

REVIEW

Open Access



Understanding racial bias through electroencephalography

Mirella Manfredi^{1,2*}, William E. Comfort³, Lucas M. Marques⁴, Gabriel G. Rego³, Julia H. Egito³, Ruth L. Romero³ and Paulo S. Boggio³

Abstract

Research on racial bias in social and cognitive psychology has focused on automatic cognitive processes such as categorisation or stereotyping. Neuroimaging has revealed differences in the neural circuit when processing social information about one's own or another's ethnicity. This review investigates the influence of racial bias on human behaviour by reviewing studies that examined changes in neural circuitry (i.e. ERP responses) during automatic and controlled processes elicited by specific tasks. This systematic analysis of specific ERP components across different studies provides a greater understanding of how social contexts are perceived and become associated with specific stereotypes and behavioural predictions. Therefore, investigating these related cognitive and neurobiological functions can further our understanding of how racial bias affects our cognition more generally and guide more effective programs and policies aimed at its mitigation.

Keywords Event-related potentials, Social neuroscience, Racial prejudice, Face recognition memory, Implicit prejudice, Empathy to pain

Understanding racial bias through electroencephalography

The year 2020 was marked by large anti-racism demonstrations in many western countries and a broader discussion of racial bias and discrimination in the media. Although these demonstrations gained media attention due to George Floyd's death, they are only the most recent manifestation of a continuous and enduring struggle to address social inequalities between racial groups.

In addition to adopting policy measures targeting these inequalities, an essential part of this broad movement has been the role of scientific investigations to identify the cognitive aspects involved in racial bias, racial prejudice and racial stereotypes. Racial bias is defined as pre-conceived beliefs, attitudes, and expectations about racial group members [1]. Racial prejudice is an attitude, usually negative, directed towards another ethnic group). Finally, Racial stereotypes are a broad group generalisations that tend to disproportionately position some groups as better than others [2]. Although the terms are often used interchangeably in the scientific literature, in this review we will mainly use the term racial bias.

Research on racial bias in social and cognitive psychology has focused on automatic cognitive processes such as categorisation or stereotyping [3, 4]. More recently, neuroimaging has revealed differences in the activation of brain regions, including amygdala, anterior cingulate cortex (ACC) and dorsolateral and ventrolateral prefrontal cortex, when processing social information about

*Correspondence:

Mirella Manfredi
mirella.manfredi@uzh.ch

¹ Department of Psychology, University of Zurich, Zurich, Switzerland

² Jacobs Center for Productive Youth Development, University of Zurich, Zurich, Switzerland

³ Social and Cognitive Neuroscience Laboratory, Developmental Disorders Program, Center for Health and Biological Sciences, Mackenzie Presbyterian University, Sao Paulo, Brazil

⁴ Institute of Physical Medicine and Rehabilitation, Clinical Hospital, University of Sao Paulo Medical School, São Paulo, Brazil



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

one's own or another's ethnicity. These brain areas are also implicated in related cognitive functions such as face detection, affective evaluation and behavioural regulation [5]. The investigation of these related cognitive and neurobiological functions can further our understanding of how racial bias affects our cognition more generally and guide more effective programs and policies aimed at its mitigation, as suggested by Denson and Chang, [6]. They demonstrate that programs aimed at greater racial or ethnic diversity in education can reduce racial bias.

One of the main difficulties in this field of research has frequently been to consistently identify prejudice or bias through explicit instruments such as questionnaires or scales, for two reasons: participants may modulate their behaviour not to be seen as prejudiced given the sensitive content of the topic [7], and because of the effect of implicit prejudices and stereotypes. These phenomena might include negative attitudes an individual holds towards a social group in the absence of conscious awareness, which are resistant to detection in behavioural tasks and may negatively influence an individual's perception and behaviour in situations of low cognitive control such as quick or automatic decisions (e.g., [7, 8]).

Given this, the adoption of techniques from the neurosciences has been of great value for identifying patterns in brain activity associated with racial bias that do not require explicit responses. Of particular relevance are those techniques that rely on brain electrophysiology recordings, such as electroencephalography (EEG) and magnetoencephalography (MEG). EEG has been consistently used to investigate temporal aspects in socially-relevant information's perceptual and semantic processing. Because accurate temporal resolution is crucial for pinpointing and analysing the different processes involved in racial bias, event-related potentials (ERPs) measured via electroencephalography (EEG) are particularly well-suited to this aim. ERPs reflect the postsynaptic activity of an extensive set of neurons activated in close temporal proximity and can be recorded by placing several electrodes on the surface of the participant's head [9]. The electric wave detected in the ERP is composed by the summing of subjacent components, where each component represent a brain activity associated with a specific cognitive process and which is characterised by: (a) *latency*, at what moment in time they occur after the onset of a specific stimulus; (b) *direction*, either positive- or negative-going; (c) *amplitude* of the wave, which represents the demand of brain activity and indirectly the level of cognitive engagement of a specific process; (d) *localisation* over the scalp, indicating in what electrode region (from an international measurement system) the component is usually detected; and (e) *source*, i.e., the probable cortical brain region (or regions) originating

the signal [10]. Differences in ERP component amplitude usually indicate distinct cognitive demand levels of cognitive processes (e.g., attention, memory, motivation, etc.) or distinct cognitive mechanisms involved in information processing. ERP latency indicates the specific moment where this process occurs after stimulus appearance.

ERP components involved in racial bias

A significant body of literature has employed an ERP design to investigate electrophysiological responses associated with the racial bias process phases. These studies identified the following ERP components: N1/N170, P2, N2/N250, P300, ERN, Pe, N400 and LPP.

The N1 is the first negative deflection on posterior scalp regions, which can be observed in response to any visual information. The N1 has a peak latency of 130–200 ms. The N170 is an early component that has been associated with the perceptual processing of faces and peaks on average at about 160–170 ms for these stimuli. Various kinds of information can also be extracted quickly and efficiently from the face to categorise the person's gender, facial expression, ethnic origin, and gaze direction [11]. The anterior N2/N250 peaks approximately 250 ms after stimulus presentation and appears to index selective attention to specific stimulus features (e.g., colour, shape, form; [12]. An N2 effect has been observed for stimuli containing conflicting features, as in the flanker and Stroop paradigms.

Similarly to N2, the MFN is sensitive to conflict, but it appears to be modulated by expectations in context: when most trials within a block are congruent, the MFN is larger for incongruent trials [13].

The P300 usually peaks between 300 and 500 ms after stimulus onset and is classically associated with task relevance. The P300 is larger for infrequent target stimuli and is strongly modulated by attention [14]. The P300 is also sensitive to emotional information (both pleasant and unpleasant stimuli), suggesting the P300 may reflect a general motivation [15, 16].

The error-related negativity (ERN) reflects error detection in many tasks. It is sensitive to different types of sensory information, such as auditory, visual, and somatosensory. This component typically occurs at the moment of the erroneous button press and peaks around 100 ms later. This component also represents a self-monitoring mechanism that allows volunteers to improve their cognitive control over prepotent responses continually.

The error positivity (Pe) component usually follows the ERN. It is usually larger in response to errors that the participant reports compared with those that are not reported. This component has been associated with an affective response to the error, awareness of the error, or adapting response strategies after error detection.

A significant body of literature has investigated semantic processing by analysing the N400, a negative event-related brain potential that peaks roughly 400 ms after the onset of a stimulus [17, 18]. The N400 peaks roughly 400 ms after the onset of a stimulus and reflects the spreading activation in the access of semantic information by a stimulus in relation to its preceding context [18]. This component is sensitive to semantic violations related to a large array of meaningful information. The late positivity (LP/P600) is a slow late positive shift that typically onsets around 500 ms after the onset of a stimulus. It has been related to a process of reanalysis of the incongruent situation produced by inconsistent meaning [19–21], particularly when the incoming information disconfirms predictions created by a preceding context [22].

Aim of the present review

When studying racial bias, the ERP technique allows us to understand further the role of subjacent cognitive processes (e.g., attention, memory) involved in the detection, categorisation, and reaction to stimuli involving same- or other-race. It also enlightens us about the correlation of the ERP components to other psychological aspects such as personality traits or explicit prejudice/stereotyping. Besides, understanding when and how social information is processed using ERP could help us make inferences about brain processes dynamics involved in racial bias and social cognition overall [5].

