it seems more likely that the current findings reflect that participants were not primed to manage their impressions because the questions about racial attitudes—something people may not be willing to communicate honestly—were mixed amongst a very large number of other items (35 items out of 430).

This study also provides the first evidence that self-reports of motivated individuation and cognitive disregard towards a racial group predict identification accuracy. Although these variables have been explored, they have thus far only been manipulated experimentally (e.g., Hugenberg et al., 2007), and the results have been contradictory (e.g., Cruz et al., 2023).

As expected, we found that self-reports of one’s quantity of cross-race interracial contact predicted identification accuracy (Singh et al., 2022). However, what is more important is that now there is a reliable and comprehensive tool for assessing interracial contact—something researchers have been calling for (e.g., Singh et al., 2022). Upon validation, this scale can easily be used in any context in which researchers wish to measure interracial contact. It is worth highlighting that the contact scale includes items relevant to both life-long and recent contact levels. This is relevant theoretically because some theories of the CRE argue for only one or the other as contributing to the CRE (see Chiroro & Valentine, 1995; Sangrigoli et al., 2005; Wan et al., 2015). That items for both were retained during factor analysis and contributed to the predictiveness of the scale in the current study suggests both are relevant. However, researchers may wish to investigate which plays a greater role and the circumstances under which each plays a larger or smaller role.

Regarding the contact hypothesis, there is controversy over whether quality versus quantity of contact is important. Only individual differences in quantity of contact predicted identification accuracy in this study. Other studies have found neither type of contact is predictive (e.g., Havard et al., 2023; McKone et al., 2019). However, research by Stelter et al. (2023) found that frequency of contact during adulthood was predictive of identification accuracy, so there is precedent for the role of quantity of contact. One explanation for why the CRE-I's quantity of contact scale was predictive, in contrast to Havard et al., may be that all the items were presented randomly. Doing this may reduce social desirability bias by making it difficult to discern the dimensions being measured. It remains to be seen whether this scale will remain predictive when the 100 (cf. 430) items are presented.

***Implications for Practice***

Currently, there are no robust training programs or interventions that reliably reduce the CRE and people who regularly engage in face perception do not appear to perform any better than anyone else (Yates et al., 2023). Consequently, a tool to measure the dimensions that contribute to the CRE may be very useful to researchers and practitioners. The CRE-I requires validation, but if successful, it could be provided to eyewitnesses and could act as a reflector variable (Wells, 2020). That is, it could improve the ability of the criminal justice system to predict whether an eyewitness is likely to be accurate and, therefore, to improve their ability to weigh eyewitness evidence. The scale uses Likert-style items that are easy to administer on paper or online, and most people are familiar with responding to these types of questions. Moreover, the items were constructed using simple language; therefore, completing the CRE-I would require limited police resources and could be completed in a relatively short period of time. Thus, the CRE-I may be used by police at the investigation level if desired, or, because the variables measured are unlikely to be time-sensitive, it could be used by the court if the need to evaluate an eyewitness’ accuracy arises. However, police may hesitate to use the CRE-I if they might perceive that a poor response from an eyewitness could undermine the credibility of the identification. Research with police officers would provide further insight into their perspective. We anticipate that the CRE-I is more likely to be used at the court level rather than at the police level, and to advance theory about the CRE.

**Limitations**

In this study, we used the term race to differentiate between the participant groups, individuals self-identifying as White or Asian, while acknowledging variations in ethnicity within each group. These variations may be partially responsible for the lack of CRE elicited from Asian participants. Indeed, an other-ethnicity effect has been shown in the literature. For instance, Black South Africans show better memory for Black South African faces than for African American faces (Chiroro et al., 2008). Future research is needed to discern whether Asians show an other-ethnicity effect, a minority effect, a CRE, or none of these. Moreover, Hunt and Shepherd (2023) noted the importance of considering how race intersects with other social characteristics (e.g., gender, class, sexual orientation, (dis)ability status, immigration status) to understand the impact of race on communities fully. This advice is almost certainly applicable to cross-race identifications. Future research should account for intersectionality and consider the interaction of these types of variables with the CRE.

Given the resources available, several methodological choices, which align with the choices of previous CRE studies, were necessary to achieve reasonable power. Firstly, due to limited resources, we recruited nearly all Asian participants via MT/CR whereas we recruited most White participants via university participant pools. Different recruitment platforms may have influenced the participant demographics and characteristics, potentially leading to variations in the CRE results. Additionally, differences in participant populations recruited through MT/CR compared to university participant pools could introduce variability in factors such as familiarity with diverse racial groups or motivation to complete the study, which in turn might impact their susceptibility to the CRE. Future research should replicate these findings with more balanced White and Asian recruitment approaches. Secondly, we utilized a short delay between each mock-crime video and its associated lineup. Delays in the real world are typically days or weeks— memory has considerably more opportunity to deteriorate. Similarly, we showed participants videos rather than exposing them to staged crimes. Thus, the emotional experience and the perceived importance of the event were likely different than if participants had witnessed an actual crime. This is a limitation of CRE studies in general rather than specific to the study at hand. Thirdly, future research should include more than two targets for each target race condition to ensure findings are not due to any unique characteristics (e.g., distinctiveness) of targets and collect data on the distinctiveness of the targets employed, which would inform generalizability. Fourth, some of the Asian target lineup sizes were smaller than we hoped, based on their resultant Tredoux’s *Es*. This was unexpected because Asians played a key role in lineup construction. Nonetheless, the results in this study may have been affected by differences in fairness. Finally, we utilized British-type sequential lineups. Valentine, Darling, and Memon (2007) found that British sequential lineups (photo or video) resulted in similar correct identifications but fewer false identifications than North American sequential lineups. Thus, replication of this study’s findings across lineup types is needed.

