

The neuroscience of prejudice and stereotyping

David M. Amodio

Abstract | Despite global increases in diversity, social prejudices continue to fuel intergroup conflict, disparities and discrimination. Moreover, as norms have become more egalitarian, prejudices seem to have ‘gone underground’, operating covertly and often unconsciously, such that they are difficult to detect and control. Neuroscientists have recently begun to probe the neural basis of prejudice and stereotyping in an effort to identify the processes through which these biases form, influence behaviour and are regulated. This research aims to elucidate basic mechanisms of the social brain while advancing our understanding of intergroup bias in social behaviour.

Social motivations

Motives that operate in social contexts and satisfy basic, often universal, goals and aspirations, such as to affiliate (for example, form relationships and communities) or to achieve dominance (for example, within a social hierarchy).

Stereotypes

Conceptual attributes associated with a group and its members (often through over-generalization), which may refer to trait or circumstantial characteristics.

Prejudices

Evaluations of or affective responses towards a social group and its members based on preconceptions.

Social motivations, such as the desire to affiliate or compete with others, rank among the most potent of human drives¹. Not surprisingly, the capacity to discern ‘us’ from ‘them’ is fundamental in the human brain. Although this computation takes just a fraction of a second^{2,3}, it sets the stage for social categorization, stereotypes, prejudices, intergroup conflict and inequality, and, at the extremes, war and genocide. Thus, although prejudice stems from a mechanism of survival, built on cognitive systems that ‘structure’ the physical world, its function in modern society is complex and its effects are often deleterious.

For the neuroscientist, the domain of prejudice provides a unique context for examining neural mechanisms of the human mind that guide complex behaviour. Social prejudices are scaffolded by basic-level neurocognitive structures, but their expression is guided by personal goals and normative expectations, played out in dyadic and intergroup settings; this is truly the human brain *in vivo*. Although probing the neural basis of prejudice is a challenging endeavour — in which the rigours of reductionism are balanced with the richness of context — it offers neuroscientists the opportunity to connect their knowledge to some of society’s most pressing problems, such as discrimination, intergroup conflict and disparities in health and socioeconomic status.

In this article, I review research on the role of the brain in social prejudice and stereotyping. The term prejudice is used broadly to refer to preconceptions — often negative — about groups or individuals based on their social, racial or ethnic affiliations⁴. Within the field of social psychology, prejudice refers more specifically to evaluations (that is, attitudes) and emotional responses towards a group and its members. Stereotypes, by comparison, are generalized

characteristics ascribed to a social group, such as personal traits (for example, unintelligent) or circumstantial attributes (for example, poor)⁵. Although they are distinguishable by content and process, prejudices and stereotypes often operate in combination to influence social behaviour⁶. Moreover, both forms of bias can operate implicitly, such that they may be activated and influence judgements and behaviours without conscious awareness^{7–9}.

Despite the persistence of prejudices and stereotypes in contemporary society, their effects on behaviour are often countered by people’s egalitarian personal beliefs and pro-social norms⁷. Guided by these beliefs and norms, people frequently engage self-regulatory processes to mitigate the effects of bias on their behaviour. Hence, a theoretical analysis of prejudice and stereotyping is incomplete without a consideration of these regulatory processes. Here, self-regulation refers to the process of acting in an intentional manner, often through mechanisms of cognitive control.

The neuroscientific research conducted on prejudice and stereotyping over the past decade suggests that these complex forms of human behaviour involve different interacting networks of neural structures. In this article, I describe the functions of key structures in each network, including both their broader neurocognitive functions and their specific roles in prejudice and stereotyping. This article extends previous reviews on this topic — which were guided by a social psychological analysis¹⁰ or emphasized a particular neuroimaging method^{11,12} — by providing a comprehensive overview of the literature from a neural-systems perspective. Although many of the conclusions drawn from this emerging literature rely heavily on reverse inference from neuroimaging data,

New York University,
Department of Psychology,
6 Washington Place,
New York, New York 10003,
USA.
e-mail: david.amodio@nyu.edu
doi:10.1038/nrn3800
Published online
4 September 2014

Self-regulation

The process of responding in an intentional manner, often involving the inhibition or overriding of an alternative response tendency.

these inferences are strengthened by converging theory and behavioural data from the extensive psychological literature on intergroup bias and self-regulation^{13,14}.

The majority of the research reviewed here concerns racial prejudice — a form of prejudice with clearly defined social categories linked to identifiable physical attributes (BOX 1). In particular, prejudice of white Americans towards black people (that is, individuals of African or Caribbean descent) has deep historical roots and contemporary relevance to social issues, and the majority of studies have examined prejudice in this context. Nevertheless,

many findings in this literature concern basic mechanisms of social cognition that, to varying extents, underlie other forms of bias, such as those based on ethnicity, gender, sexual preference and nationality.

Neural basis of prejudice

In the modern social psychology literature, prejudice is defined as an attitude towards a person on the basis of his or her group membership. Prejudice may reflect preference towards ingroup members or dislike of outgroup members, and it is typically imbued with affect, with emotions ranging from love and pride to fear, disgust and hatred^{15,16}. Consequently, research on the neural basis of prejudice has primarily focused on neural structures involved in emotion and motivation, such as the amygdala, insula, striatum and regions of orbital and ventromedial frontal cortices (FIG. 1). Although they are often examined independently, these structures appear to form a core network for the experience and expression of prejudice.

Amygdala. Research on the neural basis of prejudice has most frequently examined the amygdala, a complex subcortical structure located bilaterally in the medial temporal lobes (FIG. 1). Although the amygdala is sometimes described as a neural locus of emotion (for example, fear), it in fact comprises approximately 13 distinct nuclei that, in conjunction, perform multiple functions to support adaptive behaviour¹⁷ (FIG. 2).

The amygdala receives direct (or nearly direct) afferents from all sensory organs into its lateral nucleus, enabling it to respond very rapidly to immediate threats in advance of more elaborative processing of a stimulus¹⁸. Within the amygdala, the central nucleus (CeA) has been implicated in Pavlovian (classical) fear conditioning in both rats and humans^{19–22}, and signals emerging from the CeA activate hypothalamic and brainstem structures to induce arousal, attention, freezing and preparation for fight or flight — a response that is often characterized as ‘fear’. By comparison, output from the basal nucleus guides appetitive and instrumental responses via projections to the ventral striatum^{22,23}. Both the fear-related and appetitive functions of the amygdala involve motivation and attention, but to different ends, and they probably correspond to different aspects of a prejudice-based response. In humans, the amygdala is integral to the processing of fear in facial expressions as well as other salient social cues²⁴. Given the amygdala’s ability to respond rapidly to potential social threat, researchers interested in the neural substrate of implicit prejudice first looked to this brain structure.

A pair of early functional MRI (fMRI) studies examined the amygdala activity of white research subjects in response to blocked presentations of black and white faces^{25,26}. Although neither study found that amygdala activity to faces varied significantly as a function of ‘race’, their results were suggestive: one study showed that the relative difference in subjects’ amygdala activity to black versus white faces was correlated with a behavioural indicator of implicit prejudice (BOX 2) and with relative differences in the startle eyeblink response to black

Box 1 | Seeing race: the role of visual perception

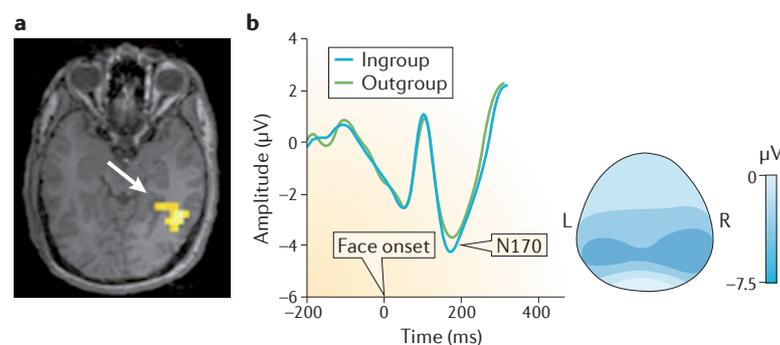
Social interactions often begin with the perception of a face. Mounting evidence reveals that social motivations can alter the way a face is seen, which presumably reflects the modulatory influences of signals from the temporal cortex, prefrontal cortex and amygdala to the fusiform gyrus¹⁴⁹. This insight suggests that prejudices and stereotypes may alter early face processing.

Early functional MRI (fMRI) research demonstrated greater fusiform activity (see the figure, panel **a**) in response to faces of one’s own racial group (that is, the ingroup) — an effect that was associated with better recognition of ingroup faces than outgroup faces⁵⁰. Research examining the N170 component of the event-related potential (ERP), which indexes the degree of initial configural face encoding at just ~170 ms, revealed enhanced processing of ingroup versus outgroup faces (see the figure, panel **b**), even when groups were defined arbitrarily³. This finding is consistent with fMRI data showing that faces of ‘coalition members’ elicited greater activity in the fusiform gyrus than did other faces, regardless of race¹⁴⁹. Hence, social group membership, even when defined on the basis of minimal categories, promotes greater visual encoding. These findings dovetail with behavioural research showing that biased visual representations of outgroup members facilitate discriminatory actions towards them^{150,151}.

In the context of race, outgroup members are often viewed as threatening and therefore may elicit vigilant attention. Indeed, larger N170 ERP amplitudes in response to viewing black versus white faces (equated in luminance) have been observed in subjects with stronger implicit prejudice¹⁵² and in subjects who were made to feel anxious about appearing biased⁴². These and other findings suggest that the visual processing of race is malleable and depends on social motivations and contexts^{153–158}.

Neural representations of race (black versus white), as determined by multivoxel pattern analysis (MVPA), have been observed in the fusiform gyrus, and these neural representations have been associated with behavioural indices of implicit prejudice and stereotyping^{49,52,159,160}. It is notable that MVPA has also identified race representation in the medial occipital cortex; however, because full-colour photos were used in these studies, the effect may reflect differences in luminance associated with skin tone rather than the race of the people depicted. Nevertheless, the broader body of findings suggests that social category cues modulate the early visual processing of ingroup and outgroup members’ faces in ways that support the perceivers’ biased or egalitarian social goals.