This review investigates how racial bias, implicit or explicit, influences human behaviour. To this aim, we identify the social and cognitive neuroscience topics that adopted ERP and the characteristics of the detected ERP components. The investigation of specific ERP components allows us to analyse changes in neural circuitry during automatic and controlled processes elicited by specific tasks. Therefore, a systematic analysis of specific ERP components across different studies provides a greater understanding of how social contexts are perceived and become associated with specific stereotypes and behavioral predictions.

Based on the available studies on this topic, we will break the discussion down into the following main sections: (i) *Face recognition*; (ii) *Implicit bias*, including studies addressing implicit prejudice and stereotyping; and (iii) *Pain empathy*.

Face recognition

A widely studied topic in neuroscience of racial bias is detecting and recognising same- or other-race faces. Two phenomena have been observed: the “Other-Race Effect” (ORE), where people identify faces of individuals from own-race easier than other-race faces (e.g., [23–27]), and the “Other-Race Advantage” (ORA), where people are

faster to categorise other- compared to same-race faces [28, 29].

In the literature, several explanations have been proposed regarding the mechanisms that differentiate the ORA from the ORE. According to the model of Valentine and colleagues [30], faces are encoded as nodes in an n -dimensional space in which the distance between two nodes is inversely related to their subjective similarity. The classification of a face results from the sum of the total activation of all neighbouring nodes, and increasing the total activation of nodes in a given group leads to faster classification. Thus, the higher homogeneity of other-race faces compared to own-race faces in representational space could lead to an easier spread of activation among other-race faces. This phenomenon could explain the faster and more accurate classification of other-race faces than own-race faces according to race (ORA) and the slower and less accurate identification of individual other-race faces than own-race faces (ORE).

In addition, the literature has shown constraining results in neurocognitive responses to these phenomena, which will be discussed in the following sections.

Early perceptual processing

Face detection and recognition is a fast and automatic feature in humans. Thus several early latency ERP components (i.e., occurring before 200 ms) are implied in this process, such as P1, N170, P2 and N2 [28, 31, 32], where some studies also report findings with middle or late components such as N250 or P3 (e.g., [33–35]).

With regards to racial features detection, there are contradictory findings. Some studies did not find a modulation of face-specific N170 response to same- or other-race faces (e.g., [28, 31, 32]), whereas other studies found differences but in both directions (i.e., larger for other-race, or larger for own-race, [36, 37]).

The conflicting N170 responses across studies suggest that this ERP response is sensitive to the context, material and demands of the experiment task. Further investigation would be useful to clarify the interaction between these elements and the effects on the N170 response.

Studies conducted on the ORE frequently manipulate facial features such as shape and colour to explore how processing of SR and OR faces occurs. Brebner et al. [24] observed an increase in N170 amplitude for OR faces linked to changes along a gradient of skin colour (from light-skinned to dark-skinned), with a significantly greater N170 amplitude in response to darker-skinned OR faces than lighter-skinned faces. Similar increases in amplitude have been observed for both the N170 [27, 33, 38–41], and other early components related to attentional modulation (i.e., N250, N200) and memory encoding (i.e., LPP) in response to the perception of OR compared

to SR faces [26, 33–35, 42–45]. Furthermore, these ERP responses (e.g. N200) were modulated by colour and facial structure and correlated with subsequent memory performance. In two studies [40, 41], the repetition suppression paradigm was used to directly test the effect of the other race. Using the EEG, they found stronger individuation processing for SR and stronger categorization processing for OR, indexed by the N170 and P2 components. Given this evidence, these results suggest that racial differences in skin colour and facial structure are detected during the encoding of unfamiliar faces and that the detection of these features may be a critical factor in the ability to remember them later [42].

Face recognition memory

Several studies have found better memory performance for same-race (SR) compared other-race faces [46–50]. In 2011 Lucas et al. investigated neural mechanisms of this memory bias by recording event-related potentials during a same-race (SR), and other-race (OR) faces memory task. They found late positive responses (LP) to memory recall for SR faces and, to a lesser extent, for OR faces. However, they also found greater N200 and P2 responses to OR faces with race-atypical features than faces that were race-stereotypical. According to the authors, these earlier potentials index the processing of unique facial information, crucial for remembering a face. This phenomenon, defined as Individuation, is usually greater for SR faces but lower and less reliable for OR faces. However, this recent evidence suggests that it is more frequent for OR faces that appear less stereotypical. In Herzmann et al. [51] analysed the neural responses during an associative-memory task, in which Caucasian and East Asian participants learned and recognised own-race and other-race faces. Participants showed a greater positivity to own-race faces, indicating more detailed retrieval of own-race faces and suggesting a better memory performance in recognising own-race faces than other race faces.

As previous studies suggest specific deficits in perception and subsequent memory of faces in older adults, some recent studies explored the impact of ageing on the processing and memorisation of young/old and White/Asian faces. Considering ORE, Komes et al. [25] tested older adults grouped according to their level of performance on tests of cognitive ability (high- and low-performing). Both groups showed a clear ORE for young faces from their racial ingroup (White), with significantly larger N170 amplitudes in response to memorising young Asian faces but not old Asian faces. However, group differences emerged independently of facial ethnicity: while low-performing subjects showed a lateralised N170 to the right, high-performing subjects showed a more

bilateral response. The authors argue that this surprising result may be due to a compensatory mechanism counteracting age-related decline in face perception enabling more efficient encoding into memory in high performers. This result is consistent with two previous studies [44, 52] and reveals the interaction between two perceptual outgroups (specifically Young and Asian) leading to a minimisation of the salience of ethnic facial cues with age and a greater perceptual marker for young (outgroup) faces.

More recently, Proverbio and De Gabriele, [53] conducted a similar study in which participants categorised three images as straight or inverted following the presentation of Caucasian and African faces of infants and adults. They observed larger N2 amplitudes originating in frontal structures in response to infants' faces compared to adults, with no effects attributable to race. However, an increase in P3 amplitude was observed in response to Caucasian faces compared to African faces, for adult faces only. These results confirm the visual primacy of age-related cues of facial perception over cues related to ethnicity or race.

Taken together, these results indicate that the ORE may be a generalised phenomenon which is applicable to other social outgroups beyond racial ethnicity, and that repeated exposure to a specific outgroup may lead to greater expertise in face recognition for individuals from that outgroup.

Facial cues for gender categorisation also appear to modulate the effect of the ORE by minimising the salience of face ethnicity [52], highlighting the task-specific nature of face recognition [54] and the secondary priority of facial cues for ethnicity compared to cues of gender, age and emotional expression. Wolff et al. [55] observed an own-gender bias for both male and female subjects, with increased performance in recognising familiar faces from one's gender compared to the opposite gender. This own-gender bias was also reflected in the electrophysiological markers associated with initial face processing, with significantly smaller P2 and N2 amplitudes in response to processing the subject's own face gender compared to faces of the opposite gender. In Wiese and Schweinberger study [56], male and female participants showed own-race biases, reflected in a larger N170 in memory performance. In fact, previous evidence show a correlation between between the magnitude of the N170 ethnicity effect during the learning phases and the magnitude of one's race bias at the time of the test (see [27]). In addition, only in female participants, better memory for own-gender faces as reflected in the N250 response. The authors concluded that since different face memory biases occur at temporally distinct stages of face processing, they would be based on different neurocognitive mechanisms.

In conclusion, the results of these studies suggest no differences in perceptual competence, but instead, the differences in ERPs reflect neural correlates of the effect of perceived social category membership on facial recognition memory. These findings align with socio-cognitive theories that highlight the significance of categorising a face as being either a member of a social ‘ingroup’ or an ‘outgroup’ [57, 58]. Consequently, bias in face memory could be caused by the fact that faces belonging to an “outgroup” are processed at a categorical level, whereas “ingroup” faces are individualised. Such individualisation would involve deeper processing and thus more accurate recognition memory.