There are several limitations regarding the CRE-I findings. Some individual differences in scale responses, such as those to the quality of interracial contact, were not predictive of identification accuracy. There are several potential explanations as to why. It is possible that those theoretical hypotheses are not accurate in their predictions of the CRE; however, it is also possible that the self-report measures failed to capture the essential mechanisms of these theoretical variables. The CRE-I should be validated with new samples with similar demographics and explored for generalizability with different demographics. The predictive validity of the CRE-I should be assessed, including across lineup types (showups, simultaneous lineups, video lineups). Furthermore, it is acknowledged that self-reports may not always reflect behavior (Baumeister et al., 2007) because of demand characteristics, desirability bias, and memory limitations. Nonetheless, they can be reliable tools and are more easily administered than many objective measures. Critically, we found, despite the limitations of self-reports, that the CRE-I could account for significant variance in identification accuracy. The criminal justice system often relies on inventories (e.g., for risk assessment). Presenting numerous items in a mixed order may reduce reporting biases for the CRE-I. Furthermore, potentially, the inventory could be administered online at the witness's convenience, ensuring their comfort outside of law enforcement contexts.

Finally, we administered a large number of items to select the best items to postdict accuracy, ensuring a comprehensive view of the CRE. To minimize the impact of participant fatigue on responses to individual items, we randomized the items, spreading fatigue randomly across items and scales. While this prevents one scale from being more affected by fatigue than others, concerns about low effort remain. Encouragingly, we did not observe straight-lining (e.g., answering "I agree" to all items). This may be because participants noted that some items were reverse-worded and did not want to be caught providing inconsistent responses, and therefore engaged more thoughtfully with the questionnaire (e.g., DeVellis, 2017). Additionally, the reliability of most of the CRE-I scales was comparable to existing scales’ reliabilities (were other scales existed), suggesting that fatigue may not have significantly impacted our findings. However, fatigue could certainly have introduced noise into our data that would not occur with a shorter scale. Future research should include attention-check questions within the CRE-I scales to quantify participant engagement.

**Conclusions**

In conclusion, we replicated the CRE for White participants with Asian and White targets and did so with British sequential lineups (cf. simultaneous lineups, North American sequential lineups, or face recognition tasks). There was no CRE for Asian participants, which may be due to the minority-race status of Asian participants or the lower lineup size for our Asian compared to White lineups. Perhaps most importantly, we produced the CRE-I, comprising 14 reliable scales that reflected identification accuracy, to be further validated and, if successful, used in the literature and practice. Two scales were generic and are likely to predict performance regardless of target race (i.e., general face recognition ability, motivation to individuate), while the remaining scales were race-specific. The findings indicated that a hybrid model of the CRE is preferable because the scales reflected dimensions suggested only by the perceptual expertise hypothesis (quantity of contact with Asians, race-specific face recognition ability), only by the social-cognitive hypothesis (cognitive disregard of Asians, racial attitudes towards Whites), or only by integrative models (motivated individuation of Whites). The scales also supported the contact hypothesis, as the quantity of contact with Asians scale was predictive of identification accuracy. Notably, three of the developed scales were predictive of identification accuracy beyond the predictive ability of confidence, indicating they provide more information than what is provided by confidence judgements regarding the likely accuracy of eyewitness identifications. The CRE-I contributes to the CRE literature in terms of both theory—via showing which factors among many relate to recognition here and potentially in future studies—and practice—by providing a tool that could improve evaluations of eyewitness reliability.

**References**

Allport, G. W., & Kramer, B. M. (1946). Some roots of prejudice. *The Journal of Psychology*, *22*(1), 9-39.

Anthony, T., Copper, C., & Mullen, B. (1992). Cross-racial facial identification: A social cognitive integration. *Personality and Social Psychology Bulletin*, 18(3), 296–301. https://doi.org/10.1177/0146167292183005

Anzures, G., Mildort, M., Bell, C., & Soethe, E. (2022). Visual examination of own- and other-race face identities and implicit racial bias. *Current Psychology*. <https://doi.org/10.1007/s12144-022-03580-3>

Bache, S. M., & Wickham, H. (2020). Magrittr: A forward-pipe operator for R. *R package version 2.0.1*. <https://CRAN.R-project.org/package=magrittr>

Baldwin, M. R., 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*(4), 505–515. <https://doi.org/10.1177/1368430212460893>

Bate, S., Bennetts, R., Hasshim, N., Portch, E., Murray, E., Burns, E., & Dudfield, G. (2019). The limits of super recognition: An other-ethnicity effect in individuals with extraordinary face recognition skills. *Journal of Experimental Psychology: Human Perception and Performance*, 45, 363– 377. https://doi.org/10.1037/xhp0000607

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1). <https://doi.org/10.18637/jss.v067.i01>

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*(4), 396-403.

Beaudry, J. L., Lindsay, R. C., Leach, A. M., Mansour, J. K., Bertrand, M. I., & Kalmet, N. (2015). The effect of evidence type, identification accuracy, line‐up presentation, and line‐up administration on observers' perceptions of eyewitnesses. *Legal and Criminological Psychology*, *20*(2), 343-364.

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, *18*(8), 706-712.

Bobak, A. K., Mileva, V. R., & Hancock, P. J. B. (2017). Stirling Face Recognition Scale (SFRS): A self-report tool for assessing individual differences in face recognition. *ResearchGate*. <https://www.researchgate.net/publication/312173873>

Brigham, J. C. (1993). College students’ racial attitudes. *Journal of Applied Social Psychology*, *23*(23), 1933-1967.

Brigham, J. C., & Barkowitz, P. (1978). Do “They all look alike?” The effect of race, sex, experience, and attitudes on the ability to recognize faces. *Journal of Applied Social Psychology*, *8*(4), 306-318.

Brigham, J. C., Woodmansee, J. J., & Cook, S. W. (1976). Dimensions of verbal racial attitudes: Interracial marriage and approaches to racial equality. *Journal of Social Issues*, *32*(2), 9-21.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77(3), 305-327.

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior research methods*, *42*(1), 286-291.