Panel **a** of the figure is from REF. 50, Nature Publishing Group. Panel **b** of the figure is reprinted from *J. Exp. Soc. Psychol.*, 49, Ratner, K. G. & Amodio, D. M., Seeing “us versus them”: minimal group effects on the neural encoding of faces, 298–301. Copyright (2013), with permission from Elsevier.



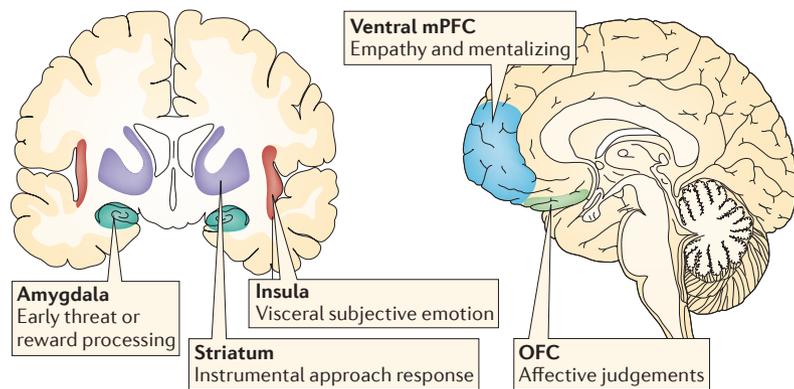


Figure 1 | Prejudice network. An interactive set of neural structures that underlie components of a prejudiced response. The amygdala is involved in the rapid processing of social category cues, including racial groups, in terms of potential threat or reward. Approach-related instrumental responses are mediated by the striatum. The insula supports visceral and subjective emotional responses towards social ingroups or outgroups. Affect-driven judgements of social outgroup members rely on the orbital frontal cortex (OFC) and may be characterized by reduced activity in the ventral medial prefrontal cortex (mPFC), a region involved in empathy and mentalizing.

versus white faces²⁵. The other study, which also included black subjects²⁶, showed that amygdala responses habituated more slowly to racial outgroup faces. Together, these studies identified the amygdala as a candidate substrate of implicit prejudice.

To examine the timing and function of the amygdala response to race more precisely, a later study used the startle eyeblink method to index CeA-dependent amygdala activity at very brief intervals following the presentation of a face image²⁷. This study revealed significantly greater startle activity in response to black faces relative to white or Asian faces — an effect that varied with subjects' self-reported motivations to respond without prejudice. By demonstrating this differential response to race with an index associated with CeA activity, this study more directly implicated fear conditioning as a mechanism underlying implicit prejudice. This link suggested that the extensive literature on fear conditioning can, to some extent, inform our understanding of implicit prejudice, specifically regarding how this form of bias may be learned, expressed and potentially extinguished^{28–31}.

The role of the amygdala in implicit prejudice has been examined in many subsequent studies. Much of this research suggests that amygdala activation reflects an immediate (or implied) threat response to racial outgroup members^{10,12,32}. For example, in white subjects viewing images of black faces, amygdala activation is greater in response to faces with darker rather than lighter skin tone³³; when the eyegaze of the target face is direct rather than averted³⁴; when judgements of faces are made on the basis of superficial information³⁵; and in contexts evoking interracial threat³⁶. Moreover, some evidence suggests that the amygdala response is stronger when ingroup and outgroup faces are presented very briefly, presumably because the brief presentation precludes the regulation of this response³⁷. By contrast, familiarity with racial outgroup members is associated with an attenuated difference in the amygdala response

to outgroup versus ingroup faces, both in children and adults^{25,38–40}. Together, these findings corroborate social psychology theories of implicit prejudice as reflecting a form of threat processing and suggest new links between implicit prejudice, Pavlovian fear conditioning and affective processes.

It is also possible that the amygdala response in some studies reflects not a direct threat from an outgroup member but rather the threat of appearing prejudiced in the presence of others who may disapprove of bias. Indeed, in white subjects, anxiety about appearing prejudiced to others has been shown to enhance eyegaze fixations and early visual processing of black faces^{41,42}, and low-prejudice individuals who worried about appearing prejudiced to others showed larger startle eyeblink responses to black versus white faces compared with low-prejudice individuals without this concern²⁷. This possibility — that amygdala activity in response to racial outgroups is due to the threat of appearing prejudiced to others — is consistent with findings from many social psychology studies^{43,44} but has not been tested directly.

More recent studies have emphasized that the amygdala response to an ingroup or outgroup member depends on a perceiver's goals: when exposure to images of people from a different racial group is combined with an unrelated secondary task (for example, to detect the appearance of a small dot on the image), race no longer drives the amygdala response^{45,46}. In fact, in a study in which the subject's goal was to identify white and black individuals in terms of coalition (for example, whether each individual belonged to one's own sports team, irrespective of race), it was coalition, and not race, that drove the amygdala response⁴⁷. Specifically, amygdala activity was highest in response to the subject's own team members. Still other studies have found no differences in amygdala activity in response to different racial groups, presumably because the study designs focused subjects' attention on task features other than race^{48–53}. Although these findings may seem to contradict other research linking amygdala activity to threat, they are consistent with a broader model of amygdala function, which proposes that it responds to motivationally relevant cues — aversive or rewarding — to guide adaptive behaviours^{22,23,54–56}.

To date, the research literature suggests that there are three main patterns of amygdala function with respect to intergroup responses. One pattern reflects a learned threat response to racial outgroups, which is ostensibly rooted in fear conditioning. A second, but still speculative, pattern may reflect the threat experienced by a perceiver who worries about appearing prejudiced in the eyes of others when viewing faces of racial outgroup members. Both of these patterns probably represent activity of the CeA, given its known role in fear conditioning and anxiety. A third pattern seems to reflect instrumental (that is, goal-directed) responses, suggesting approach-related motivation and attention towards members of the ingroup (which can be based on race or other social categories). This instrumental response probably reflects output from the basal nucleus, given the involvement of this nucleus in

Configural face encoding

The visual encoding of a face in terms of its basic structural characteristics (for example, the eyes, nose, mouth and the relative distances between these elements). Configural encoding may be contrasted with featural encoding, which refers to the encoding of feature characteristics that make an individual's face unique.

Instrumental responses

Actions performed to achieve a desired outcome (that is, goal-directed responses).

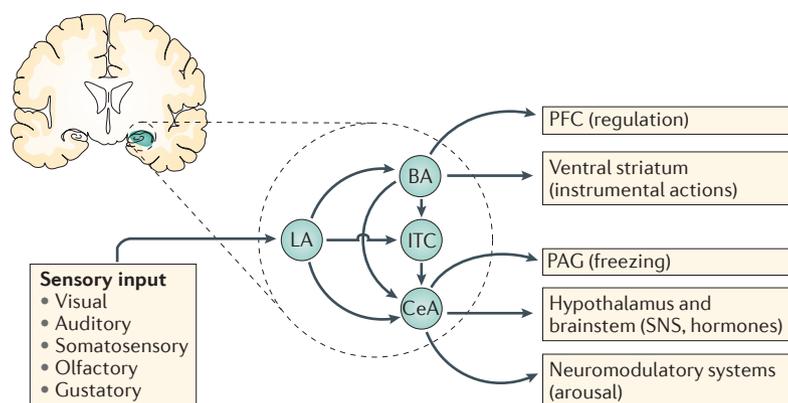


Figure 2 | The amygdala and its role in prejudice. Amygdala activity is frequently observed in individuals while they view members of racial outgroups, but it has also been found in response to viewing members of one’s own group independently of race⁴⁷. This mixed finding may reflect the different functions of nuclei within the amygdala. The figure depicts three amygdala nuclei that probably contribute to these two forms of prejudice: sensory inputs enter via the lateral nucleus of the amygdala (LA) and, depending on the context and nature of the stimuli, this signal is directed to the central nucleus of the amygdala (CeA), which supports a threat response, or to the basal nucleus of the amygdala (BA), which supports an instrumental response¹⁸. Because of the inhibitory nature of within-amygdala projections, activating signals involve connections through intercalated masses (ITCs). PAG, periaqueductal grey; PFC, prefrontal cortex; SNS, sympathetic nervous system.

goal-directed behaviour. Together, these findings identify the amygdala as a major substrate of different forms of implicit prejudice. However, it is important to note that behavioural expressions of bias, such as in social interactions or on a laboratory task (for example, the implicit association test (IAT)), may reflect other processes — such as conceptual associations, intentions and cognitive control — in addition to an amygdala-based response⁵⁷. As the contributions of different amygdala nuclei become better understood, and with more refined behavioural assessments of implicit bias, the role of the amygdala in prejudice and other social processes will become increasingly clear.

Orbital frontal cortex. The orbital frontal cortex (OFC) (FIG. 1), which is often considered to include the inferior ventral medial prefrontal cortex (mPFC), is associated with the processing of affective cues, contingency-based learning, evaluation and decision making^{58–60}. In the social domain, the OFC supports the monitoring of social cues and subsequent adjustment of one’s behaviour⁶¹. This function is crucial in intergroup situations involving social norms, in which responses may be influenced by others’ expectations⁶². Moreover, the OFC is anatomically interconnected with brain regions involved in all sensory modalities and with structures that are known to represent emotional and reward processes (such as the basal nuclei of the amygdala and striatum) and social knowledge (such as the medial frontal cortex and temporal poles)⁶³. In comparison with the amygdala, the OFC seems to support more complex and flexible evaluative representations that are more directly applicable to the intricacies of social behaviour.

To date, relatively few studies have examined the role of the OFC in prejudice, most likely because the field has primarily focused on comparatively basic responses to racial group members (for example, through passive viewing) rather than the kind of complex evaluative processes that are known to involve the OFC. However, findings from these studies are generally consistent with the OFC’s proposed role in complex evaluations of people based on group membership, beyond its potential role in implicit racial attitudes⁶⁴. For example, OFC activity has been associated with subjects’ deliberative judgements regarding the prospect of befriending black, relative to white, individuals⁴⁹ (FIG. 3). OFC activity has also been associated with subjects’ preference for members of their own team independently of race, indicating that the OFC may have a broader role in group-based evaluation⁴⁷. Given its role in the regulation of social behaviour⁶¹, the OFC is likely to emerge as an important substrate of more elaborated forms of intergroup evaluation.