It is interesting to note that context and changes in task demands can modulate the electrophysiological correlates of the ORE. Minor modulations by visual expertise have been observed in Asian subjects following previous experience with the majority ethnicity (White in the US), leading to smaller increases in N170 and N250 amplitudes when categorising outgroup faces which form part of the majority ethnicity [26, 42]. A more direct investigation of the effect of face expertise on N170 amplitude was conducted by Balas and Saville [59]. Essentially, they found that individuals recruited from small towns (500–1000 inhabitants) showed similar N170 amplitudes for both faces and objects compared to a more face-selective N170 amplitude for participants recruited from big cities (30,000 to 100,000 inhabitants). In addition, small-town participants were significantly less accurate at memorising previously-presented faces and discriminating between a novel and unfamiliar faces. Since ORE can be influenced by visual competence [60], the individual level of visual competence seems to be the main determinant of inefficient coding of facial stimuli in a racial outgroup.

The great variety and number of faces one may interact with during education influences the neurocognitive mechanisms underlying face processing. Poor exposure to faces belonging to other categories may decrease memory and a reduced ability to discriminate between members of these categories compared to faces belonging to one’s own categories.

In conclusion, these studies on the detection and recognition of same- or other-race faces suggest that the ORE is due to differences in the processing of other-race and same-race faces at an early stage of encoding associated with the visual familiarity of the face as defined by recognition memory paradigms. The inefficient encoding of other-race faces into long-term memory is primarily due to a lack of visual expertise and/or previous experience with faces from a different ethnicity to one’s own. While an electrophysiological component at approximately 250 ms is associated with race perception and may act as a perceptually-salient “tag” for further visual

processing [27], facial cues related to ethnicity (eg. skin color, eye width) can be selectively modulated by facial cues corresponding to gender and age. This evidence suggests that facial cues related to ethnicity might be secondary to more socially-relevant facial information.

Implicit bias

Implicit bias is the unconscious attribution of particular qualities to a member from a particular social group [61]. It is based on internal schemas originating from associative learning, where social categories are associated with affects or beliefs, primarily influenced by cultural and social context and that can modify an individual’s subjective evaluation, attitude and reaction toward social objects [62–64].

There are several experimental paradigms to investigate implicit bias, such as the Weapon Identification Task (WIT; [8]), the First-Person Shooter task (FPS; [65]), the Implicit Association Test (IAT; [66]), or the Affect Misattribution procedure [67]. These paradigms investigate implicit attitudes or stereotypes through tests that require quick responses in associating or categorising a series of stimuli. Rapid responses tend to hinder the use of cognitive control, and variations in response time and accuracy may indicate the existence of biases in perception and reaction to other-race stimuli compared to ingroup stimuli. Among those tasks, IAT, WIT, and FPS were adopted to study racial bias through ERP analysis. Thus, we will briefly describe those tasks and the main findings concerning the detected electrophysiological components.

Weapon identification task

The Weapons Identification task (WIT) is a computer-designed paradigm where the participant must categorise an object as a weapon or a tool as quickly as possible, after seeing a face from an unknown person from the same- or other-race (a priming stimulus). In two consecutive studies, Amodio et al. [3, 68] demonstrated that enhanced control, defined as increased response time and greater accuracy, to race-biased responses were positively correlated with the Error-Related Negativity potential (ERN) amplitude. This component is usually detected after unexpected stimuli, such as errors or false alarm, and might also represent a self-monitoring mechanism. More recently, Fleming et al. [69] observed that North-American military cadets fire more quickly and accurately at targets holding a gun when paired with a Middle-Eastern male body in traditional Arabic clothing. However, this increase in firing rate was accompanied by a correspondingly greater bias toward classifying objects in the hands of these targets as a gun compared to targets from other ethnic groups. Consistent with its

role in conflict monitoring, greater ERN amplitudes originating in the medial-frontal cortex were observed in the misclassification of tools associated with these targets. In conclusion, these studies showed that larger ERNs to racial responses correspond to higher levels of control during the task suggesting that race-based responses can be made despite activation of neural systems designed to detect bias and recruit controlled processing.

First-person shooter task

Other classical experiment is the First-Person Shooter task (FPST; [65] which is composed of a series of trials viewing different background images in which Black or White males may appear, holding a gun or a neutral object such as a wallet or cell phone. During the FPST, participants pretend to be in the police office dilemma where they must press a specific button to shoot armed targets, which indicate social threat, and press a different button not to shoot unarmed targets. Correll, Urland and Ito [70] found significantly higher P2 amplitudes in Caucasian participants in response to Black armed targets than White armed or unarmed targets. In contrast, they found larger N2 amplitudes for unarmed targets compared to armed targets and White targets compared to Black targets. These findings demonstrate that these components show a high degree of specificity: a greater P2 amplitudes seem to be associated with alertness while a higher N2 amplitude is associated with a greater inhibition. The data suggest that threat perception and conflict detection play a crucial role. According to the authors, this finding would suggest that factors that modulate the relationship between threat and race in the real world could impact racial bias.

Additionally, Xu and Inzlicht, [71] investigated the association between neural (ERN and Pe) and behavioural responses to errors in a shooting game. Participants completed a Shooter go/no-go task, which required shooting armed targets using a game gun and avoiding shooting innocent non-targets. The results showed that shooting errors elicited greater ERN and Pe amplitudes than correct shooting and that greater amplitudes of these neural responses were correlated with more accurate behavioural performance.

Overall, these studies suggest that it is possible to obtain online measures of brain response to shooting responses and that neural responses to shooting are predictive of response bias.

Implicit association test

Finally, the Implicit Association Test (IAT [66]); a reaction time-based measure of implicit social attitudes [72], has been one of the benchmarks for research into racial bias. Research participants may not always want to or be

capable of reporting their true attitudes about socially sensitive issues. For these reasons, implicit measures that could reveal the types and strengths of evaluative associations without reliance on self-report promise to significantly advance the scientific study of attitudes.

A primary concern with most measures based on self-report, and even measures based on observable behaviour, is that participants can control their responses to obscure underlying evaluations. Thus, the ideal for an implicit measure is that it be structured to avoid the influence of control-related processes, whereby responses are based purely on automatically activated evaluations and associations.

In behavioural studies, the IAT has been employed to compare the reaction times in response to associations between stimuli and attributes, such as Black or White faces and positive- or negative-valence words. Studies employing an IAT-EEG design have explored how automatic racial bias can modulate the direction of attention, cognitive control and decision-making over time. Therefore, this instrument has been of great importance and utility in understanding racial bias processing.

The most common ERPs studied during implicit association processing are N170, N200, P300 and MFN. Ofan, Rubin and Amodio [73] explored anxiety effects on the neural processing of faces versus non-faces during a priming task, examining modulations on N170. They found that participants with higher social anxiety exhibited greater processing, reflected in a larger N170, of Black faces compared to White faces just in public conditions, such as monitoring by the experimenter. Additionally, higher N200, which often reflects monitoring of response conflict [74, 75], was observed after presenting uncommon associations stimuli, such as Black faces with Positive attributes for White participants. This higher amplitude reflects a cognitive demand to inhibit automatic responses, as the association between outgroup faces with Negative attributes [72].

In another study, Coates and Campbell [72] found that common associations on IAT resulted in a single P300 peak, while uncommon associations resulted in a sequence of two peaks. Thus, the authors argue that the reduced P300 amplitudes in response to uncommon associations may be related to a lack of certainty with respect to the decision.

In 2015, Hilgard and colleagues conducted a study investigating congruency and task switching effects in the IAT. During the IAT, the participants must switch between semantically categorising attitude objects on some trials (e.g. classifying names as Black or White) and evaluating words (as good or bad) on other trials. The authors found higher MFN amplitude on IAT blocks in response to uncommon associations (e.g. Black/Positive

and White/Negative) in this study. This response was more negative for incongruent than congruent trials but more positive for switch than for no-switch trials, suggesting separable control processes are engaged by these two factors.

Overall, these studies showed that it is easier to categorise familiar stimuli such as faces from our own social or ethnic ingroup, and we may hold an implicit bias to associate positive attributes with our ingroup and negative attributes with members of other outgroups. These give rise to the significant effects commonly observed in the IAT. Based on data collected with the IAT, implicit preferences for a specific ethnic group seem to form part of human social cognition and are relatively stable over time and shared among members of that ingroup [76]. In sum, since testing implicit bias is associated with a large degree of individual variability and different levels of self-control, the methodological standardisation of investigation in this area is necessary to understand small levels of inter-subject variability in IAT-ERP studies accessing different ethnic samples. Unbiased participants may not show any differences in cognitive effort regarding their electrophysiological markers during the classification of common and uncommon associations, but we have no clear data of how unbiased subjects process stimulus associations differently. For example, it is unclear whether they have an automatic association specific to a certain ethnic outgroup or have a greater level of cognitive control. Considering all of these factors, it would be enlightening for IAT to be applied in samples of, for example, children from interracial couples or from clinical populations that are extremely sociable and have impairments in cognitive control, such as Williams Syndrome.