Caldara, R., & Fusaroli, R. (2016). Investigating social attention: A case for increasing ecological validity in the study of eye-tracking. *Frontiers in Psychology*, 7, 684.

Canty, A., & Ripley, B. (2017). Package ‘boot’. *Bootstrap Functions. CRAN R Proj*.

Chance, J. E., & Goldstein, A. G. (1996). Eyewitness identification. *Psychological Issues in Eyewitness Identification*, 153.

Chiroro, P., & Valentine, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. *The Quarterly Journal of Experimental Psychology*, *48*(4), 879–894. <https://doi.org/10.1080/14640749508401421>

Chiroro, P. M., Tredoux, C. G., Radaelli, S., & Meissner, C. A. (2008). Recognizing faces across continents: The effect of within-race variations on the own-race bias in face recognition. *Psychonomic Bulletin & Review*, 15, 1089-1092.

Correll, J., Ma, D. S., & Davis, J. P. (2021). Perceptual tuning through contact? Contact interacts with perceptual (not memory-based) face-processing ability to predict cross-race recognition. *Journal of Experimental Social Psychology*, *92*, 104058.

Cruz, F., Palma, T. A., Bansemer, E., Correll, J., Fonseca, S., Gonçalves, P., & Santos, A. S. (2023). Do individuation instructions reduce the cross-race effect? A registered replication of. *Journal of Experimental Social Psychology*, *104*, 104423

DeCarlo, L. T. (1998). Signal detection theory and generalized linear models. *Psychological methods*, *3*(2), 186.

DeVellis, R. F. (2017). Scale development: Theory and applications*.* *Sage Publications.*

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5.

Dowle, M., & Srinivasan, A. (2023, February 17). Extension of “data.frame” [R package data.table version 1.14.8]*. The Comprehensive R Archive Network*. https://cran.r-project.org/web/packages/data.table/

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*(4), 576-585.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics,* 63(6), 1004-1013.

Estudillo, A. J. (2021). Self-reported face recognition abilities for own and other-race faces. *Journal of Criminal Psychology*, 11(2), 105-115.

Fallshore, M., & Schooler, J. W. (1995). Verbal vulnerability of perceptual expertise. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(6), 1608

Ferguson, D. P., Rhodes, G., Lee, K., & Sriram, N. (2001). “They all look alike to me”: Prejudice and cross-race face recognition. *British Journal of Psychology*, *92*, 567-577.

Firke, S. (2023). Janitor: Simple Tools for Examining and Cleaning Dirty Data. <https://CRAN.R-project.org/package=janitor>.

Fitzgerald, R. J., Rubínová, E., & Juncu, S. (2021). Eyewitness identification around the world. In *Methods, Measures, and Theories in Eyewitness Identification Tasks* (pp. 294-322). Routledge.

Ge, L., Zhang, H., Wang, Z., Quinn, P. C., Pascalis, O., Kelly, D., Slater, A., Tian, J., & Lee, K. (2009). Two faces of the other-race effect: Recognition and categorization of Caucasian and Chinese faces. *Perception,* 38(8), 1199–1210. https://doi.org/10.1068/p6136

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*(5), 1105–1122. <https://doi.org/10.1037/a0016548>

Gray, K. L., Bird, G., & Cook, R. (2017). Robust associations between the 20-item prosopagnosia index and the Cambridge Face Memory Test in the general population. *Royal Society Open Science*, *4*(3), 160923.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464.

Gross, T. F. (2009). Own-ethnicity bias in the recognition of Black, East Asian, Hispanic, and White faces. *Basic and Applied Social Psychology*, 31, 128–135. http://dx.doi.org/ 10.1080/01973530902880381.

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other‐race effect in face recognition. *British Journal of Psychology*, *99*(1), 45-56.

Havard, C., Memon, A., & Humphries, J. E. (2017). The own-race bias in child and adolescent witnesses: Evidence from video line-ups. *International Journal of Police Science & Management*, 19(4), 261-272.

Havard, C., Breese, E., Thirkettle, M., Kask, K., Leol, K. L., & Mädamürk, K. (2023). A background of bias: Subtle changes in lineup backgrounds increase the own race bias. *Journal of Police and Criminal Psychology*, 1-12.

Hayward, L. E., Tropp, L. R., Hornsey, M. J., & Barlow, F. K. (2017). Toward a comprehensive understanding of intergroup contact. *Personality and Social Psychology Bulletin*, *43*(3), 347–364. https://doi.org/10.1177/0146167216685291

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*(6), 996-1002.

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*(2), 334-340.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological review*, *117*(4), 1168.

Hunt, J. S., & Shepherd, S. M. (2023). Racial justice in psycholegal research and forensic psychology practice: Current advances and a framework for future progress. *Law And Human Behavior*, *47*(1), 1–11. <https://doi.org/10.1037/lhb0000526>

Hourihan, K. L., Benjamin, A. S., & Liu, X. (2012). A cross-race effect in metamemory: Predictions of face recognition are more accurate for members of our own race. *Journal of Applied Research in Memory and Cognition*, *1*(3), 158–162. <https://doi.org/10.1016/j.jarmac.2012.06.004>

Innocence project: Annual report 2020-2021. Annual Report. (2021a, December 6). <https://report2021.innocenceproject.org/>

Jackiw, L. (2007). Social influences of the cross-race effect. *Faculty of Graduate Studies and Research*, University of Regina

Jacques, C., & Rossion, B. (2006). The speed of individual face categorization. *Psychological science*, 17(6), 485-492.

Koehnken, G., Malpass, R. S., & Wogalter, M. S. (1996). Logical foundations for the identification of individuals. *Psychological Issues in Eyewitness Identification*, 205.

Lavrakas, P. J., Buri, J. R., & Mayzner, M. S. (1976). A perspective on the recognition of other-race faces. *Perception & Psychophysics*, *20*(6), 475-481.