Insula. The insula (FIG. 1a) is a large cortical region that runs medial to the temporal lobes, adjacent to the frontal cortex, and broadly functions to represent somatosensory states (including visceral responses) and emotions related to such states (such as disgust)⁶⁵. Posterior insula regions are thought to provide primary representation of interoceptive signals, whereas anterior regions support the cognitive re-representation of these signals. This re-representation in the anterior insula provides an interface with the anterior cingulate cortex (ACC) and PFC, which are involved in subjective awareness of emotion and cognitive control⁶⁶. It is the anterior insula, rather than the posterior insula, that is most frequently associated with aspects of social cognition and social emotion.

Although the insula is rarely of focal interest in neuroimaging studies of prejudice, its activity is frequently associated with responses to racial outgroup versus ingroup members in experimental tasks^{33,45,48}. This finding has been interpreted as reflecting a negative visceral reaction, such as disgust, to racial outgroups⁶⁷, and it has been specifically associated with white subjects’ implicit negative attitudes towards black people^{51,64}. Thus, the insula seems to contribute to the subjective affect that is often experienced as part of a prejudiced response. It could be speculated that the representation of this affective response in the anterior insula may — through its connections with the ACC and PFC — facilitate the ability to detect and regulate one’s behaviour on the basis of a prejudicial affective response.

It is notable that the insula is also implicated in pro-social emotions, such as empathy, towards liked individuals^{68–70}. For example, insula activity was found to increase when subjects viewed another person being exposed to a painful stimulus, but only if that person was of the same racial group⁷¹. Similarly, another study observed insula activity when members of liked, but not disliked, outgroups were harmed⁶⁷. Both findings suggest that empathy-related activity in the insula depends on the victim’s social affiliation. In an interesting twist,

Implicit bias
Prejudiced or stereotype-based perceptions or responses that operate without conscious awareness.

Deliberative judgements
Judgements that result from thoughtful considerations (often involving cognitive control) as opposed to rapid, gut-level, ‘snap’ judgements.

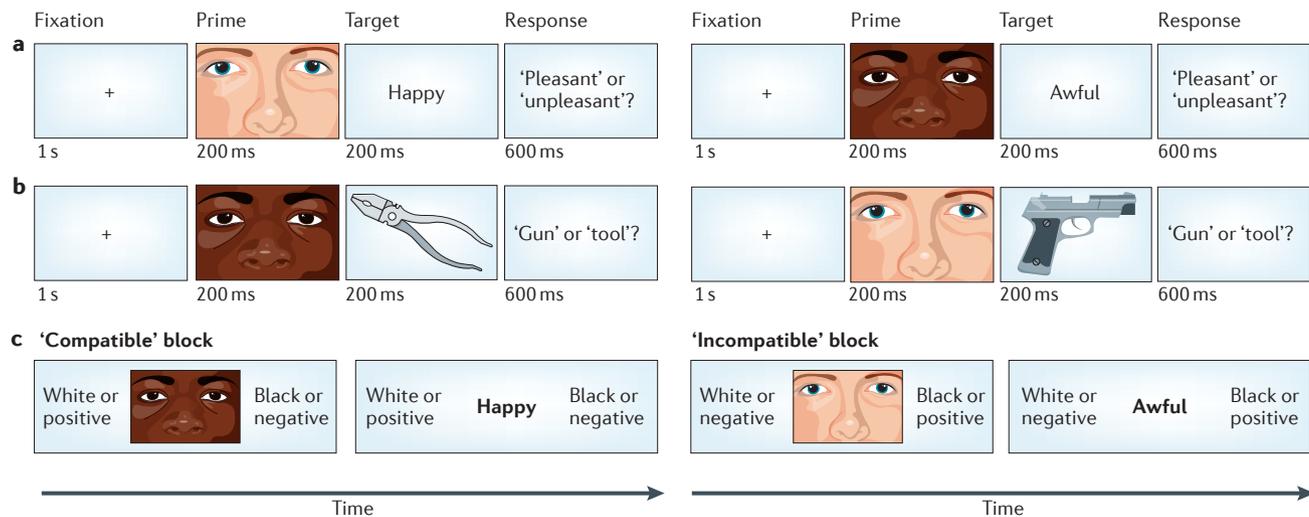
Box 2 | **Measuring implicit prejudice and stereotyping**

Unlike explicit racial beliefs, implicit attitudes and stereotypes reflect associations in the mind that operate without conscious awareness⁹. Implicit attitudes associated with race are formed through direct or indirect exposure to members of these racial groups in negative (or sometimes positive) contexts. Such implicit racial associations are typically assessed using computerized priming tasks; the priming effect is considered to be ‘implicit’ because subjects may be unaware that they possess racial associations or may be otherwise unaware of how their racial associations affect their task responses. Racial bias assessed by implicit measures such as these has been shown to predict a wide range of behavioural forms of discrimination¹⁶¹.

In an example of a sequential priming task, subjects view and classify target words as either ‘pleasant’ or ‘unpleasant’ (see the figure, part a). Each target word is preceded by a prime stimulus that represents a social category: for example, white and black faces. Implicit prejudices are revealed in task performance: among white Americans, negative (versus positive) words are often classified more quickly following black faces than following white faces⁹.

A variant used to assess implicit stereotype associations is the weapons identification task, in which white and black face stimuli (primes) are followed by images of handguns and handtools¹⁶² (see the figure, part b). Black primes typically facilitate the categorization of guns and interfere with the categorization of tools, reflecting the stereotype of black Americans as dangerous. Because this task creates stereotype-based interference (on black-face prime–tool trials), it is also used to elicit and index the cognitive control of stereotyping.

In the implicit association test (IAT), subjects view a series of stimuli, such as white and black faces and positive and negative words¹⁶³ (see the figure, part c). During ‘compatible’ trials, white faces and positive words are categorized using one key, whereas black faces and negative words are categorized with a different key. During ‘incompatible’ trials, categories are rearranged: white faces and negative words are categorized with one key, and black faces and positive words with the other key. A tendency to respond more quickly on compatible than incompatible blocks is taken to indicate an anti-black and/or pro-white attitude. The IAT effect represents the difference in average response latency between these two trial blocks, with higher scores indicating stronger implicit prejudice.



insula activity has also been observed when a disliked outgroup member is rewarded — a case of outgroup envy — and the degree of this activation predicted subjects’ intention to harm that individual⁷². Although our understanding of insula function in social contexts is still developing, these findings highlight a role of visceral responses to other people that has been largely overlooked in past social-cognition research but that may nonetheless be crucial for guiding intergroup social behaviour.

Striatum. The striatum is a component of the basal ganglia that comprises the caudate nucleus and putamen (FIG. 1). This structure is broadly involved in instrumental learning and reward processes, including the coordination of goal-directed and habit-based responses through bidirectional connections with the PFC (via the caudate nucleus) and with motor areas (via the putamen), respectively⁷³. Findings from functional neuroimaging research on economic bargaining and reinforcement learning suggest that striatal activation is associated with the computation of value (that is, value placed on a potential action) and anticipated outcomes^{74,75}.

Consistent with a role of the striatum in reward processing, fMRI studies of social perception have revealed increased striatal activity in response to viewing pictures of ingroup versus outgroup members⁴⁷. In a study in which white subjects completed an IAT that assessed preferences for black versus white individuals, caudate activity was stronger when subjects viewed white faces compared with black faces, and this difference was associated with an implicit preference for white ingroup members⁶⁴. In an economic bargaining game, the degree of trust shown by white subjects towards a black partner was associated with striatal activity⁷⁶. These initial findings suggest that the striatum has a role in guiding positive intergroup interactions through instrumental and approach-related responses.

Medial prefrontal cortex. The medial frontal cortex — which encompasses Brodmann area 8 (BA8), BA9 and BA10 along the medial wall of the frontal cortex, superior and anterior to the ACC — has emerged as a particularly important structure for the processing of social information^{62,77–79}. This highly associative region has prominent

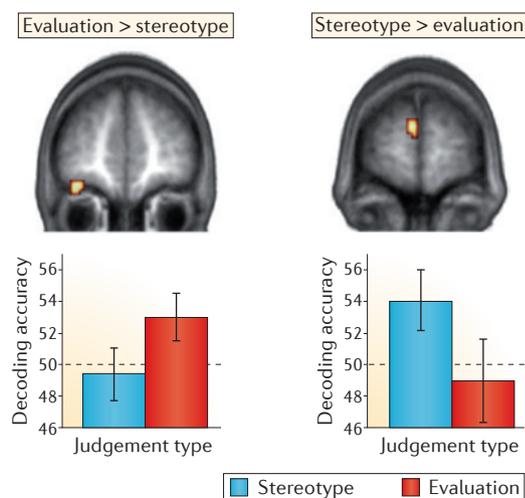


Figure 3 | Neural representation of racial bias in affect-based and stereotype-based judgements.

A multivoxel pattern analysis approach revealed a unique representation (that is, decoding accuracy) of race in the orbital frontal cortex when subjects judged images of black and white males according to an evaluative dimension (who is more likely to be a friend?), and a unique representation of race in the medial prefrontal cortex when judging black and white males on a stereotypical trait dimension (who is more interested in athletics?)⁴⁹. Reprinted from *Neuropsychologia*, 50, Gilbert, S. J., Swencionis, J. K. & Amodio, D. M., Evaluative versus trait representation in intergroup social judgments: distinct roles of anterior temporal lobe and prefrontal cortex, 3600–3611, Copyright (2012), with permission from Elsevier.

interconnections with the ACC, the insula, the OFC and the dorsolateral PFC (dlPFC), as well as other structures⁶².