The present section highlighted the importance of studying the temporal organisation of the mental processes involved in implicit prejudice. Furthermore, the findings mentioned above demonstrate that electrophysiological components constitute an effective means of objectively assessing automatic preferences and personal prejudices.

The study of implicit prejudice as mapped by ERPs may help us understand conflict-monitoring processes, which form an essential function for regulating much of our everyday behaviour.

Empathy to pain

Our understanding of the neural mechanisms involved in processing aversive stimuli such as pain has been long-established via physiological investigations [77–79]. More recently, experiments in cognitive psychology have shown that individuals report a heightened level of discomfort, either when directly stimulated with an aversive or painful stimulus or when observing the application

of an aversive stimulus to others [80–84]. The perception of other individuals in painful or aversive conditions typically results in high levels of aversive self-report [81]. More recently, advances in social neuroscience have demonstrated that under both conditions of self-exposure to painful stimuli and the perception of pain stimulation in others, a common network of cortical and subcortical neural structures are recruited in the processing of aversive stimuli (see [83, 84]). Within this network, the principal cortical structures involved in aversive stimuli perception includes the Prefrontal Cortex (PFC) and the Anterior Cingulate Cortex (ACC), and the main subcortical structures include the anterior Insula (aI) and amygdala [84], also observed for different kinds of perceived pain [85, 86]. Furthermore, some studies have shown that features such as social context [80] and cultural influence [82] can modulate an individual's empathic responses to the pain of others. So arises the question: are empathic response levels related to another individual's pain perception or determined by ethnic differences? In the context of the current review: how can the ethnic differences observed in individuals influence the electrophysiological processing of a pain-related situation?

It has been observed that empathic neural responses to perceived pain are stronger for ingroup members than for outgroup members. In other words, social relationships influence empathy in people [87]. This phenomenon is known as racial bias in empathy (RBE).

A recent meta-analysis [88] identifies several ERP components that would be directly associated with the observation of pain in others. These components are the centre-parietal P3 component and the late positive potential. These components have been observed in studies investigating the interaction between empathy and racial bias. However, in the studies reviewed here, the early N1 and N2 components not always associated with the observation of vicarious pain, were also observed. As the authors of previous studies also note, these early components could be associated with the first ethnic group identification that precedes the empathic response.

Racial bias in empathy in static tasks

In 2012, Sheng & Han conducted a study to test whether perceiving a person of another race as a symbol of a racial group decreases empathy for that person. Initially, it has been observed a higher amplitude of P2 in response to Asian faces (ingroup) expressing pain than White (outgroup) painful faces. However, subsequent experiments showed that paying attention to the observed individual's feelings of pain and including individuals of other races in one's team increased neural responses to expressions of pain in the faces of other races, thus suppressing RBE.

This finding suggests that manipulations of intergroup relations may decrease RBE-related brain activity.

More recently, Sheng et al. [89] found similar effects: higher P2 amplitudes in Chinese subjects in response to Chinese faces than White faces, even while recording during a perceptual and not a race- or empathy-based task. Participants were instructed to classify each face for their orientation and not their ethnicity (Chinese or White) or expression (painful or neutral). This study confirms Sheng and Han's [87] finding that there is a greater response to pain to ingroup faces than outgroup faces.

In 2019, Li and Han tested the pain sensitivity of own-race and other-race faces in images with different pixel-scale definitions. It was observed that fewer painful pixels were necessary to perceive the pain expressions of same-race faces compared to faces of other races and that P2 amplitude increased faster in response to increasing the number of pain pixels of same-race faces compared to faces of other races. These results support the evidence of increased sensitivity of neural responses associated with subtle variations in pain expressions of same-race faces compared to other-race faces.

In contrast to the paradigm used by Sheng and Han [87], in the study run by Contreras-Huerta et al. [90], no physical clues were associated with group members, such that the group and outgroup members could not be distinguished solely by low-level visual features. The results showed no specific electrophysiological modulation due to ingroup manipulation for the N2 or P3 components. However, they found an increase in N1 amplitude in response to the observation of ingroup (White) individuals receiving painful stimulation compared to non-painful stimulation. Higher P3 amplitudes were elicited in response to the perception of individuals receiving painful stimulation, regardless of the ethnic group (White or Asian) at a later stage of processing. These results support a model of empathy for pain that consists of early, automatic empathic responses to people from their own race, as indexed by N1, and a subsequent general salience response to painful stimuli that does not differentiate between races associated with the P3 response.

Despite differences in early empathic responses in the two groups due to methodological differences [90], taken together, these studies demonstrate the influence of the perception of individuals in the group, leading to an increase in the magnitude of early ERP components in empathic responses to pain.

Sheng et al. [91] investigated whether the subsequent presentation of faces of different ethnic groups expressing pain can modulate electrophysiological response. Accordingly, Chinese participants were presented with adaptor faces (with pain or neutral expressions) and target faces (with only pain expression) in rapid succession.

The face pair could be either congruent (both faces had the same ethnicity) or incongruent (ex. White and Asian). Corroborating the previously cited studies, the authors found increased P2 amplitude relative to ingroup faces compared to outgroup faces. However, this effect was only observed when the face sequences were congruent with face ethnicity. The increase in P2/N2 occurred only when an adaptor and a target were of the same race, and these effects were observed in both Chinese and Caucasian participants. This finding suggests that perceiving pain expressions from different races may engage distinct neurocognitive processing.

Racial bias in empathy in dynamic tasks

Paradigms involving static tasks, such as observing images of faces expressing pain, provide important evidence and scientific insights into the influence of ethnic differences on the processing of a pain-related situation. However, everyday social interactions occur dynamically, leaving the doubt: would the same electrophysiological patterns found with the presentation of images be found with videos presenting faces of pain?

Sessa et al. [92] tested white participants performing a pain assessment task based on the videos of the faces of Whites and African individuals being touched by a needle or cotton swab. They observed an increase in the amplitude of the ERP components N2 and N3 in response to the processing of White faces in painful situations compared to situations without pain. This effect was not observed for African faces. In addition, no differences between ethnic groups were observed at later stages of processing (P3 response). Furthermore, they observed a pole of activation over the left medial frontal gyrus (MFG) for the P3 component in response to the observation of ingroup pain, while the source of P3 activation in response to the observation of outgroup stimuli was located in the TPJ. Thus, in line with the studies mentioned above, the authors concluded that earlier processing stages are more susceptible to modulation by ethnic group, as observed by the modulation of N2/N3. Later stages of processing, while still sensitive to race-related modulation, represent a greater degree of control over higher cognitive functions such as mentalising and cognitive reprocessing, as observed at P3. It is important to note that, despite using different techniques and tasks, Sheng et al. [93] findings showing TPJ recruitment during the tasks of pain discrimination are consistent with the results from Sessa et al. [92]. Together, they highlight the importance of a region in the posterior temporal cortex, including the TPJ, in the cognitive processing or inference about the mental state of others, also known as Theory of Mind (TOM, [94]).

Another interesting issue involving these studies is the relationship between the electrophysiological response of pain empathy processing and levels of empathy. These levels are typically measured through behavioural scales, such as the Interpersonal Reactivity Scale (IRI) initially designed by Davis [95]. Sessa et al. [92] found a positive correlation between white individuals' performance on the IRI subtest "Empathic Concern" and N2/N3 amplitude when observing ingroup faces in pain situations. However, no significant correlation was observed for outgroup faces in pain situations. This result suggests that empathy levels can be predicted from the electrophysiological responses to other individuals in painful situations, but only at the early stages of processing and only for individuals from the same ethnic group.

Similarly, a study by Han et al. [96] sought to test the influence of different oral mimicry (contracted and relaxed oral musculature) on neurocognitive processing during the observation of facial imagery expressing pain in an ingroup (Asian) and outgroup (White) individuals. This design builds on classic studies testing the influence of mimicry and emotional embodiment on subsequent cognitive and emotional processing (see [97] for a review). Consistent with the literature [87, 90, 91, 93], greater amplitudes of N1, N2 and P3 were observed for faces expressing pain compared to emotionally neutral faces, with the greatest amplitudes observed in response to faces belonging to the same ethnic group.