Lee, J., & Penrod, S. D. (2022). Three‐level meta‐analysis of the other‐race bias in facial identification. *Applied Cognitive Psychology*, *36*(5), 1106–1130. <https://doi.org/10.1002/acp.399>

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

Lindzey, G., & Rogolsky, S. (1950). Prejudice and identification of minority group membership. *The Journal of Abnormal and Social Psychology*, *45*(1), 37.

Luus, C. E., & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior*, *15*(1), 43-57.

Lüdecke, D., Patil, I., Ben-Shachar, M. S., Wiernik, B. M., Waggoner, P., & Makowski, D. (2021). see: An R package for visualizing statistical models. *Journal of Open-Source Software*, *6*(64), 3393.

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, and Law*, *7*(1), 98.

MacLin, O. H., & Malpass, R. S. (2003). The ambiguous face illusion. *Perception, 32* (2), 249-252.

Mansour, J. K., Lindsay, R. S., Brewer, N., & Munhall, K. G. (2009). Characterizing visual behaviour in a lineup task. *Applied Cognitive Psychology*, *23*(7), 1012–1026. https://doi.org/10.1002/acp.1570

Marsh, B. J., Pezdek, K., & Ozery, D. H. (2016). The cross-race effect in face recognition memory by bicultural individuals. *Acta Psychologica*, *169*, 38–44. <https://doi.org/10.1016/j.actpsy.2016.05.003>

McKone, E., Dawel, A., Robbins, R. A., Shou, Y., Chen, N., & Crookes, K. (2023). Why the other‐race effect matters: Poor recognition of other‐race faces impacts everyday social interactions. *British Journal of Psychology*, *114*, 230-252.

McKone, E., Wan, L., Pidcock, M., Crookes, K., Reynolds, K. J., Dawel, A., Kidd, E., & Fiorentini, C. (2019). A critical period for faces: Other-race face recognition is improved by childhood but not adult social contact. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-49202-0

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*(1), 3–35. <https://doi.org/10.1037/1076-8971.7.1.3>

Meissner, C. A., Brigham, J. C., & Butz, D. A. (2005). Memory for own‐and other‐race faces: A dual‐process approach. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *19*(5), 545-567.

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition*, *33*, 783-792.

National Registry of Exonerations. (2024). Exonerations in the United States map [Interactive Map]. http://www.law.umich.edu/special/exoneration/Pages/Exonerations-in-the-United-States-Map.aspx

Navon, D. (1992). Selection of lineup foils by similarity to the suspect is likely to misfire. *Law and Human Behavior*, *16*(5), 575-593.

Ng, W., & Lindsay, R. C. L. (1994). Cross-race facial recognition: Failure of the contact hypothesis. *Journal of Cross-Cultural Psychology*, 25(2), 217–232. https://doi.org/10.1177/0022022194252004

O'Toole, A. J., Deffenbacher, K. A., Valentine, D., & Aladi, H. (1995). Structural aspects of face recognition and the other race effect. Memory and Cognition, 22, 208–224. https://doi.org/10.3758/bf03208892

Platz, S. J., & Hosch, H. M. (1988). Cross‐racial/ethnic eyewitness identification: A field study. *Journal of Applied Social Psychology*, *18*(11), 972-984.

Quigley-McBride, A., & Wells, G. L. (2021). Methodological considerations in eyewitness identification experiments. In *Methods, measures, and theories in eyewitness identification tasks* (pp. 85-112). Routledge.

R Core Team (2022). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL: <https://www.R-project.org/>.

Reardon, S. F., Matthews, S. A., O’Sullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., & Bischoff, K. (2008). The geographic scale of metropolitan racial segregation. *Demography*, *45*(3), 489-514.

Revelle, W. (2021). psych: Procedures for psychological, psychometric, and personality research (Version 2.1.6). [Computer software]. Retrieved from <https://cran.r-project.org/package=psych>

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British journal of psychology*, *80*(3), 313-331.

Rhodes, G., Ewing, L., Hayward, W. G., Maurer, D., Mondloch, C. J., & Tanaka, J. W. (2009). Contact and other‐race effects in configural and component processing of faces. *British Journal of Psychology*, 100(4), 717-728.

Rhodes, G., Lie, H. C., Ewing, L., Evangelista, E., & Tanaka, J. W. (2010). Does perceived race affect discrimination and recognition of ambiguous-race faces? A test of the sociocognitive hypothesis. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36(1), 217–223. [https://doi.org/10.1037/a0017680](https://psycnet.apa.org/doi/10.1037/a0017680)

Rhodes, M. G., Sitzman, D. M., & Rowland, C. A. (2014). Monitoring and control of learning own‐race and other‐race faces. *Applied Cognitive Psychology*, *27*(5), 553-563.

Robertson, D. J., Black, J., Chamberlain, B., Megreya, A. M., & Davis, J. P. (2020). Super‐recognisers show an advantage for other race face identification. *Applied Cognitive Psychology*, *34*(1), 205-216

Rodin, M. J. (1987). Who is memorable to whom: A study of cognitive disregard. *Social Cognition*, *5*(2), 144-165.

Rule, N. O., Ambady, N., Adams, R. B., & Macrae, C. N. (2007). Us and them: Memory advantages in perceptually ambiguous groups. *Psychonomic Bulletin & Review, 14*, 687-692.

Sedlacek, W. E., & Brooks Jr, G. C. (1970). The measurement of attitudes of Whites toward Blacks with certain beliefs. *U.S. Department of Health, Education, and Welfare.*

Samuels, P. (2016). Advice on exploratory factor analysis *Birmingham City University.* <https://doi.org/10.13140/RG.2.1.5013.9766>

Sangrigoli, S., Pallier, C., Argenti, A.-M., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science*, 16(6), 440–444. https:// doi.org/10.1111/j.0956-7976.2005.01554.x

Scheck, B., Neufeld, P., & Dwyer, J. (2003). Actual innocence: When justice goes wrong and how to make it right. *LARC Cardozo Law*. https://larc.cardozo.yu.edu/faculty-books/78/

Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015). The 20-item prosopagnosia index (PI20): a self-report instrument for identifying developmental prosopagnosia. *Royal Society open science*, *2*(6), 140343

Shriver, E. R., Young, S. G., Hugenberg, K., Bernstein, M. J., & Lanter, J. R. (2008). Class, race, and the face: Social context modulates the cross-race effect in face recognition. *Personality and Social Psychology Bulletin*, *34*(2), 260-274.