In the context of social cognition, mPFC activity has been primarily associated with the formation of impressions about other people, especially impressions that require mentalizing — the process of considering a person's unique perspective and motives (that is, engaging in theory of mind)⁸⁰. Because the mPFC is typically activated during judgements about other people (as opposed to inanimate objects)⁷⁷, mPFC activity, particularly in ventral, perigenual regions (FIG. 1), has been interpreted by some theorists as reflecting a 'humanization' process and, by extension, empathy^{67,81}. Hence, a lack of mPFC activity in response to a social target may indicate a form of prejudice that is characterized by a lack of humanization (that is, dehumanization) and empathy. Indeed, the ventral mPFC has been shown to be more highly activated when a subject views members of esteemed groups associated with pride and admiration than when a subject views members of low-status groups associated with disgust and disregard (for example, homeless people)⁶⁷. Moreover, in a study in which Chinese and Caucasian subjects viewed images of people being exposed to a painful or non-painful stimulus (that is, a needle penetrating the cheek versus a Q-tip touch), mPFC and ACC activity was elicited only in response to seeing racial ingroup

members in pain⁷¹. In a conceptually related study of gender bias, men who reported highly sexist attitudes exhibited lower mPFC activity when viewing sexualized images of female (but not male) bodies than men with less sexist views — a pattern consistent with idea that sexual objectification involves a form of dehumanization⁸². Hence, the mPFC's role in prejudice seems to be marked by an absence of activity, which may reflect a lack of humanization and empathy regarding disliked or disrespected outgroup members.

A neural network for prejudice. Prejudice is a complex social cognitive process that seems to be supported by a network of neural structures (FIG. 1). The amygdala supports threat-based associations, which are thought to underlie the most common form of implicit prejudice, and it is also involved in initial responses to salient positive or negative cues, including cues regarding group membership. Thus, depending on the situation and nuclei of interest, amygdala activity may be associated with social threat or with social reward. Activity in the anterior insula supports the subjective experience of negative affect (which often accompanies a prejudiced response), whereas the mPFC is involved in mentalizing and perspective taking, which may be engaged more strongly towards ingroup than outgroup members. Neural projections from the amygdala and insula to the ventral mPFC may support the integration of affective responses with mentalizing and empathy processes. Finally, appetitive responses such as positive attitudes and approach-related behavioural tendencies, which are often expressed towards ingroup members, are primarily supported by the striatum. These brain regions may function in concert to support the learning, experience and expression of prejudice.

Neural basis of stereotyping

In contrast to prejudice, which reflects an evaluative or emotional component of social bias, stereotypes represent the cognitive component — the conceptual attributes linked to a particular group as defined by a culture or society. This process involves the encoding and storage of stereotype concepts, the selection and activation of these concepts into working memory and their application in judgements and behaviours^{5,83}. As such, stereotyping involves cortical structures that support more general forms of semantic memory, object memory, retrieval and conceptual activation, such as the temporal lobes and inferior frontal gyrus (IFG), as well as regions that are involved in impression formation, such as the mPFC^{10,84,85}. Although there is substantial overlap between these structures and those implicated in prejudice, as described above, these structures appear to form the core of a stereotyping network in the brain that may operate separately from a prejudice network (FIG. 4).

Temporal lobe. Stereotypes reflect conceptual associations between social groups and a particular set of attributes — associations that are thought to reside in semantic memory⁸³. As such, stereotype associations are posited to involve regions of the lateral temporal lobe that

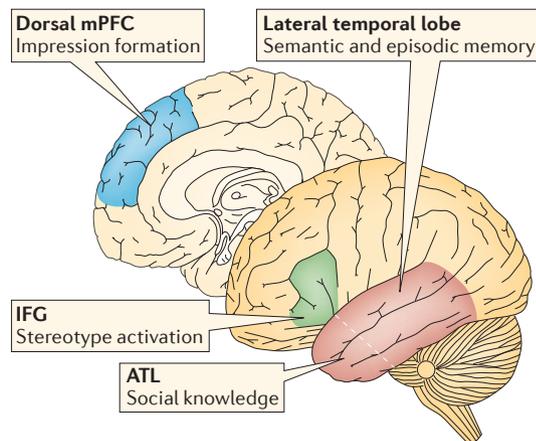


Figure 4 | **Stereotyping network.** Neural structures that underlie components of intergroup stereotyping. Semantic information stored in the lateral temporal lobe — especially representations of stereotype-related knowledge about people and social groups in the anterior temporal lobe (ATL) — is recruited into the dorsal medial prefrontal cortex (mPFC) to support the formation of impressions (that is, stereotypes) and, in conjunction, into the inferior frontal gyrus (IFG) to support goal-directed actions that are guided by these stereotypes.

underpin semantic knowledge^{10,85–89} (FIG. 4). In particular, the anterior temporal lobe (ATL) is associated with the representation of social knowledge, such as attributes that describe people but not inanimate objects^{84,90,91}. The dorsal part of the ATL, which is implicated more specifically in the representation of social objects (that is, people), is densely interconnected with the regions of the mPFC that are associated with trait judgement and impression formation⁹². This suggests that social information represented in the ATL is selected into the mPFC to support the process of social cognition.

Not surprisingly, the ATL is frequently implicated in studies of stereotype representation. In one fMRI study examining the neural basis of stereotyping, subjects considered either social or non-social categories (for example, men versus women or violins versus guitars) and judged which category was more likely to be characterized by a particular feature (for example, enjoys romantic comedies or has six strings). A contrast of brain activity between social and non-social conditions revealed that ATL activity was uniquely activated during stereotype-relevant judgements of social categories⁹³. A different fMRI study used multivoxel pattern analysis (MVPA) to examine neural activity representing judgements of black and white individuals on the basis of stereotype traits (athleticism) versus evaluations (potential for friendship)⁴⁹. Results showed that when subjects made trait judgements, a behavioural index of implicit stereotyping correlated with ATL activity, and when they made evaluative judgements, a behavioural index of implicit racial attitudes correlated with activity in the same part of the ATL. Consistent with these findings, the disruption of ATL activity by transcranial magnetic stimulation attenuated the behavioural expression of implicit gender stereotype associations⁹⁴, suggesting

that the ATL is necessary for stereotype representation. Thus, knowledge of social stereotypes appears to reside in the ATL.

Medial prefrontal cortex. As discussed above, the mPFC is consistently involved in the representation of an individual's traits, preferences and mental states during impression formation^{77,80}. Although relevant to aspects of prejudice, the mPFC is more directly involved in stereotyping.

To date, the neural substrates of stereotyping have mainly been examined within the domains of gender and political orientation. These studies have linked mPFC activity, typically in dorsal regions, with the activation of gender-related and political concepts during behavioural tasks such as the IAT^{51,89,95–98}. The mPFC has been implicated in the domain of racial stereotyping as well, during tasks that require subjects to infer personal traits of individuals from racial minority groups (for example, African Americans)³⁵. In an fMRI study designed to distinguish the neural representation of stereotype-based judgements of black versus white people from judgements based on affective responses, MVPA results identified the rostral dorsal mPFC as the only region representing stereotype judgements⁴⁹.

Although the mPFC has been linked to stereotyping, its precise role in this process remains a point of inquiry. Some authors have conceptualized the anterior mPFC as a repository of social knowledge^{79,99} or as a region that integrates information about social knowledge with goals in order to coordinate social behaviour^{49,62,93}. Researchers are beginning to investigate these alternative functions^{100,101}. Nevertheless, in either case, the mPFC seems to be centrally involved in the stereotype-based processing of people.

It is notable that the mPFC is often considered to function as part of a social-cognition (or mentalizing) network, together with the temporoparietal junction, superior temporal sulcus, precuneus and ATL^{78,102–104}. As discussed, the mPFC and ATL have been directly linked to social stereotyping processes, whereas the other regions seem to be primarily associated with theory-of-mind processing, action understanding and self-consciousness — processes that are less directly relevant to stereotyping and prejudice. Hence, the set of regions involved in functional networks proposed for one psychological function (for example, making mental state inferences) may not cohere in the context of another (for example, stereotyping) despite the fact that both functions represent aspects of social cognition. In this case, an involvement of the mPFC in stereotyping does not necessarily implicate other components of networks associated with mentalizing and social cognition.

Lateral prefrontal cortex. The lateral PFC — more specifically, the regions often referred to as the IFG (FIG. 4) (BA44, BA45 and BA47) — is associated with the selection of concepts into working memory to support goal-directed action^{87,105–109}. William James famously observed that 'thinking is for doing', and the left IFG, in particular, reflects this notion: strong reciprocal connections of the

lateral PFC with the basal ganglia and motor cortex support the coordination of complex actions that are guided by working memory and high-level cognition^{73,110}. Stereotypes are a form of social cognition that guide behaviour, and indeed the process of applying stereotypes in judgement and behaviour has been shown to specifically involve activity in the IFG¹¹¹.

Whereas the retrieval of conceptual knowledge typically involves the left IFG, activity in the right IFG has been observed in research subjects who were judging whether gender-stereotyped traits applied to a series of male and female individuals (as compared with traits that were unrelated to gender stereotypes)¹¹¹. Given other evidence that the right IFG has a role in domain-general response inhibition¹¹², it is possible that activation of the right IFG during stereotype judgement tasks reflects an individual's efforts to inhibit the influence of stereotypes on behaviour. This pattern of lateralized function — stereotype retrieval and implementation on the left and response inhibition on the right — suggests a useful distinction in the processes through which stereotypes are applied in behaviour.

A neural network for stereotyping. The research described above suggests that a network of neural structures supports stereotyping processes (FIG. 4). The ATL is believed to represent stereotype-related knowledge, and it provides input to the mPFC, possibly also during the online formation of impressions about an individual. In this way, social stereotypes 'stored' in the ATL may influence trait impression processes associated with dorsal mPFC activity. The application of stereotypes to behaviour seems to involve regions of the lateral PFC that are associated with goal representation and response inhibition. Together, the structures in this putative network may support the storage, activation and behavioural expression of social stereotypes.