Finally, several behavioural studies [98–100] have shown that individuals typically tend to conform to the opinion of other ingroup members in response to anxiety generated by priming to death or threat stimuli (an effect known as mortality salience or terror management). In support of this hypothesis, a Chinese study used a mortality salience induction method to investigate the effects of pain empathy in individuals from different ethnic groups [99]. In the early stages of processing, they found higher P2 amplitudes only when observing Asian individuals (ingroup) in painful situations regardless of the experimental condition (induction to mortality salience compared to the control condition). However, in subsequent processing steps, the induction of mortality salience led to an increase in the amplitude of the P3 component in response to the observation of ingroup individuals in painful situations. These results demonstrate that induction of mortality salience leads to increased empathy responses towards ingroup individuals and that this type of induction could also be involved in the electrophysiological modulation of final phase processing related to pain empathy.

Overall, these results indicate that the perception of pain expressions of different races may elicit a greater empathic response to ingroup faces than to outgroup

faces. This response was observed in both early (N2, P2) and later processing stages (P3) of pain information. Furthermore, the P3 potential was more influenced by cognitive processes, such as the theory of mind, related to the activity of regions such as the TPJ and subjective measures such as the Interpersonal Reactivity Scale.

Visually witnessing ethnic individuals from the group in pain-related situations leads to more rapid recruitment of brain areas related to empathy and TOM than observing individuals from the outgroup, which may facilitate a greater degree of empathic response (Table 1).

The impact of racial bias on mental and physical well-being

Racial prejudices observed in the laboratory using the EEG technique have serious effects on our society and can also lead to alterations in psychophysical well-being. In this section we will present some studies in which, using techniques other than EEG, the effects on the health of individuals affected by racial biases were investigated. We believe that this section, although differing from the central corpus of the review in terms of techniques, is complementary to the previous sections as it provides a measure of the actual risks posed by this observable phenomenon in the laboratory on the health of ethnic minority individuals. In fact, a growing body of literature has emphasised that racial discrimination is an important factor potentially causing health problems in ethnic minorities. Berger and Sarnyai, in a recent review [103], provided evidence on the hormonal and neural effects of racial discrimination. It has been observed that chronically elevated cortisol levels and a dysregulated hypothalamic–pituitary–adrenal (HPA) axis mediate the effects of racial discrimination on disease. Furthermore, it was found that racial discrimination can impair prefrontal cortex (PFC) function, showing significant similarities to chronic social stress. In the same year, Paradies et al. [104] conducted a meta-analysis in which they reviewed data from 293 studies published between 1983 and 2013 and examined the relationship between reported racism and mental and physical health outcomes. Overall, the results showed that racism is associated with poorer physical and mental health, including depression, anxiety and psychological distress.

More recently, Korous et al. [105] conducted a meta-analysis on racial discrimination and cortisol levels. The results suggested that the association between racial discrimination and cortisol is complex and that effect sizes varied depending on the method used. Results from another meta-analysis [106] noted that adults' reported experiences of discrimination are negatively related to mental health and physical health indicators, such as pre-clinical indicators of disease, treatment utilisation and adherence to medical regimens. Emerging evidence also

Table 1 Summary of studies included in the systematic review

Topic	Subtopic	ERP	Function	Studies
Face recognition	Early perceptual processing	N170	Increase in N170 amplitudes for OR	Brebner et al. [24]
		N200	Higher N200 amplitudes to the perception/categorisation of OR compared to SR faces	Ito and Urland [43], Herrmann et al. [35], Stahl et al. [26], Herzmann et al. [42], Wiese [44]
		N250	Higher N250 amplitudes response to the perception/categorisation of OR compared to SR faces	Ito and Urland [43], Herrmann et al. [35], Herzmann et al. [42], Wiese [44]
		LPP	Higher LPP amplitudes to the perception/categorisation of OR compared to SR faces	Ito and Urland [43], Wiese [44]
	Face recognition memory	N200	Higher N200 to OR faces with race-atypical features than faces that were race-stereotypical	Lucas et al. [101]
		P2	Higher P2 to OR faces with race-atypical features than faces that were race-stereotypical	Lucas et al. [101]
Implicit bias	Weapon identification task	ERN	Higher ERN amplitude is positively related to enhanced control during WIT	[3, 68]
			Higher ERN amplitudes observed for misclassification of tools associated with these targets	Fleming et al. [69]
	First-person shooter task	P2	Higher P2 amplitudes in Caucasian in response to Black armed targets than White armed or unarmed targets	Correll, Urland and Ito [70]
			Higher N2 amplitudes for unarmed targets compared to armed targets and White targets compared to Black targets	Correll, Urland and Ito [70]
			Higher ERN amplitudes related to shooting errors compared to correct shooting	Xu and Inzlicht, [71]
	Implicit association test (IAT)	N170	Higher N170 related to higher social anxiety faced with Black faces compared to White faces	Ofan, Rubin and Amodio [73]
			Higher N200 when presenting uncommon associations stimuli	Ofan, Rubin and Amodio [73]
			Single P300 peak observed for common associations, while two peaks observed for uncommon associations	Coates and Campbell [72]
			Higher MFN amplitude in response to uncommon associations	Hilgard et al. [13]
			Higher N1 amplitudes when facing ingroup pain faces, compared to outgroup faces	Contreras-Huerta et al. [90]
Empathy to Pain	Racial bias in empathy in static tasks	P2	Higher P2 amplitudes when facing ingroup pain faces, compared to outgroup faces	Li and Han [102], Sheng and Han [87], Sheng et al. [91], Sheng et al. [89]
			Higher P3 amplitudes when facing ingroup pain faces, compared to outgroup faces	Contreras-Huerta et al. [90]
			Higher N2 amplitudes when facing ingroup pain faces, compared to outgroup faces	Sessa et al. [92]
	Racial bias in empathy in dynamic tasks	N3	Higher N3 amplitudes when facing ingroup pain faces, compared to outgroup faces	Sessa et al. [92]

WIT—"Weapon Identification Task"

ERN—"Error-related negativity potential"

IAT—"Implicit Association Test"

suggests that discrimination may affect the health of children and adolescents.

Overall, these studies suggest that discriminatory experiences can, like other stressors, trigger physiological responses that negatively affect the maintenance of body homeostasis. In conclusion, exposure to discrimination is an important risk factor for the disease, affecting racial and ethnic minorities significantly.

Can racial prejudice be reduced?

Recently, a few studies have tried to verify whether racial prejudices can be reduced, which could be of great relevance to society. In Pech and Caspar [107] conducted a study to investigate whether a video game designed to reduce prejudice against minorities in a fictional society has the potential to decrease prejudice against minorities in real life as well. A positive effect of the test game compared to the control game on the attenuation of prejudice against an out-group individual was observed. These results appear very promising as they support the evidence that the use of fictional characters in video games can induce positive changes towards non-fictional individuals.

In another study [108], it was observed that racial prejudice to others' pain is modulated by experience and contact with individuals of a different ethnicity, suggesting that experience reduces racial prejudice. Another work [109] showed that another effective technique to reduce racial prejudice could be group therapy, which, by addressing emotions related to racial prejudice through practical techniques, offers a rich context for dealing with this phenomenon.

Finally, a general but, in our opinion, fundamental suggestion comes from the review of Dotson and Duarte [110]. Here, the authors emphasise the urgency of taking demographic characteristics such as race and ethnicity into account when investigating the neural basis of human cognition. They argue that, although challenging, greater diversity in cognitive neuroscience research is needed to improve reproducibility and to meet the treatment needs of a diverse population.

Conclusions

The review and study of the ERP components associated with implicit prejudice bias, empathy for pain and recognition of ethnicity via face processing pinpoint the distinct stages involved in the perception and emotional response towards social outgroups. The rationale for the selection of ERP-EEG design as the inclusion criterion for studies in this review is that it offers an objective lens for understanding results frequently drawn from a heterogeneous set of methodologies for studying prejudice, ranging from anthropology and

social psychology to cognitive neuroscience. In other words, ERPs allows us to follow the time course of how an observer processes same- or other-race stimuli and forms unfavourable evaluations.