Shriver, E. R., & Hugenberg, K. (2010). Power, individuation, and the cross-race recognition deficit. *Journal of Experimental Social Psychology*, *46*(5), 767-774.

Singh, B., Mellinger, C., Earls, H. A., Tran, J., Bardsley, B., & Correll, J. (2022). 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.

Slone, A. E., Brigham, J. C., & Meissner, C. A. (2000). Social and cognitive factors affecting the own-race bias in whites. *Basic and Applied Social Psychology*, *22*(2), 71–84. https://doi.org/10.1207/s15324834basp2202\_1

Smith, S. M., Lindsay, R. S., Pryke, S., & Dysart, J. E. (2001). Postdictors of eyewitness errors: Can false identifications be diagnosed in the cross-race situation? *Psychology, Public Policy and Law*. https://doi.org/10.1037/1076-8971.7.1.153

Smith, S.M., Stinson, V., & Prosser, M.A. (2004). Do they all look alike? An exploration of decision-making strategies in cross-race facial identifications. *Canadian Journal of Behavioral Science/Revue Canadienne des sciences du Behaviour*, 36 (2), 146.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, *7*(1), 36.

Stephan, W. G., & Stephan, C. W. (1993). Cognition and affect in stereotyping: Parallel interactive networks. In *Affect, Cognition and Stereotyping* (pp. 111-136). Academic Press.

Steblay, N. K. (2011). What we know now: The Evanston Illinois field lineups. *Law and Human Behavior*, *35*, 1-12.

Stelter, M., Simon, D., Calanchini, J., Christ, O., & Degner, J. (2023). Real‐life outgroup exposure, self‐reported outgroup contact and the other‐race effect. *British Journal of Psychology*, *114*, 150-171.

Streiner, D. L. (1994). Figuring out factors: the use and misuse of factor analysis. *The Canadian Journal of Psychiatry*, *39*(3), 135-140.

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*(1), B1-B9.

Thompson, S. G. (2007). Beyond a reasonable doubt-reconsidering uncorroborated eyewitness identification testimony. *UC Davis L. Rev*., 41, 1487.

Tracy, R. E., Zomberg, D., & Young, S. G. (2023). On the role of experience versus motivation in predicting the own‐race effect. *British Journal of Psychology*, *114*, 172-187.

Tredoux, C. G. (1998). Statistical inference on measures of lineup fairness. *Law and Human Behavior*, 22(2), 217-237.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion and race in face recognition. *The Quarterly Journal of Experimental Psychology, 43A,* 161-204.

Valentine, T., Darling, S., & Memon, A. (2007). Do strict rules and moving images increase the reliability of sequential identification procedures?. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *21*(7), 933-949.

Walker, L. (1994). Attitudes to minorities: Survey evidence of Western Australians' attitudes to Aborigines, Asians, and women. *Australian Journal of Psychology*, *46*(3), 137-143.

Walker, P. J., & Hewstone, M. (2006). A perceptual discrimination investigation of the own-race effect and intergroup experience. *Applied Cognitive Psychology*, *20*(4), 461–475. <https://doi.org/10.1002/acp.1191>

Walker, P. M., & Hewstone, M. (2008). The influence of social factors and implicit racial bias on a generalized own‐race effect. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *22*(4), 441-453.

Walker, P. J., & Tanaka, J. W. (2003). An encoding advantage for own-race versus other-race faces. *Perception*, *32*(9), 1117–1125. <https://doi.org/10.1068/p5098>

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

Wells, G. L., & Olson, E. S. (2001). The other-race effect in eyewitness identification: What do we do about it? *Psychology, Public Policy and Law*, *7*(1), 230–246. <https://doi.org/10.1037/1076-8971.7.1.230>

Wells, G. L. (2020). Psychological science on eyewitness identification and its impact on police practices and policies. *American Psychologist*, *75*(9), 1316.

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, 25(9), 1115-1125.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., ... & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open-Source Software*, *4*(43), 1686.

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, *18*(1), 10-65.

Wright, D. B., Boyd, C., & Tredoux, C. (2001). A field study of own-race bias in South Africa and England. *Psychology, Public Policy and Law*, *7*(1), 119–133. <https://doi.org/10.1037/1076-8971.7.1.119>

Wright, D. B., & London, K. (2009). Multilevel modelling: Beyond the basic applications. *British Journal of Mathematical and Statistical Psychology*, *62*(2), 439-456.

Wylie, L. E., Bergt, S., Haby, J., Brank, E. M., & Bornstein, B. H. (2015). Age and lineup type differences in the own-race bias. *Psychology, Crime & Law*, *21*(5), 490-506.

Yates, A. N., Cavazos, J. G., Jeckeln, G., Hu, Y., Noyes, E., Hahn, C. A., ... & Phillips, P. J. (2023). Who is that? Perceptual expertise on other-race face comparisons, disguised face comparisons, and face memory. *Technical report, Psyarxiv*.