Interacting networks

The framework described above suggests separate networks for prejudice and stereotyping, but in most cases these two processes operate in concert. Thus, although they are largely rooted in distinct neural systems, their effects converge in higher-level cognition and behavioural expression. Neuroscience studies suggest several places at which this convergence may occur, although this is primarily based on studies of connectivity in non-human animals. For example, the anatomical connectivity of the amygdala and OFC with the ATL, via the uncinate fasciculus⁹², is consistent with behavioural evidence that affective responses may influence the activation of stereotype concepts, and vice versa^{113,114}. Similarly, signals from several structures that are involved in both prejudice and stereotyping — including the amygdala, insula, striatum, OFC and ATL — converge in regions of the mPFC, where information seems to be integrated in support of elaborate person representations^{62,63}. Finally, the joint influences of prejudiced affect and stereotype concepts on behaviour are likely to converge in the striatum, which receives inputs from the amygdala, OFC, lateral PFC and ATL, as well as other regions^{73,115}. Although

the coherence of these proposed functional networks for prejudice and stereotyping and their interaction has yet to be tested, their existence is consistent with known anatomical connectivity (which has been primarily observed in non-human animals) and is supported by decades of theory and behavioural research in the social psychology literature^{4,5}.

Regulation of prejudice and stereotyping

In an era of increasing diversity, international relations, global communication and awareness of civil rights issues, intergroup biases are often deemed to be both personally and socially unacceptable. Preferences based on racial and ethnic categories that may have been adaptive in less complex societies are no longer so. Fortunately, the human mind is adept at self-regulation, and although stereotypes and prejudices may come to mind automatically in intergroup contexts, their expression can often be moderated. Neuroscience research on the mechanisms supporting the control of intergroup responses incorporates existing domain-general models of cognitive control into broader models that consider the influence of social factors. For example, the impetus for the control of racial bias may arise from internal cues (for example, the personal rejection of prejudice) or external cues (for example, social pressure to respond without prejudice), and engagement in control is frequently associated with social emotions such as social anxiety or guilt⁴³. A neural model of prejudice control should account for these different impetuses and emotion effects. In this way, neuroscience research on prejudice has inspired an expanded view of the neural and psychological processes involved in control.

Anterior cingulate cortex. The ACC has been widely implicated in the monitoring and detection of response conflict^{116,117}. In particular, the dorsal region of the ACC (FIG. 5) is often activated during cognitive control tasks, such as the Stroop or Flanker tasks, on trials involving a high degree of conflict between one's desired response and a countervailing tendency^{118,119}. Conflict monitoring theory posits that as the conflict signal in the ACC rises, the ACC increasingly engages dlPFC regions that function to implement goal-directed behaviour¹²⁰. This model is consistent with the ACC's connectivity with PFC regions involved in high-level goal representation and with the PFC's connectivity with the striatum, through which top-down control is implemented in behaviour^{73,110,121}.

In a social context, cognitive control is needed to curb the unwanted influence of implicit stereotypes and prejudices on behaviour^{7,122}. Building on conflict monitoring theory, it has been proposed that the control of implicit bias requires the detection of a conflict between a biased tendency and one's goal to act without bias¹²³. Support for this proposal was provided by a study that assessed ACC activity, which was indexed by the error-related negativity component of the event-related potential (ERP). In this study, subjects performed a task that required them to inhibit the automatic expression of racial stereotypes on some trials but not others (the

Event-related potential (ERP). An electrical signal produced by summated postsynaptic potentials of cortical neurons in response to a discrete event, such as a stimulus or a response in an experimental task. Typically recorded from the scalp in humans, ERPs can be measured with extremely high temporal resolution and can be used to track rapid, real-time changes in neural activity.

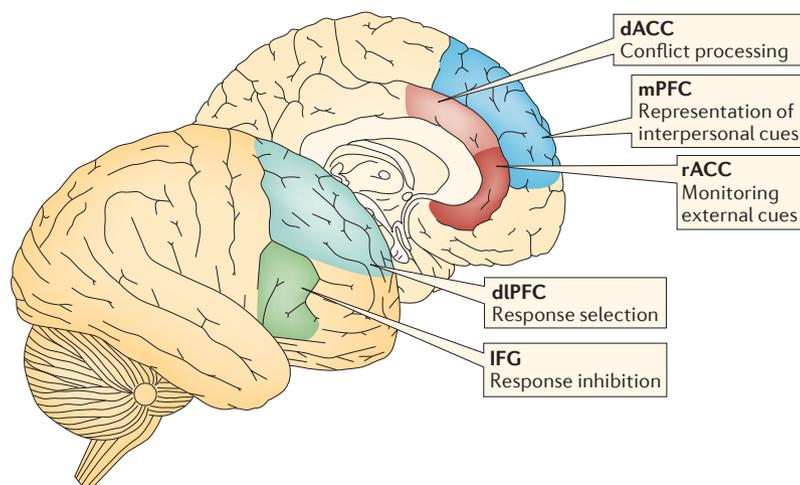


Figure 5 | Regulation network. Neural structures supporting the regulation of intergroup responses. Conflicts between a biased tendency and either internal goals or external cues (for example, social norms) are processed in the dorsal anterior cingulate cortex (dACC) and rostral ACC (rACC) and medial prefrontal cortex (mPFC). The medial prefrontal cortex (mPFC) is involved in perspective taking and mentalizing, and activation in this region provides further representation of interpersonal cues to guide regulatory processing. Intergroup response goals are represented in the lateral PFC and implemented in behaviour in coordination with the striatum and motor cortex. dlPFC, dorsolateral PFC; IFG, inferior frontal gyrus.

weapons identification task; see BOX 2). ACC activity was selectively greater on trials requiring stereotype inhibition, and the degree of subjects' ACC activity on these trials predicted their success at controlling the expression of stereotypes in task behaviour¹²³. This finding has been replicated and extended in several studies using different tasks and with alternative ERP indices of ACC activity (for example, the error-related negativity and response-locked N2)^{124–127}. These studies revealed, for example, that ACC activation in response to stereotype conflict occurs implicitly and without deliberation, is observed on a range of cognitive control tasks that require the inhibition of stereotype-based responses and is associated with an individual's motivation to respond without prejudice.

In several fMRI studies, ACC activation has also been observed in white American subjects while they viewed images of black (versus white) faces. However, this activation typically occurred in the absence of an opportunity for cognitive control (that is, during tasks that did not require control, such as passive face viewing)^{37,45,48}. Thus, it is difficult to interpret ACC activations in these studies in terms of a control process. Nevertheless, these findings suggest that exposure to black faces may spontaneously elicit conflict detection processes. In one notable exception — a study in which fMRI was recorded during performance on a racial attitude IAT, which requires cognitive control — ACC activity was associated with the ability to identify the correct (that is, non-biased) task response⁶⁴. This finding is conceptually consistent with evidence from ERP studies that link ACC activity to conflict processing and the engagement of top-down control¹²⁸. Similar patterns have been observed when people are confronted with explicit

feedback about their bias. In two fMRI studies in which subjects completed IAT measures of racial attitudes, false IAT feedback to the subjects indicating that he or she showed racial prejudice elicited heightened ACC activity, and this degree of activity was associated with feelings of guilt — a self-regulatory emotion that promotes prosocial behaviours^{129,130}. Together, these findings suggest that the ACC supports the detection of one's unwanted social biases and the engagement of cognitive control in order to avoid the expression of bias.

Lateral prefrontal cortex. As noted above, the PFC is associated with working memory, response selection and the representation of high-level goals^{131,132}, and it governs most goal-directed responses in humans. Lateral PFC regions (FIG. 5), in particular, coordinate the control of action and attention; most findings indicate that activity in the left lateral PFC is linked to the implementation of action, whereas activity in the right lateral PFC is linked to action inhibition (in right-handed individuals)^{133,134}. On the basis of research in cognitive neuroscience, lateral PFC regions have been proposed as primary substrates of cognitive control of prejudice^{27,37,64}.

Until recently, fMRI studies of responses towards ingroup versus outgroup faces were not designed to elicit or assess cognitive control. That is, most studies used tasks in which pictures of racial ingroup and outgroup members were viewed passively by subjects. Interestingly, these studies consistently revealed activity in regions of the PFC, most often the right IFG, in response to explicit presentations of black faces compared with white faces^{37,45,48}. Although this finding is difficult to interpret in terms of control, the established role of the right IFG in response inhibition suggests that exposure to black faces in these tasks may have spontaneously elicited a form of inhibitory control, perhaps owing to subjects' concern about appearing prejudiced. In addition, right IFG activity has been reported to correlate negatively with amygdala activity in response to viewing black faces. This could suggest a potential regulatory circuit for prejudice control^{37,45}; however, as these data are merely correlational and as there is little direct connectivity between the IFG and amygdala¹³⁵, it is more likely that right IFG activity reflects a form of response inhibition rather than the direct downregulation of amygdala activity.

The role of lateral PFC activity in the cognitive control of race biased behaviour was examined directly in an electroencephalography (EEG) study, in which subjects completed a task that assessed the behavioural inhibition of stereotypes (the weapons identification task)¹³⁶. Greater dlPFC activity was found to be associated with better behavioural control (as modelled using the process dissociation procedure), indicating a direct link between the dlPFC and control of stereotyping. Furthermore, the relation between dlPFC activity and stereotype control was mediated by greater attentional orienting to black faces than white faces, as indexed by the P2 component of the ERP. This pattern suggested that PFC activity tuned perceptual attention to relevant stimuli, which in turn facilitated behavioural

control (BOX 2). Findings consistent with this idea have been reported in studies using EEG, fMRI and brain lesion approaches in combination with behavioural tasks designed to assess elements of prejudice control^{51,64,126,137,138}. Together, these studies have begun to identify the specific pathways through which the PFC guides the control of intergroup responses.

Medial prefrontal cortex. As noted in previous sections, the mPFC contributes to aspects of both stereotyping and prejudice. However, this region is large, heterogeneous and widely interconnected, and emerging theories and research suggest that its function may be closely tied to regulatory processes as well. Amodio and Frith⁶² proposed that, given its role in mentalizing, the mPFC supports the regulation of behavioural responses according to social cues. Early evidence from ERP data suggested that activity in the mPFC and/or rostral ACC was uniquely associated with behavioural control that is guided by external social cues, whereas activity in the dorsal ACC was associated with internally controlled behaviour¹²⁴. In addition, ventral portions of the mPFC are interconnected with the OFC and amygdala, and through these connections it may support the top-down modulation of emotional responses¹³⁹. Hence, the mPFC may support the regulation of intergroup affect, such as threat or contempt, although this hypothesis remains to be tested. Considered broadly, emerging evidence regarding mPFC function suggests that it has a larger role in cognitive control than previously thought, particularly in the context of regulating complex social responses.