Investigating specific ERP components associated with implicit prejudice allows us to understand the impact of racial bias on attentional orientation, specifically the delay or errors directly attributed to the difficulty in inhibiting implicit or explicit prejudiced responses. This difficulty is due to conflicts on a cognitive and social level, which tend to modulate the participant's task engagement and his/her future responses.

This review provides an overview of recent studies investigating racial bias using an event-related potential EEG collection and analysis design. The data gathered here focus on the widespread use and application of univariate analysis of evoked potentials for the study of racial bias. According to Ito et al. [111] the use of ERP design in EEG experimentation allows for a greater understanding of how social contexts are perceived and associated with specific stereotypes and behavioural predictions. Thus, ERPs evaluate changes in the neural circuitry and its activity at the level of automatic and controlled processes for the applied tasks. While prejudice is a complex social phenomenon that requires several levels of analysis and a variety of different methods of investigation, EEG and ERP analysis is a relatively modern and efficient tool for testing aspects of social cognition. We hope the combined use of ERP-EEG and both implicit and explicit behavioural measures will yield further insights into our understanding of this phenomenon by facilitating the investigation of more subtle forms of prejudice and of the cognitive processes recruited during social evaluation.

Abbreviations

ACC	Anterior cingulate cortex
EEG	Electroencephalography
ERN	Error-related negativity
Pe	Error-related positivity
ERP	Event-related potential
FPST	First-person shooter task
fMRI	Functional magnetic resonance imaging
IAT	Implicit association test
al	Anterior insula
IRI	Interpersonal reactivity scale
LPP	Late positive potential
MFG	Medial frontal gyrus
MFN	Medial frontal negativity
OR	Other-race
ORE	Other-race effect
PFC	Prefrontal cortex
SR	Same-race
SPCN	Sustained posterior contralateral negativity
TPJ	Temporo-parietal junction
TOM	Theory of mind
VWM	Visual working memory
WIT	Weapon identification test

Acknowledgements

We would like to thank Beatriz Sanchez and Fabiana Felississimo for the initial review and selection of relevant articles.

Author contributions

All authors contributed to manuscript writing and approved the final version of the manuscript.

Funding

GGR is supported by a postdoctoral research Grant #2019/26665-5, FAPESP. LMM is supported by a postdoctoral research Grant #2021/05897-5, São Paulo Research Foundation (FAPESP). WEC is supported by a postdoctoral research fellowship: nº 2017/12752-8, from the São Paulo Research Foundation (FAPESP). JHE was supported by a PhD grant: nº 2016/17242-5, from the São Paulo Research Foundation (FAPESP). PSB holds a fellowship from the Brazilian National Research Council—CNPq (304164/2012-7). Paulo S. Boggio is supported by: 1. CNPq researcher fellowship [grant n. 309905/2019-2]], 2. CAPES - PRINT (Programa Institucional de Internacionalização [grant n. 88887.310255/2018 - 00]), 3. CAPES - PROEX (grant n. 04236/2021), and 4. CNPq - INCT (National Institute of Science and Technology on Social and Affective Neuroscience, grant n. 406463/2022-0). Mirella Manfredi was partially funded by the Jacobs Foundation.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors have no competing interests.

Received: 29 June 2022 Accepted: 14 March 2023

Published online: 27 March 2023

References

- Powell JA. Structural racism: building upon the insights of John Calmore. *NCL Rev.* 2007;86:791.
- Embrick DG, Henricks K. Discursive colorlines at work: How epithets and stereotypes are racially unequal. *Symb Interact.* 2013;36(2):197–215.
- Amodio DM, Devine PG. Stereotyping and evaluation in implicit race bias: evidence for independent constructs and unique effects on behavior. *J Pers Soc Psychol.* 2006;91(4):652.
- Lieberman Z, Woodward AL, Kinzler KD. The origins of social categorization. *Trends Cogn Sci.* 2017;21(7):556–68.
- Ito TA, Bartholow BD. The neural correlates of race. *Trends Cogn Sci.* 2009;13(12):524–31.
- Denson N, Chang MJ. Racial diversity matters: The impact of diversity-related student engagement and institutional context. *Am Educ Res J.* 2009;46(2):322–53. <https://doi.org/10.3102/0002831208323278>.
- Greenwald AG, Poehlman TA, Uhlmann EL, Banaji MR. Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *J Personal Soc Psychol.* 2009;97(1):17.
- Payne BK. Prejudice and perception: the role of automatic and controlled processes in misperceiving a weapon. *J Pers Soc Psychol.* 2001;81(2):181.
- Baggio G. *Meaning in the brain.* MIT Press; 2018.
- Luck SJ, Kappenman ES (2012) ERP components and selective attention
- Bruce V, Young A. *In the eye of the beholder: the science of face perception.* Oxford University Press; 1998.
- Codispoti M, Ferrari V, Bradley MM. Repetitive picture processing: automatic and cortical correlates. *Brain Res.* 2006;1068(1):213–20.
- Hilgard J, Bartholow BD, Dickter CL, Blanton H. Characterizing switching and congruency effects in the implicit association test as reactive and proactive cognitive control. *Soc Cogn Affect Neurosci.* 2015;10(3):381–8.
- Duncan-Johnson CC, Donchin E. On quantifying surprise: the variation of event-related potentials with subjective probability. *Psychophysiology.* 1977;14(5):456–67.
- Manfredi M, Adorni R, Proverbio A. Why do we laugh at misfortunes? An electrophysiological exploration of comic situation processing. *Neuropsychologia.* 2014;61(1):324–34. <https://doi.org/10.1016/j.neuropsychologia.2014.06.029>.
- Proverbio AM, Riva F, Paganelli L, Cappa SF, Canessa N, Perani D, Zani A. Neural coding of cooperative versus affective human interactions: 150 ms to code the action's purpose. *PLoS ONE.* 2011;6(7):e22026. <https://doi.org/10.1371/journal.pone.0022026>.
- Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science.* 1980. <https://doi.org/10.1126/science.7350657>.
- Kutas M, Federmeier KD. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu Rev Psychol.* 2011;62:621–47. <https://doi.org/10.1146/annurev.psych.093008.131123>.
- Ibáñez A, Toro P, Cornejo C, Urquina H, Hurquina H, Manes F, Weisbrod M, Schröder J. High contextual sensitivity of metaphorical expressions and gesture blending: a video event-related potential design. *Psychiatry Res.* 2011;191(1):68–75. <https://doi.org/10.1016/j.psychres.2010.08.008>.
- Sitnikova T, Kuperberg G, Holcomb PJ. Semantic integration in videos of real-world events: an electrophysiological investigation. *Psychophysiology.* 2003;40(1):160–4.
- Van Petten C, Luka BJ. Prediction during LANGUAGE COMPREHENSION: BENEFITS, COSTS, and ERP components. *Int J Psychophysiol Off J Int Organ Psychophysiol.* 2012;83(2):176–90. <https://doi.org/10.1016/j.ijpsycho.2011.09.015>.
- Quante L, Bölte J, Zwitserlood P. Dissociating predictability, plausibility and possibility of sentence continuations in reading: evidence from late-positivity ERPs. *PeerJ.* 2018. <https://doi.org/10.7717/peerj.5717>.
- Anzures G, Mildort M. Do perceptual expertise and implicit racial bias predict early face-sensitive ERP responses? *Brain Cogn.* 2021. <https://doi.org/10.1016/j.bandc.2020.105671>.
- Brebner JL, Krigolson O, Handy TC, Quadflieg S, Turk DJ. The importance of skin color and facial structure in perceiving and remembering others: an electrophysiological study. *Brain Res.* 2011;1388:123–33. <https://doi.org/10.1016/j.brainres.2011.02.090>.
- Komes J, Schweinberger SR, Wiese H. Preserved fine-tuning of face perception and memory: evidence from the own-race bias in high- and low-performing older adults. *Front Aging Neurosci.* 2014;6:1–10. <https://doi.org/10.3389/fnagi.2014.00060>.
- Stahl J, Wiese H, Schweinberger SR. Learning task affects ERP-correlates of the own-race bias, but not recognition memory performance. *Neuropsychologia.* 2010;48(7):2027–40. <https://doi.org/10.1016/j.neuropsychologia.2010.03.024>.
- Wiese H, Kaufmann JM, Schweinberger SR. The neural signature of the own-race bias: evidence from event-related potentials. *Cereb Cortex.* 2014;24(3):826–35. <https://doi.org/10.1093/cercor/bhs369>.
- Caldara R, Rossion B, Bovet P, Hauert C-A. Event-related potentials and time course of the 'other-race' face classification advantage. *NeuroReport.* 2004;15(5):905–10.
- Sun G, Zhang G, Yang Y, Bentin S, Zhao L. Mapping the time course of other-race face classification advantage: a cross-race ERP study. *Brain Topogr.* 2014;27(5):663–71.
- Valentine T. A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Q J Exper Psychol Sect A.* 1991;43(2):161–204.
- Caldara R, Thut G, Servois P, Michel CM, Bovet P, Renault B. Face versus non-face object perception and the 'other-race' effect: a spatio-temporal event-related potential study. *Clin Neurophysiol.* 2003;114(3):515–28.
- Tanaka JW, Pierce LJ. The neural plasticity of other-race face recognition. *Cogn Affect Behav Neurosci.* 2009;9(1):122–31.