Young, S. L., Hugenberg, K., Bernstein, M. S., & Sacco, D. F. (2012). Perception and motivation in face recognition. *Personality and Social Psychology Review*, *16*(2), 116–142. <https://doi.org/10.1177/1088868311418987>

Zhou, X., Elshiekh, A., & Moulson, M. C. (2019). Lifetime perceptual experience shapes face memory for own- and other-race faces. *Visual Cognition*, 27(9-10), 687– 700. <https://doi.org/10.1080/13506285.2019.1638478>

**APPENDIX**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scales of the CRE-I and their subscales | # of items | Range of Item Loadings | Total variance accounted for (%) | Cronbach’s *α* | Cutoff |
| General Face Recognition Ability | 4 | .73 - .79 | 56 | .84 | .6 |
| Race-Specific Face Recognition Ability - Asian |  |  |  |  |  |
| Difficulty  Noticing | 3  3 | .52 - .84  .61 - .74 | 28  25 | .71 | .5 |
| Race-Specific Face Recognition Ability-White |  |  |  |  |  |
| Difficulty  Noticing | 3  3 | .69 - .78  .49 - .78 | 27  22 | .70 | .45 |
| Racial Attitudes-Asian | 6 | .65 - .88 | 62 | .90 | .6 |
| Racial Attitudes-White | 8 | .63-.73 | 45 | .88 | .6 |
| Quantity of Contact-Asian |  |  |  |  |  |
| Interaction  Presence | 8  4 | .64 - .73  .60- .86 | 32  20 | .90 | .6 |
| Quantity of Contact-White | 9 | .60 - .75 | 46 | .88 | .6 |
| Quality of Contact-Asian |  |  |  |  |  |
| Value  Reliability | 8  6 | .71 - .85  .72 - .82 | 37  25 | .90 | .6 |
| Quality of Contact-White |  |  |  |  |  |
| Value  Reliability | 3  3 | .50 - .86  .63 - .70 | 26  22 | .77 | .45 |
| Motivated Individuation- General | 3 | .41 - .68 | 34 | .58 | .35 |
| Motivated Individuation-Asian |  |  |  |  |  |
| Attention  Disinterest | 6  3 | .41 - .75  .50 - .76 | 19  15 | .71 | .4 |
| Motivated Individuation-White |  |  |  |  |  |
| Attention  Disinterest | 6  3 | .42 - .56  .44 - .75 | 16  12 | .6 | .4 |
| Cognitive Disregard-Asian | 4 | .34 - .72 | 26 | .55 | .35 |
| Cognitive Disregard- White | 4 | .39 - .58 | 22 | .51 | .35 |

**Table 1**

*Factor Structure and Reliability of Developed Scales*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Scales of the CRE-I and their subscales | | # of items | Range of Item Loadings | Total variance accounted for (%) | Cronbach’s *α* | Cutoff |
| Positive Contact-Asian |  | |  |  |  |  |
| Amicable Experiences  Supportive Bonds | 3  3 | | .60 - .90  .64 - .67 | 26  22 | .78 | .55 |
| Positive Contact- White | 6 | | .61 - .67 | 41 | .80 | .6 |
| Negative Contact Asian | 7 | | .75 - .87 | .63 | .93 | .6 |
| Negative Contact White |  | |  |  |  |  |
| Unpleasant Interactions  Unpleasant Feelings | 4  3 | | .60 - .77  .57 - .76 | 29  21 | .79 | .6 |
| Racial Attitudes- Asian | 3 | | .67 - .85 | 61 | .82 | .6 |
| Racial Attitudes- White | 4 | | .56 - .81 | 48 | .79 | .55 |

**Table 2**

*Factor Structure and Reliability of Existing-Amended Scale*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 3**  *Correlations between Developed Scales and Existing Amended Scales (reported by Pearson`s* *r)* | | | | | | | | | | | | | | | | | | | | | |
|  | | Existing Amended Scales | | | | | | | | | | | | | | | | | | | |
|  | Hayward et al. (2017) | | | | | | | | | | | | | Brigham (1993) | | | | | | Bobak et al. (2019) | |
| Developed Scales | Positive Contact- Asian | | | Positive Contact- White | | Negative Contact - Asian | | | Negative Contact- White | | | Racial Attitudes- Asian | | | | Racial Attitudes-White | | | Stirling | | |
|  | Amicable Experiences | | Supportive Bonds | |  | |  | | | Unpleasant Interactions | Unpleasant  Feelings | |  | | | | |  | | |  |
| General Face Recognition Ability |  | |  | |  | |  |  | | | | | |  |  | |  | | | *r =* .78,  *p <* .001 | |
| Quantity of Contact-Asian |  | | | | | | | | | | | | | | | | | | | | |
| Interaction | *r =* .26,  *p <* .001 | | *r =* .15,  *p =* .03 | |  | | *r =* .28,  *p <* .001 |  | | | | | |  |  | |  | | |  | |
| Presence | *r =* .01,  *p =* .85 | | *r =* -.17,  *p =* .01 | |  | | *r =* .67,  *p <* .001 |  | | | | | |  |  | |  | | |  | |
| Quantity of Contact-White |  | |  | | *r =* .52,  *p <* .001 | |  | *r =* .18,  *p =* .007 | | | | | | *r =* -.20,  *p =* .003 |  | |  | | |  | |
| Quality of Contact- Asian |  | |  | |  | |  |  | | | | | |  |  | |  | | |  | |
| Value | *r =* .19,  *p =* .008 | | *r =* .05,  *p =* .47 | |  | | *r =* .47,  *p <* .001 |  | | | | | |  |  | |  | | |  | |
| Reliability | *r =* .63,  *p <* .001 | | *r =* .56,  *p <* .001 | |  | | *r =* -.24,  *p <* .001 |  | | | | | |  |  | |  | | |  | |
| Quality of Contact- White |  | | | | | | | | | | | | | | | | | | | | |
| Value |  | |  | | *r =* .58,  *p <* .001 | |  | *r =* .14,  *p =* .04 | | | | | | *r =* -.19,  *p =* .008 |  | |  | | |  | |
| Reliability |  | |  | | *r =* .68,  *p <* .001 | |  | *r =* .03,  *p =* .58 | | | | | | *r =* .07,  *p =* .32 |  | |  | | |  | |
| Racial Attitudes- Asian |  | |  | |  | |  |  | | | | | |  | *r =* .67,  *p <* .001 | |  | | |  | |
| Racial Attitudes- White |  | |  | |  | |  |  | | | | | |  |  | | *r =* .49,  *p <* .001 | | |  | |