A network for the regulation of prejudice and stereotyping. Self-regulation is critical for the adaptive expression of social behaviour. This is especially true with regard to stereotyping and prejudice given the potential (unwanted) influence of implicit biases on behaviour. The putative network of neural regions involved in the regulation of intergroup responses includes ACC and PFC regions that have been implicated in existing models of cognitive control, as well as additional regions that facilitate control in social contexts (FIG. 5). Specifically, although the dorsal ACC and lateral PFC may carry out the domain-general functions of detecting conflict and implementing top-down control, the mPFC and rostral ACC are important for guiding control that is based on social cues, such as norms against expressing prejudice, and intergroup emotions.

From brain to society

Research on the neural basis of prejudice occupies a special position at the interface of the natural and social sciences, and as such it is uniquely situated to bring neuroscientific advances to bear on real-world social issues. To date, the neuroscientific analysis of prejudice has advanced theories of how prejudices are formed, expressed and potentially controlled, and these can be used to inform interventions aimed at reducing discrimination. For example, research linking implicit prejudice and stereotyping to different neural substrates suggests

that these two forms of bias are subserved by different learning and memory systems — a clue that interventions to reduce prejudice and stereotyping may require different approaches^{10,28}. Implicit prejudice has been linked to fear conditioning involving the amygdala, whereas stereotype associations appear to reflect conceptual learning systems involving the temporal cortex and PFC. Importantly, learning and expression differ considerably between these systems: fear conditioning may be acquired in a single trial and expressed primarily through behavioural freezing, anxiety and heightened vigilance¹⁸, whereas conceptual associations require many exposures for acquisition and are expressed through high-level representations of impressions and goal-directed actions¹⁴⁰. Social cognition studies have begun to adopt intervention strategies that are consistent with this analysis, using tasks in which images of racial outgroup members are repeatedly paired with positive images and appetitive responses or with counter-stereotypical concepts, in an effort to selectively target the affective or semantic memory systems underlying implicit prejudices and stereotypes, respectively^{29,30,141}. By considering the operations of these different neural systems, researchers are gaining a better understanding of how and under what conditions different forms of bias are activated, expressed and potentially extinguished.

Despite some success in reducing behavioural and physiological expressions of implicit bias in the laboratory^{29,30,141,142}, most forms of implicit learning are resistant to extinction^{140,143}. Implicit racial biases are particularly difficult to change in a cultural milieu that constantly reinforces racial prejudices and stereotypes (for example, in mainstream media). Thus, although attempts to undo learned intergroup associations are laudable, such strategies may be ineffective for reducing the expression of bias in behaviour outside the laboratory. Instead, interventions that enhance the cognitive control of behaviour should be more effective. Such control-based strategies may not reduce prejudice in the mind, but they can prevent its effect on potential victims. Over time, control-driven changes in behaviour may become habitual, and prejudiced and stereotypical associations in the mind may weaken^{7,144}. Neuroscience models suggest that control-based interventions should focus on (at least) two separate processes: those for monitoring unwanted racially biased tendencies and those involved in the top-down control of behaviour. Strategies to enhance ACC-mediated conflict-monitoring processes include interventions that increase people's awareness of the potential for bias, increase attention to specific cues indicating that control may be needed (for example, the appearance of an outgroup member in an interaction) and increase the sensitivity of conflict monitoring systems (for example, by activating cognitive conflict prior to an intergroup response)^{145,146}. For the effective control of behaviour, conflict monitoring processes must be paired with top-down response plans. To this end, psychological research has shown that goal strategies that link a specific cue (for example, 'if I meet a black person') with a pre-planned response (for example, 'I will ignore his or her race' or 'I will respond more carefully') are especially

effective at facilitating the control of implicit stereotypes in behaviour¹⁴⁷. By helping to inspire and explicate strategies such as these, the neuroscience of prejudice is already beginning to inform policy and interventions aimed at reducing prejudice in society.

Beyond its implications for social issues regarding intergroup relations, research on the neural basis of prejudice provides a context for understanding neural function as it relates to the real-world lives of human beings. Many areas of social neuroscience consider the

effects of social factors on neural function, but the neuroscience of prejudice is a particularly rich topic as it considers the roles of personal attitudes and motivations, social norms and social emotions as they relate to complex interpersonal behaviours. If the human brain has evolved to support survival and prosperity in a complex social environment, then a research approach that considers this range of factors will be needed to truly understand neural function. Research on the neural basis of prejudice is an important step in this direction.

1. Brewer, M. B. The psychology of prejudice: ingroup love and outgroup hate? *J. Soc. Issues* **55**, 429–444 (1999).
2. Ito, T. A. & Urland, G. R. Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *J. Pers. Soc. Psychol.* **85**, 616–626 (2003).
3. Ratner, K. G. & Amodio, D. M. Seeing “us versus them”: minimal group effects on the neural encoding of faces. *J. Exp. Soc. Psychol.* **49**, 298–301 (2013).
4. Allport, G. W. *The Nature of Prejudice* (Addison-Wesley, 1954).
5. Fiske, S. T. in *The Handbook of Social Psychology* Vol. 2, 4th edn (eds Gilbert, D. T., Fiske, S. T. & Lindzey, G.) 357–411 (McGraw Hill, 1998).
6. Amodio, D. M. & Devine, P. G. Stereotyping and evaluation in implicit race bias: evidence for independent constructs and unique effects on behavior. *J. Pers. Soc. Psychol.* **91**, 652–661 (2006).
7. Devine, P. G. Stereotypes and prejudice: their automatic and controlled components. *J. Pers. Soc. Psychol.* **56**, 5–18 (1989).
8. Fazio, R. H., Jackson, J. R., Dunton, B. C. & Williams, C. J. Variability in automatic activation as an unobtrusive measure of racial attitudes: a *bona fide* pipeline? *J. Pers. Soc. Psychol.* **69**, 1013–1027 (1995).
9. Greenwald, A. G. & Banaji, M. R. Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychol. Rev.* **102**, 4–27 (1995).
10. Amodio, D. M. The social neuroscience of intergroup relations. *Eur. Rev. Soc. Psychol.* **19**, 1–54 (2008).
11. Ito, T. A. & Bartholow, B. D. The neural correlates of race. *Trends Cogn. Sci.* **13**, 524–531 (2009). **In a field dominated by fMRI studies, this review presents important research on ERP approaches to probing the neural underpinnings of sociocognitive processes involved in prejudice and stereotyping.**
12. Kubota, J. T., Banaji, M. R. & Phelps, E. A. The neuroscience of race. *Nature Neurosci.* **15**, 940–948 (2012).
13. Amodio, D. M. Can neuroscience advance social psychological theory? Social neuroscience for the behavioral social psychologist. *Soc. Cogn.* **28**, 695–716 (2010).
14. Poldrack, R. A. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* **10**, 59–63 (2006).
15. Mackie, D. M. & Smith, E. R. in *The Social Self: Cognitive, Interpersonal, and Intergroup Perspectives* (eds Forgas, J. P. & Williams, K. D.) 309–326 (Psychology Press, 2002).
16. Cottrell, C. A. & Neuberg, S. L. Different emotional reactions to different groups: a sociofunctional threat-based approach to “prejudice”. *J. Pers. Soc. Psychol.* **88**, 770–789 (2005).
17. Swanson, L. W. & Petrovich, G. D. What is the amygdala? *Trends Neurosci.* **21**, 323–331 (1998).
18. LeDoux, J. E. Emotion circuits in the brain. *Annu. Rev. Neurosci.* **23**, 155–184 (2000).
19. Davis, M. Neural systems involved in fear and anxiety measured with fear-potentiated startle. *Am. Psychol.* **61**, 741–756 (2006).
20. Fendt, M. & Fanselow, M. S. The neuroanatomical and neurochemical basis of conditioned fear. *Neurosci. Biobehav. Rev.* **23**, 743–760 (1999).
21. LeDoux, J. E., Iwata, J., Cicchetti, P. & Reis, D. J. Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *J. Neurosci.* **8**, 2517–2529 (1988).
22. Holland, P. C. & Gallagher, M. Amygdala circuitry in attentional and representational processes. *Trends Cogn. Sci.* **3**, 65–73 (1999).
23. Rolls, E. T. & Rolls, B. J. Altered food preferences after lesions in the basolateral region of the amygdala in the rat. *J. Comp. Physiol. Psychol.* **83**, 248–259 (1973).
24. Adolphs, R. Fear, faces, and the human amygdala. *Curr. Opin. Neurobiol.* **18**, 166–172 (2008).
25. Phelps, E. A. *et al.* Performance on indirect measures of race evaluation predicts amygdala activation. *J. Cogn. Neurosci.* **12**, 729–738 (2000). **An early fMRI investigation of neural correlates of implicit prejudice, focusing on the amygdala. Relative differences in amygdala activity in response to black compared with white faces were associated with a behavioural index of implicit prejudice and startle responses to black versus white faces.**
26. Hart, A. J. *et al.* Differential response in the human amygdala to racial outgroup versus ingroup face stimuli. *Neuroreport* **11**, 2351–2355 (2000).
27. Amodio, D. M., Harmon-Jones, E. & Devine, P. G. Individual differences in the activation and control of affective race bias as assessed by startle eyeblink responses and self-report. *J. Pers. Soc. Psychol.* **84**, 738–753 (2003). **The first reported difference in amygdala activity in response to black compared with white (and Asian) faces. By using the startle-eyeblink method to assess amygdala activity related to the CeA, the results link implicit prejudice more directly to a Pavlovian conditioning mechanism of learning and behavioural expression.**
28. Amodio, D. M. & Ratner, K. G. A memory systems model of implicit social cognition. *Curr. Dir. Psychol. Sci.* **20**, 143–148 (2011).
29. Olson, M. A. & Fazio, R. H. Reducing automatically-activated racial prejudice through implicit evaluative conditioning. *Pers. Soc. Psychol. Bull.* **32**, 421–433 (2006).
30. Kawakami, K., Phills, C. E., Steele, J. R. & Dovidio, J. F. (Close) distance makes the heart grow fonder: improving implicit racial attitudes and interracial interactions through approach behaviors. *J. Pers. Soc. Psychol.* **92**, 957–971 (2007).
31. Amodio, D. M. & Hamilton, H. K. Intergroup anxiety effects on implicit racial evaluation and stereotyping. *Emotion* **12**, 1273–1280 (2012).
32. Chekroud, A. M., Everett, J. A. C., Bridge, H. & Hewstone, M. A review of neuroimaging studies of race-related prejudice: does amygdala response reflect threat? *Front. Hum. Neurosci.* **8**, 179 (2014).
33. Ronquillo, J. *et al.* The effects of skin tone on race-related amygdala activity: an fMRI investigation. *Soc. Cogn. Affect. Neurosci.* **2**, 39–44 (2007).
34. Richeson, J. A., Todd, A. R., Trawalter, S. & Baird, A. A. Eye-gaze direction modulates race-related amygdala activity. *Group Process. Interg. Relat.* **11**, 235–246 (2008).
35. Freeman, J. B., Schiller, D., Rule, N. O. & Ambady, N. The neural origins of superficial and individuated judgments about ingroup and outgroup members. *Hum. Brain Mapp.* **31**, 150–159 (2010).
36. Forbes, C. E., Cox, C. L., Schmader, T. & Ryan, L. Negative stereotype activation alters interaction between neural correlates of arousal, inhibition, and cognitive control. *Soc. Cogn. Affect. Neurosci.* **7**, 771–781 (2012).
37. Cunningham, W. A. *et al.* Separable neural components in the processing of black and white faces. *Psychol. Sci.* **15**, 806–813 (2004).
38. Telzer, E. H., Humphreys, K., Shapiro, M. & Tottenham, N. L. Amygdala sensitivity to race is not present in childhood but emerges in adolescence. *J. Cogn. Neurosci.* **25**, 234–244 (2013).
39. Cloutier, J., Li, T. & Correll, J. The impact of childhood experience on amygdala response to perceptually familiar black and white faces. *J. Cogn. Neurosci.* **26**, 1992–2004 (2014).
40. Olsson, A., Ebert, J. P., Banaji, M. R. & Phelps, E. A. The role of social group in the persistence of learned fear. *Science* **309**, 785–787 (2005).
41. Richeson, J. A. & Trawalter, S. The threat of appearing prejudiced and race-based attentional biases. *Psychol. Sci.* **19**, 98–102 (2008).
42. Ofan, R. H., Rubin, N. & Amodio, D. M. Situation-based social anxiety enhances the neural processing of faces: evidence from an intergroup context. *Soc. Cogn. Affect. Neurosci.* <http://dx.doi.org/10.1093/scan/nst087> (2013).
43. Plant, E. A. & Devine, P. G. Internal and external motivation to respond without prejudice. *J. Pers. Soc. Psychol.* **75**, 811–832 (1998).
44. Stephan, W. G. & Stephan, C. W. Intergroup anxiety. *J. Soc. Issues* **41**, 157–175 (1985).
45. Lieberman, M. D., Hariri, A., Jarcho, J. M., Eisenberger, N. I. & Bookheimer, S. Y. An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nature Neurosci.* **8**, 720–722 (2005).
46. Wheeler, M. E. & Fiske, S. T. Controlling racial prejudice: social-cognitive goals affect amygdala and stereotype activation. *Psychol. Sci.* **16**, 56–63 (2005).
47. Van Bavel, J. J., Packer, D. J. & Cunningham, W. A. The neural substrates of in-group bias: a functional magnetic resonance imaging investigation. *Psychol. Sci.* **11**, 1131–1139 (2008). **This study compared the independent effects of race and coalition (team membership) on amygdala responses to faces and showed that when team membership was salient, the amygdala responded to coalition (subject’s team or other team) and not race (black or white). This finding demonstrated that the amygdala response to race is not inevitable but rather corresponds to the subject’s particular task goal.**
48. Richeson, J. A. *et al.* An fMRI investigation of the impact of interracial contact on executive function. *Nature Neurosci.* **6**, 1323–1328 (2003).
49. Gilbert, S. J., Swencionis, J. K. & Amodio, D. M. Evaluative versus trait representation in intergroup social judgments: distinct roles of anterior temporal lobe and prefrontal cortex. *Neuropsychologia* **50**, 3600–3611 (2012). **Using fMRI, the authors distinguished neural processes involved in stereotyping and prejudiced attitudes. Stereotype-based judgements of black compared with white people uniquely involved the mPFC; affect-based judgements uniquely involved the OFC.**
50. Golby, A. J., Gabrieli, J. D. E., Chiao, J. Y. & Eberhardt, J. L. Differential fusiform responses to same- and other-race faces. *Nature Neurosci.* **4**, 845–850 (2001). **The first demonstration of race effects in visual processing, using fMRI. Differences in fusiform responses to black and white faces predicted the degree of ‘own race bias’ in memory in white subjects.**
51. Knutson, K. M., Mah, L., Manly, C. F. & Grafman, J. Neural correlates of automatic beliefs about gender and race. *Hum. Brain Mapp.* **28**, 915–930 (2007).
52. Brosch, T., Bar-David, E. & Phelps, E. A. Implicit race bias decreases the similarity of neural representations of black and white faces. *Psychol. Sci.* **24**, 160–166 (2013).
53. Schreiber, D. & Iacoboni, M. Huxtables on the brain: an fMRI study of race and norm violation. *Polit. Psychol.* **33**, 313–330 (2012).
54. Whalen, P. J. Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. *Curr. Dir. Psychol. Sci.* **7**, 177–188 (1998).