33. Balas B, Nelson CA. The role of face shape and pigmentation in other-race face perception: an electrophysiological study. *Neuropsychologia*. 2010;48(2):498–506. <https://doi.org/10.1016/j.neuropsychologia.2009.10.007>.
34. Dickter CL, Bartholow BD. Racial ingroup and outgroup attention biases revealed by event-related brain potentials. *Soc Cogn Affect Neurosci*. 2007;2(3):189–98. <https://doi.org/10.1093/scan/nsm012>.
35. Herrmann MJ, Schreppel T, Jäger D, Koehler S, Ehliis AC, Fallgatter AJ. The other-race effect for face perception: an event-related potential study. *J Neural Transm*. 2007;114(7):951–7. <https://doi.org/10.1007/s00702-007-0624-9>.
36. Stahl J, Wiese H, Schweinberger SR. Expertise and own-race bias in face processing: an event-related potential study. *NeuroReport*. 2008;19(5):583–7.
37. Walker PM, Silvert L, Hewstone M, Nobre AC. Social contact and other-race face processing in the human brain. *Soc Cogn Affect Neurosci*. 2008;3(1):16–25.
38. Caharel S, Montalan B, Fromager E, Bernard C, Lalonde R, Mohamed R. Other-race and inversion effects during the structural encoding stage of face processing in a race categorization task: an event-related brain potential study. *Int J Psychophysiol*. 2011;79(2):266–71. <https://doi.org/10.1016/j.ijpsycho.2010.10.018>.
39. van Elk M, Duizer M, Sligte I, van Schie H. Transcranial direct current stimulation of the right temporoparietal junction impairs third-person perspective taking. *Cogn Affect Behav Neurosci*. 2017;17(1):9–23. <https://doi.org/10.3758/s13415-016-0462-z>.
40. Vizioli L, Rousselet GA, Caldara R. Neural repetition suppression to identity is abolished by other-race faces. *Proc Natl Acad Sci USA*. 2010;107(46):20081–6. <https://doi.org/10.1073/pnas.1005751107>.
41. Zhou Y, Gao T, Zhang T, Li W, Wu T, Han X, Han S. Neural dynamics of racial categorization predicts racial bias in face recognition and altruism. *Nat Hum Behav*. 2020;4(1):69–87. <https://doi.org/10.1038/s41562-019-0743-y>.
42. Herzmann G, Willenbockel V, Tanaka JW, Curran T. The neural correlates of memory encoding and recognition for own-race and other-race faces. *Neuropsychologia*. 2011;49(11):3103–15. <https://doi.org/10.1016/j.neuropsychologia.2011.07.019>.
43. Ito TA, Urland GR. Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *J Pers Soc Psychol*. 2003;85(4):616–26. <https://doi.org/10.1037/0022-3514.85.4.616>.
44. Wiese H. The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biol Psychol*. 2012;89(1):137–47. <https://doi.org/10.1016/j.biopsycho.2011.10.002>.
45. Zheng X, Segalowitz SJ. Putting a face in its place: In- and out-group membership alters the N170 response. *Soc Cogn Affect Neurosci*. 2014;9(7):961–8. <https://doi.org/10.1093/scan/nst069>.
46. Bernstein MJ, Young SG, Hugenberg K. The cross-category effect: mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychol Sci*. 2007;18(8):706–12.
47. MacLin OH, Malpass RS. Racial categorization of faces: the ambiguous race face effect. *Psychol Public Policy Law*. 2001;7(1):98.
48. Shriver ER, Young SG, Hugenberg K, Bernstein MJ, Lanter JR. Class, race, and the face: social context modulates the cross-race effect in face recognition. *Pers Soc Psychol Bull*. 2008;34(2):260–74.
49. Tüttenberg SC, Wiese H. Learning own- and other-race facial identities: Testing implicit recognition with event-related brain potentials. *Neuropsychologia*. 2019. <https://doi.org/10.1016/j.neuropsychologia.2019.107218>.
50. Tüttenberg SC, Wiese H. Recognising other-race faces is more effortful: The effect of individuation instructions on encoding-related ERP Dm effects. *Biol Psychol*. 2021. <https://doi.org/10.1016/j.biopsycho.2020.107992>.
51. Herzmann G, Minor G, Adkins M. Neural correlates of memory encoding and recognition for own-race and other-race faces in an associative-memory task. *Brain Res*. 2017;1655:194–203. <https://doi.org/10.1016/j.brainres.2016.10.028>.
52. Wallis J, Lipp OV, Vanman EJ. Face age and sex modulate the other-race effect in face recognition. *Atten Percept Psychophys*. 2012;74(8):1712–21. <https://doi.org/10.3758/s13414-012-0359-z>.
53. Proverbio AM, De Gabriele V. The other-race effect does not apply to infant faces: an ERP attentional study. *Neuropsychologia*. 2019;126:36–45. <https://doi.org/10.1016/j.neuropsychologia.2017.03.028>.
54. Ishai A. Let's face it: it's a cortical network. *Neuroimage*. 2008;40(2):415–9. <https://doi.org/10.1016/j.neuroimage.2007.10.040>.
55. Wolff N, Kemter K, Schweinberger SR, Wiese H. What drives social in-group biases in face recognition memory? ERP evidence from the own-gender bias. *Soc Cogn Affect Neurosci*. 2014;9(5):580–90. <https://doi.org/10.1093/scan/nst024>.
56. Wiese H, Schweinberger SR. Inequality between biases in face memory: event-related potentials reveal dissociable neural correlates of own-race and own-gender biases. *Cortex*. 2018;101:119–35. <https://doi.org/10.1016/j.cortex.2018.01.016>.
57. Hugenberg K, Young SG, Bernstein MJ, Sacco DF. The categorization-individuation model: an integrative account of the other-race recognition deficit. *Psychol Rev*. 2010;117(4):1168.
58. Sporer SL. Recognizing faces of other ethnic groups: an integration of theories. *Psychol Public Policy Law*. 2001;7(1):36.
59. Balas B, Saville A. N170 face specificity and face memory depend on hometown size. *Neuropsychologia*. 2015;69:211–7.
60. Haque A, Cottrell GW. Modeling the other-race advantage with PCA. In: *Proc. 27th Annu. Cogn. Sci. Conf.* 2005. pp. 899–904.
61. Celis LE, Mehrotra A, Vishnoi NK (2020) Interventions for ranking in the presence of implicit bias. In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*, 369–380
62. Greenwald AG, Banaji MR. Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychol Rev*. 1995;102(1):4.
63. Greenwald AG, Krieger LH. Implicit bias: scientific foundations. *Calif Law Rev*. 2006;94(4):945–67.
64. Pritlove C, Juando-Prats C, Ala-Leppilampi K, Parsons JA. The good, the bad, and the ugly of implicit bias. *The Lancet*. 2019;393(10171):502–4.
65. Correll J, Park B, Judd CM, Wittenbrink B. The police officer's dilemma: using ethnicity to disambiguate potentially threatening individuals. *J Pers Soc Psychol*. 2002;83(6):1314.
66. Greenwald AG, McGhee DE, Schwartz JLK. Measuring individual differences in implicit cognition: the implicit association test. *J Pers Soc Psychol*. 1998;74(6):1464.
67. Payne BK, Hall DL, Cameron CD, Bishara AJ. A process model of affect misattribution. *Pers Soc Psychol Bull*. 2010;36(10):1397–408.
68. Amodio DM, Harmon-Jones E, Devine PG, Curtin JJ, Hartley SL, Covert AE. Neural signals for the detection of unintentional race bias. *Psychol Sci*. 2004;15(2):88–93.
69. Fleming KK, Bandy CL, Kimble MO. Decisions to shoot in a weapon identification task: the influence of cultural stereotypes and perceived threat on false positive errors. *Soc Neurosci*. 2010;5(2):201–20.
70. Correll J, Urland GR, Ito TA. Event-related potentials and the decision to shoot: the role of threat perception and cognitive control. *J Exp Soc Psychol*. 2006;42(1):120–8.
71. Xu X, Inzlicht M. Neurophysiological responses to gun-shooting errors. *Int J Psychophysiol*. 2015;95(3):247–53.
72. Coates MA, Campbell KB. Event-related potential measures of processing during an implicit association test. *NeuroReport*. 2010;21(16):1029–33.
73. Ofan RH, Rubin N, Amodio DM. Situation-based social anxiety enhances the neural processing of faces: evidence from an intergroup context. *Soc Cogn Affect Neurosci*. 2014;9(8):1055–61.
74. Dickter CL, Bartholow BD. Ingroup categorization and response conflict: interactive effects of target race, flanker compatibility, and infrequency on N2 amplitude. *Psychophysiology*. 2010;47(3):596–601.
75. Liotti M, Woldorff MG, Perez R III, Mayberg HS. An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*. 2000;38(5):701–11.
76. Neto F (2015). O paradigma do Teste de Associação Implícita
77. Bishop GH, Landau WM. Evidence for a double peripheral pathway for pain. *Science*. 1958;128(3326):712–3.
78. Loeser JD, Melzack R. Pain: an overview. *The Lancet*. 1999;353(9164):1607–9.
79. Willis WD, Westlund KN. Neuroanatomy of the pain system and of the pathways that modulate pain. *J Clin Neurophysiol Offic Publ Am Electroencephalogr Soc*. 1997;14(1):2.