**Table 4**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Scales of the CRE-I and their subscales | *z* | *p* | *corrected p+* | *OR* |
| General Face Recognition Ability | 2.43 | .01\* | .02\* | 1.24 |
| Race-Specific Face Recognition Ability-Asian  Difficulty  Noticing | 2.30  0.50 | .02\*  .61 | .02\*  .78 | 1.36  0.94 |
| Race-Specific Face Recognition Ability-White  Difficulty  Noticing | 2.36  0.34 | .01\*  .72 | .03\*  .72 | 1.29  1.04 |
| Racial Attitudes-Asian | 1.73 | .08 | .08 | 1.19 |
| Racial Attitudes-White | 3.22 | .001\* | .002\* | 1.42 |
| Quantity of Contact-Asian  Interaction  Presence | 1.21  2.09 | .22  .03\* | .22  .053 | 1.14  0.79 |
| Quantity of Contact-White | 1.34 | .17 | .17 | 1.21 |
| Quality of Contact-Asian  Value  Reliability | 1.42  0.25 | .15  .79 | .22  .79 | 0.87  1.03 |
| Quality of Contact-White  Value  Reliability | 1.14  0.54 | .25  .58 | .38  .58 | 1.19  0.92 |
| Motivated Individuation- General | 0.37 | .70 | .70 | 1.04 |
| Motivated Individuation-Asian  Attention  Disinterest | 1.41  0.78 | .15  .43 | .23  .43 | 0.86  1.31 |
| Motivated Individuation-White  Attention  Disinterest | 1.14  2.92 | .25  .003\* | .25  .01\* | 1.17  0.92 |
| Cognitive Disregard-Asian | 2.07 | .03\* | .03\* | 1.28 |
| Cognitive Disregard- White | 1.57 | .11 | .11 | 1.24 |
| Positive Contact-Asian  Amicable Experiences  Supportive Bonds | 1.22  0.26 | .22  .79 | .33  .79 | 1.16  1.03 |
| Positive Contact- White | 1.03 | .30 | .30 | 1.17 |
| Negative Contact Asian | 1.69 | .09 | .09 | 0.84 |
| Negative Contact White  Unpleasant Interactions  Unpleasant Feelings | 0.89  1.41 | .36  .15 | .36  .23 | 1.09  0.86 |
| Racial Attitudes- Asian | 1.99 | .04\* | .04\* | 1.20 |
| Racial Attitudes- White | 3.05 | .002\* | .004 | 1.35 |

*Inferential Results of Multiple Logistic Regression Models Including CRE-I Scales as Predictors of Identification Accuracy*

*Note.* All models include Target Race. *+p*-values corrected for the False Discovery Rate correction (Benjamini & Hochberg, 1995) a. \* *p* < .05

**Table 5**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Scales of the CRE-I | *z* | *p* | *corrected p+* | *OR* |
| General Face Recognition Ability | 1.67 | .09 | .09 | 1.20 |
| Race-Specific Face Recognition Ability-Asian *Difficulty* | 2.13 | .008\* | .008\* | 1.31 |
| Race-Specific Face Recognition Ability-White *Difficulty* | 1.76 | .07 | .07 | 1.25 |
| Racial Attitudes-White | 2.60 | .009\* | .009\* | 1.42 |
| Quantity of Contact-Asian *Presence* | 1.36 | .17 | .17 | 0.85 |
| Motivated Individuation-White *Disinterest* | 2.46 | .01\* | .01\* | 1.34 |
| Cognitive Disregard-Asian | 1.35 | .17 | .17 | 1.22 |
| Racial Attitudes-Asian | 1.81 | .06 | .06 | 1.23 |
| Racial Attitudes-White | 2.80 | .004\* | .004\* | 1.39 |

*Inferential Results of Multiple Logistic Regression Models Including Both Confidence and CRE-I Scales Predictive of Identification* *Accuracy as Predictors of Identification Accuracy.*

*Note.* *+p*-values corrected for the False Discovery Rate correction (Benjamini & Hochberg, 1995) a. \* *p* < .05