55. Said, C. P., Baron, S. & Todorov, A. Nonlinear amygdala response to face trustworthiness: contributions of high and low spatial frequency information. *J. Cogn. Neurosci.* **21**, 519–528 (2009).
56. Cunningham, W. A., Van Bavel, J. J. & Johnsen, I. R. Affective flexibility: evaluative processing goals shape amygdala activity. *Psychol. Sci.* **19**, 152–160 (2008).
57. Phelps, E. A., Cannistraci, C. J. & Cunningham, W. A. Intact performance on an indirect measure of race bias following amygdala damage. *Neuropsychologia* **41**, 203–208 (2003).
58. Bechara, A., Damasio, H. & Damasio, A. R. Emotion, decision making and the orbitofrontal cortex. *Cereb. Cortex* **10**, 295–307 (2000).
59. O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J. & Andrews, C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neurosci.* **4**, 95–102 (2001).
60. Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E. & Behrens, T. E. Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**, 1054–1069 (2011).
61. Beer, J. S., Heerey, E. H., Keltner, D., Scabini, D. & Knight, R. T. The regulatory function of self-conscious emotion: insights from patients with orbitofrontal damage. *J. Pers. Soc. Psychol.* **85**, 594–604 (2003).
62. Amodio, D. M. & Frith, C. D. Meeting of minds: the medial frontal cortex and social cognition. *Nature Rev. Neurosci.* **7**, 268–277 (2006).
63. Ongür, D. & Price, J. L. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* **10**, 206–219 (2000).
64. Beer, J. S. *et al.* The Quadruple Process model approach to examining the neural underpinnings of prejudice. *Neuroimage* **43**, 775–783 (2008).
65. Craig, A. D. How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Rev. Neurosci.* **3**, 655–666 (2002).
66. Craig, A. D. How do you feel — now? The anterior insula and human awareness. *Nature Rev. Neurosci.* **10**, 59–70 (2009).
67. Harris, L. T. & Fiske, S. T. Dehumanizing the lowest of the low: neuro-imaging responses to extreme outgroups. *Psychol. Sci.* **17**, 847–853 (2006).
68. Singer, T. *et al.* Empathy for pain involves the affective but not sensory components of pain. *Science* **303**, 1157–1162 (2004).
69. Singer, T. *et al.* Empathic neural responses are modulated by the perceived fairness of others. *Nature* **439**, 466–469 (2006).
70. Lamm, C., Meltzoff, A. N. & Decety, J. How do we empathize with someone who is not like us? A functional magnetic resonance imaging study. *J. Cogn. Neurosci.* **22**, 362–376 (2010).
71. Xu, X., Zuo, X., Wang, X. & Han, S. Do you feel my pain? Racial group membership modulates empathic neural responses. *J. Neurosci.* **29**, 8525–8529 (2009).
- In a demonstration of racial bias in empathy, mPFC and ACC activity linked to empathy was observed only when subjects viewed members of their own racial group being exposed to painful stimuli.**
72. Cikara, M. & Fiske, S. T. Bounded empathy: neural responses to outgroup targets' (mis)fortunes. *J. Cogn. Neurosci.* **23**, 3791–3803 (2011).
73. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381 (1986).
74. Knutson, B., Adams, C. S., Fong, G. W. & Hommer, D. Anticipation of monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* **21**, RC159 (2001).
75. O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452–454 (2004).
76. Stanley, D. A. *et al.* Race and reputation: perceived racial group trustworthiness influences the neural correlates of trust decisions. *Phil. Trans. R. Soc. B* **367**, 744–753 (2012).
77. Mitchell, J. P., Heatherton, T. F. & Macrae, C. N. Distinct neural systems subserve person and object knowledge. *Proc. Natl Acad. Sci. USA* **99**, 15238–15243 (2002).
78. Van Overwalle, F. Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* **30**, 829–858 (2009).
79. Krueger, F., Barbey, A. K. & Grafman, J. The medial prefrontal cortex mediates social event knowledge. *Trends Cogn. Sci.* **13**, 103–109 (2009).
80. Frith, C. D. & Frith, U. Interacting minds—a biological basis. *Science* **286**, 1692–1695 (1999).
81. Cikara, M., Bruneau, E. G. & Saxe, R. Us and them: intergroup failures of empathy. *Curr. Dir. Psychol. Sci.* **20**, 149–153 (2011).
82. Cikara, M., Eberhardt, J. L. & Fiske, S. T. From agents to objects: sexist attitudes and neural responses to sexualized targets. *J. Cogn. Neurosci.* **23**, 540–551 (2011).
83. Hamilton, D. L. & Sherman, J. W. in *Handbook of Social Cognition* Vol. 2, 2nd edn (eds Wyer, R. S. Jr & Srull, T. K.) 1–68 (Erlbaum, 1994).
84. Olson, I. R., McCoy, D., Klobusicky, E. & Ross, L. A. Social cognition and the anterior temporal lobes: a review and theoretical framework. *Soc. Cogn. Affect. Neurosci.* **8**, 123–133 (2013).
85. Quadflieg, S. & Macrae, C. N. Stereotypes and stereotyping: what's the brain got to do with it? *Eur. Rev. Soc. Psychol.* **22**, 215–273 (2011).
- A comprehensive review of neuroscience research on social stereotyping processes.**
86. Gabrieli, J. D. Cognitive neuroscience of human memory. *Annu. Rev. Psychol.* **49**, 87–115 (1998).
87. Martin, A. The representation of object concepts in the brain. *Annu. Rev. Psychol.* **58**, 25–45 (2007).
88. Patterson, K., Nestor, P. J. & Rogers, T. T. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Rev. Neurosci.* **8**, 976–987 (2007).
89. Quadflieg, S. *et al.* Exploring the neural correlates of social stereotyping. *J. Cogn. Neurosci.* **21**, 1560–1570 (2009).
90. Olson, I. R., Plotzker, A. & Ezzyat, Y. The enigmatic temporal pole: a review of findings on social and emotional processing. *Brain* **130**, 1718–1731 (2007).
91. Zahn, R. *et al.* Social concepts are represented in the superior anterior temporal cortex. *Proc. Natl Acad. Sci. USA* **104**, 6430–6435 (2007).
92. de Schotten, M. T., Dell'Acqua, F., Valabregue, R. & Catani, M. Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex* **48**, 82–96 (2012).
93. Contreras, J. M., Banaji, M. R. & Mitchell, J. P. Dissociable neural correlates of stereotypes and other forms of semantic knowledge. *Soc. Cogn. Affect. Neurosci.* **7**, 764–770 (2012).
- An fMRI study on the neural underpinnings of stereotyping that distinguishes the role of trait inference processes from that of object knowledge representation.**
94. Gallate, J., Wong, C., Ellwood, S., Chi, R. & Snyder, A. Noninvasive brain stimulation reduces prejudice scores on an implicit association test. *Neuropsychology* **25**, 185–192 (2011).
95. Milne, E. & Grafman, J. Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *J. Neurosci.* **21**, 1–6 (2001).
96. Mitchell, J. P., Macrae, C. N. & Banaji, M. R. Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron* **50**, 655–663 (2006).
97. Quadflieg, S. *et al.* Stereotype-based modulation of person perception. *Neuroimage* **57**, 549–557 (2011).
98. Forbes, C. E. *et al.* Identifying temporal and causal contributions of neural processes underlying the Implicit Association Test. (IAT). *Front. Hum. Neurosci.* **6**, 320 (2012).
99. Mitchell, J. P. Social psychology as a natural kind. *Trends Cogn. Sci.* **13**, 246–251 (2009).
100. Gilbert, S. J. *et al.* Distinct regions of medial rostral prefrontal cortex supporting social and nonsocial functions. *Soc. Cogn. Affect. Neurosci.* **2**, 217–226 (2007).
101. Baetens, K., Ma, N., Steen, J. & van Overwalle, F. Involvement of the mentalizing network in social and non-social high construal. *Soc. Cogn. Affect. Neurosci.* **9**, 817–824 (2014).
102. Saxe, R. Uniquely human social cognition. *Curr. Opin. Neurobiol.* **16**, 235–239 (2006).
103. Frith, C. D. & Frith, U. Implicit and explicit processes in social cognition. *Neuron* **60**, 503–510 (2008).
104. Uddin, L. Q., Iacoboni, M., Lange, C. & Keenan, J. P. The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends Cogn. Sci.* **11**, 153–157 (2007).
105. Blaxton, T. A. *et al.* Functional mapping of human memory using PET: comparisons of conceptual and perceptual tasks. *Can. J. Exp. Psychol.* **50**, 42–56 (1996).
106. Buckner, R. L. & Tulving, E. in *Handbook of Neuropsychology* Vol. 10 (eds Boller, F. & Grafman, J.) 439–466 (Elsevier, 1995).
107. Demb, J. B. *et al.* Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and process specificity. *J. Neurosci.* **15**, 5870–5878 (1995).
108. Thompson-Schill, S. L. Neuroimaging studies of semantic memory: inferring "how" from "where". *Neuropsychologia* **41**, 280–292 (2003).
109. Miller, E. K., Freedman, D. J. & Wallis, J. D. The prefrontal cortex: categories, concepts and cognition. *Phil. Trans. R. Soc. Lond. B* **357**, 1123–1136 (2002).
110. Balleine, B. W., Delgado, M. R. & Hikosaka, O. The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* **27**, 8161–8165 (2007).
111. Mitchell, J. P., Ames, D. L., Jenkins, A. C. & Banaji, M. R. Neural correlates of stereotype application. *J. Cogn. Neurosci.* **21**, 594–604 (2009).
112. Aron, A. R., Robbins, T. W. & Poldrack, R. A. Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* **8**, 170–177 (2004).
113. Bodenhausen, G. V., Sheppard, L. & Kramer, G. P. Negative affect and social perception: the differential impact of anger and sadness. *Eur. J. Soc. Psychol.* **24**, 45–62 (1994).
114. DeSteno, D., Dasgupta, N., Bartlett, M. Y. & Caidric, A. Prejudice from thin air: the effect of emotion on automatic intergroup attitudes. *Psychol. Sci.* **15**, 319–324 (2004).
115. Choi, E. Y., Yeo, B. T. & Buckner, R. L. The organization of the human striatum estimated by intrinsic functional connectivity. *J. Neurophysiol.* **108**, 2242–2263 (2012).
116. Bush, G., Luu, P. & Posner, M. L. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* **4**, 215–222 (2000).
117. Mansouri, F. A., Tanaka, K. & Buckley, M. J. Conflict-induced behavioral adjustment: a clue to the executive functions of the prefrontal cortex. *Nature Rev. Neurosci.* **10**, 141–152 (2009).
118. Gerhing, W. J., Goss, B., Coles, M. G. H., Meyer, D. E. & Donchin, E. A neural system for error detection and compensation. *Psychol. Sci.* **4**, 385–390 (1993).
119. Botvinick, M., Nystrom, L., Fissell, K., Carter, C. & Cohen, J. Conflict monitoring versus selection-foraction in anterior cingulate cortex. *Nature* **402**, 179–181 (1999).
120. Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652 (2001).
121. Shackman, A. J. *et al.* The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Rev. Neurosci.* **12**, 154–167 (2011).
122. Amodio, D. M. & Devine, P. G. in *Self Control in Society, Mind and Brain* (eds Hassin, R. R., Ochsner, K. N. & Trope, Y.) 49–75 (Oxford Univ. Press, 2010).
123. Amodio, D. M. *et al.* Neural signals for the detection of unintentional race bias. *Psychol. Sci.* **15**, 88–93 (2004).
- This study demonstrated the role of the ACC in the control of racial bias. Using ERPs, the authors found an increase ACC activity to stereotype-based conflict that predicted enhanced control of bias in behaviour.**
124. Amodio, D. M., Kubota, J. T., Harmon-Jones, E. & Devine, P. G. Alternative mechanisms for regulating racial responses according to internal versus external cues. *Soc. Cogn. Affect. Neurosci.* **1**, 26–36 (2006).
125. Amodio, D. M., Devine, P. G. & Harmon-Jones, E. Individual differences in the regulation of intergroup bias: the role of conflict monitoring and neural signals for control. *J. Pers. Soc. Psychol.* **94**, 60–74 (2008).
126. Bartholow, B. D., Dickter, C. L. & Sestiri, M. A. Stereotype activation and control of race bias: cognitive control of inhibition and its impairment by alcohol. *J. Pers. Soc. Psychol.* **90**, 272–287 (2006).
127. Correll, J., Urland, G. R. & Ito, T. A. Event-related potentials and the decision to shoot: the role of threat perception and cognitive control. *J. Exp. Soc. Psychol.* **42**, 120–128 (2006).
128. Gonsalkorale, K., Sherman, J. W., Allen, T. J., Klauer, K. C. & Amodio, D. M. Accounting for successful control of implicit racial bias: the roles of association activation, response monitoring, and overcoming bias. *Pers. Soc. Psychol. Bull.* **37**, 1534–1545 (2011).