80. Akitsuki Y, Decety J. Social context and perceived agency affects empathy for pain: an event-related fMRI investigation. *Neuroimage*. 2009;47(2):722–34.
81. Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry*. 1994;25(1):49–59.
82. Cheon BK, Im D, Harada T, Kim J-S, Mathur VA, Scimeca JM, Parrish TB, Park HW, Chiao JY. Cultural influences on neural basis of intergroup empathy. *Neuroimage*. 2011;57(2):642–50.
83. Decety J, Lamm C. Human empathy through the lens of social neuroscience. *Sci World J*. 2006;6:1146–63.
84. Lamm C, Decety J, Singer T. Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*. 2011;54(3):2492–502.
85. Eisenberger NI, Lieberman MD. Why rejection hurts: a common neural alarm system for physical and social pain. *Trends Cogn Sci*. 2004;8(7):294–300.
86. Eisenberger NI, Lieberman MD, Williams KD. Does rejection hurt? An fMRI study of social exclusion. *Science*. 2003;302(5643):290–2.
87. Sheng F, Han S. Manipulations of cognitive strategies and intergroup relationships reduce the racial bias in empathic neural responses. *Neuroimage*. 2012;61(4):786–97.
88. Coll MP. Meta-analysis of ERP investigations of pain empathy underlines methodological issues in ERP research. *Soc Cogn Affect Neurosci*. 2018;13(10):1003–17. <https://doi.org/10.1093/scan/nsy072>.
89. Sheng F, Du N, Han S. Degraded perceptual and affective processing of racial out-groups: an electrophysiological approach. *Soc Neurosci*. 2017;12(4):479–87.
90. Contreras-Huerta LS, Hielscher E, Sherwell CS, Rens N, Cunnington R. Intergroup relationships do not reduce racial bias in empathic neural responses to pain. *Neuropsychologia*. 2014;64:263–70.
91. Sheng F, Han X, Han S. Dissociated neural representations of pain expressions of different races. *Cereb Cortex*. 2016;26(3):1221–33.
92. Sessa P, Meconi F, Castelli L, Dell'Acqua R. Taking one's time in feeling other-race pain: an event-related potential investigation on the time-course of cross-racial empathy. *Soc Cogn Affect Neurosci*. 2014;9(4):454–63.
93. Sheng F, Liu Q, Li H, Fang F, Han S. Task modulations of racial bias in neural responses to others' suffering. *Neuroimage*. 2014;88:263–70.
94. Saxe R, Kanwisher N. People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind." *Neuroimage*. 2003;19(4):1835–42. [https://doi.org/10.1016/S1053-8119\(03\)00230-1](https://doi.org/10.1016/S1053-8119(03)00230-1).
95. Davis MH (1980) A multidimensional approach to individual differences in empathy
96. Han X, Luo S, Han S. Embodied neural responses to others' suffering. *Cogn Neurosci*. 2016;7(1–4):114–27.
97. Price TF, Peterson CK, Harmon-Jones E. The emotive neuroscience of embodiment. *Motiv Emot*. 2012;36(1):27–37.
98. Giannakakis AE, Fritsche I. Social identities, group norms, and threat: on the malleability of ingroup bias. *Pers Soc Psychol Bull*. 2011;37(1):82–93.
99. Li X, Liu Y, Luo S, Wu B, Wu X, Han S. Mortality salience enhances racial in-group bias in empathic neural responses to others' suffering. *Neuroimage*. 2015;118:376–85.
100. Renkema LJ, Stapel DA, Van Yperen NW. Go with the flow: Conforming to others in the face of existential threat. *Eur J Soc Psychol*. 2008;38(4):747–56.
101. Lucas HD, Chiao JY, Paller KA. Why some faces won't be remembered: brain potentials illuminate successful versus unsuccessful encoding for same-race and other-race faces. *Front Hum Neurosci*. 2011;5:20. <https://doi.org/10.3389/fnhum.2011.00020>.
102. Li W, Han S. Behavioral and electrophysiological evidence for enhanced sensitivity to subtle variations of pain expressions of same-race than other-race faces. *Neuropsychologia*. 2019;129:302–9. <https://doi.org/10.1016/j.neuropsychologia.2019.04.008>.
103. Berger M, Sarnyai Z. "More than skin deep": stress neurobiology and mental health consequences of racial discrimination. *Stress*. 2015;18(1):1–10.
104. Paradies Y, Ben J, Denson N, Elias A, Priest N, Pieterse A, Gupta A, Kelaheer M, Gee G. Racism as a determinant of health: a systematic review and meta-analysis. *PLoS ONE*. 2015;10(9): e0138511.
105. Korous KM, Causadias JM, Casper DM. Racial discrimination and cortisol output: a meta-analysis. *Soc Sci Med*. 2017;193:90–100.
106. Williams DR, Lawrence JA, Davis BA, Vu C. Understanding how discrimination can affect health. *Health Serv Res*. 2019;54:1374–1388. <https://doi.org/10.1111/1475-6773.13222>.
107. Pech GP, Caspar EA. Can a video game with a fictional minority group decrease intergroup biases towards non-fictional minorities? A social neuroscience study. *Int J Hum-Comput Interact*. 2022;45:45.
108. Cao Y, Contreras-Huerta LS, McFadyen J, Cunnington R. Racial bias in neural response to others' pain is reduced with other-race contact. *Cortex*. 2015;70:68–78. <https://doi.org/10.1016/j.cortex.2015.02.010>.
109. Stevens FL, Abernethy AD. Neuroscience and racism: the power of groups for overcoming implicit bias. *Int J Group Psychother*. 2018;68(4):561–84. <https://doi.org/10.1080/00207284.2017.1315583>.
110. Dotson VM, Duarte A. The importance of diversity in cognitive neuroscience. *Ann NY Acad Sci*. 2020;1464(1):181–91. <https://doi.org/10.1111/nyas.14268>.
111. Ito TA, Urland GR, Willadsen-Jensen E, Correll J. The social neuroscience of stereotyping and prejudice: Using event-related brain potentials to study social perception. *Social Neuroscience: People Thinking about Thinking People*, 2006;189–212

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