**Table 6**

*Means, Standard Deviations, and Internal Consistency Estimates for Factors and Items in the Final CRE-I Existing-Amended Scales*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Factors & Items | | *M* (*SD*) | Item-Total Correlation | *ɑ* if item deleted |
| Positive Contact-Asian (*ɑ* = .79 95% CI [.73, .83]) | |  |  |  |
| Positive Contact-Asian, *Amicable Experiences* (*ɑ* = .77 95% CI [.71, .82]) | | 4.49 (0.86) |  |  |
| I feel my Asian friends are interested in what I have to say. | | 4.65 (0.91) | .74 | .65 |
| I feel free to express myself when I encounter Asian people. | | 4.50 (1.03) | .65 | .57 |
| My Asian friends are generous to me. | | 4.38 (1.06) | .68 | .59 |
| Positive Contact-Asian, *Supportive Bonds* (*ɑ* = .70 95% CI [.61, .76]) | | 4.88 (0.78) |  |  |
| The Asian people I know have been polite to me. | | 4.86 (0.90) | .64 | .53 |
| I have had pleasant interactions with Asian people. | | 5.02 (0.83) | .60 | .49 |
| I have had comfortable interactions with Asian people. | | 4.77 (1.10) | .59 | .50 |
| Positive Contact-White (*ɑ* = .80 95% CI [.75, .84]) | | 4.74 (0.65) |  |  |
| The White people I know treat me as equals. | | 4.72 (0.90) | .60 | .54 |
| My White friends respect me. | | 4.87 (0.75) | .65 | .58 |
| My White friends make me feel valued. | | 4.67 (0.90) | .59 | .52 |
| The White people I know have been polite to me. | | 4.61 (0.93) | .60 | .53 |
| The White people I know have been kind to me. | | 4.68 (0.89) | .64 | .58 |
| The White people I know have been friendly towards me. | | 4.91 (0.82) | .64 | .57 |
| Negative Contact-Asian (*ɑ* = .93 95% CI [.91, .94]) | | 2.06 (0.97) |  |  |
| Asian people have patronized me. | | 2.33 (1.23) | .81 | .78 |
| Asian people have made me feel worthless. | | 2.05 (1.18) | .76 | .72 |
| My Asian friends have treated me as inferior. | | 2.19 (1.15) | .82 | .78 |
| Asian people have exploited me. | | 2.03 (1.22) | .75 | .72 |
| I have been threatened by Asian people. | | 2.10 (1.17) | .87 | .83 |
| Asian people have tried to fight me. | | 2.03 (1.16) | .81 | .77 |
| Asian people have physically harmed me. | | 1.95 (1.32) | .77 | .74 |
| Negative Contact-White (*ɑ* = .79 95% CI [.74, .83]) | |  |  |  |
| Negative Contact-White, *Unpleasant Interactions* (*ɑ* = .80 95% CI [.75, .84]) | | 4.25 (1.04) |  |  |
| White people have been rude to me. | | 4.40 (1.18) | .69 | .61 |
| White people have insulted me. | | 4.37 (1.24) | .71 | .63 |
| I have had hostile interactions with White people. | | 4.08 (1.39) | .60 | .54 |
| White people have verbally insulted me. | | 4.11 (1.37) | .73 | .65 |
| Negative Contact-White, *Unpleasant Feelings* (*ɑ* = .73 95% CI [.66, .79]) | | 3.09 (1.06) |  |  |
| My White friends are arrogant towards me. | | 2.93 (1.18) | .58 | .50 |
| My White friends have treated me as inferior. | | 3.26 (1.29) | .70 | .60 |
| White people have exploited me. | | 3.17 (1.38) | .66 | .56 |
| Racial Attitudes – White Scale (*ɑ* = .79 95% CI [.75, .84]) | | 4.92 (0.99) |  |  |
| I worry that my application for a job or a promotion may be denied in the next few years because of preferential treatment given to minority group members. (R) | | 4.59 (1.21) | .64 | .58 |
| Local city officials often pay less attention to a request or complaint from a White person than from an Asian person. (R) | | 4.31 (1.37) | .63 | .56 |
| I would rather not have Whites live in the same apartment building I live in. (R) | | 5.14 (1.17) | .73 | .63 |
| When I see an interracial couple, I feel that they are making a mistake in dating each other. (R) | | 5.32 (1.21) | .74 | .64 |
|  | Racial Attitudes – Asian Scale (*ɑ* = .82 95% CI [.77, .86]) | 5.16 (1.01) |  |  |
| I would rather not have Asians live in the same apartment building I live in. (R) | | 5.37 (1.03) | .67 | .63 |
| Generally, Asian people are not as smart as White people. (R) | | 4.70 (1.21) | .76 | .68 |
| When I see an interracial couple, I feel that they are making a mistake in dating each other. (R) | | 5.32 (1.21) | .76 | .68 |
| *Note*. *ɑ* = Cronbach’s alpha.  *M* = Mean scale scores. *SD* = Standard deviation. Higher values indicated greater agreement with the statement. (R) indicates reverse worded items. | | | | |

1. Race refers to a group of humankind with shared distinctive physical traits and can be understood as distinct from ethnicity, which is a population group comprised of individuals with shared cultural heritage or ancestry. However, typically in the CRE literature, it is used interchangeably with ethnicity, therefore, we will also take this approach. Operationally, in this study, race refers to individuals who self-identify as White or Asian as participants, and the actors starring in the mock-crime videos who self-identified as White or East Asian. [↑](#footnote-ref-1)
2. Other Asian countries reported by participants to be as country of origin are: 11 from India, 9 from China, 9 from Korea, 5 from Vietnam, 5 from the Philippines, 4 from Hong Kong, 3 from Indonesia, 3 from Pakistan, 2 from Taiwan, 2 from Malaysia, 1 from Bangladesh, 1 from Kazakhstan, 1 from Liberia, 1 from Nepal, 1 from Singapore, 1 from Thailand, 1 from Laos, and 1 from Trinidad. [↑](#footnote-ref-2)
3. 11 India, 9 China, 2 Hong Kong, 2 Malaysia, 2 Pakistan, 1 Kazakhstan, 1 Korea, 1 Singapore, 1 Taiwan, 1 Thailand, 1 Vietnam, 1 Pakistan. [↑](#footnote-ref-3)
4. 6 Poland, 4 Italy, 2 Hungary, 2 Ireland, 1 Egypt, 1 Finland, 1 Greece, 1 Latvia, 1 Lithuania, 1 Norway, 1 Romania, 1 South Africa [↑](#footnote-ref-4)
5. 2 Hungary, 2 Ireland, 2 Italy, 1 Egypt, 1 Finland, 1 France, 1 Greece, 1 Norway, 1 Poland, 1 Romania, 1 South Africa, 1 Netherlands, 1 United Arab Emirates [↑](#footnote-ref-5)
6. TM and © 2008 Entertainment Rights Distribution Limited. All rights reserved. [↑](#footnote-ref-6)
7. The race-neutral scale item was included in analyses along with each race-specific cognitive disregard scale. [↑](#footnote-ref-7)
8. The race-neutral scale items were included in analyses along with each race-specific scale. [↑](#footnote-ref-8)
9. For the existing amended scales, we could have conducted an EFA, considering the alterations from the original scale, or a confirmatory factor analysis (CFA), acknowledging that the scales are derived from a validated instrument. Based on discussions with an expert in scale development, we conducted an EFA to ensure only items most representative for our samples were used. [↑](#footnote-ref-9)
10. After the EFA, we did not exclude any scale based on low reliability or low factor loadings because retaining all scales allowed us to preserve the theoretical integrity of the constructs being measured. [↑](#footnote-ref-10)
11. Confidence predicted accuracy also for Asian participants, *z* = 6.55, *p* < .001. [↑](#footnote-ref-11)