129. Fourie, M. M., Thomas, K. G. F., Amodio, D. M., Warton, C. M. R. & Meintjes, E. M. Neural correlates of experienced moral emotion: an fMRI investigation of emotion in response to prejudice feedback. *Soc. Neurosci.* **9**, 203–218 (2014).
130. Van Nunspeet, F., Ellemers, N., Derks, B. & Nieuwenhuis, S. Moral concerns increase attention and response monitoring during IAT performance: ERP evidence. *Soc. Cogn. Affect. Neurosci.* **9**, 141–149 (2014).
131. Fuster, J. M. The prefrontal cortex—an update: time is of the essence. *Neuron* **2**, 319–333 (2001).
132. Badre, D. & D'Esposito, M. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Rev. Neurosci.* **10**, 659–669 (2009).
133. Koechlin, E., Ody, C. & Kouneiher, F. The architecture of cognitive control in the human prefrontal cortex. *Science* **302**, 1181–1185 (2003).
134. Aron, A. R. The neural basis of inhibition in cognitive control. *Neuroscientist* **13**, 214–228 (2007).
135. Ghashghaei, H. T., Hilgetag, C. C. & Barbas, H. Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *Neuroimage* **34**, 905–923 (2007).
136. Amodio, D. M. Coordinated roles of motivation and perception in the regulation of intergroup responses: frontal cortical asymmetry effects on the P2 event-related potential and behavior. *J. Cogn. Neurosci.* **22**, 2609–2617 (2010).
This study demonstrated the role of the PFC in the behavioural control of racial bias and showed that this effect involves changes in early perceptual attention to racial cues. The study provided initial evidence for a 'motivated perception' account of self-regulation.
137. Amodio, D. M., Devine, P. G. & Harmon-Jones, E. A dynamic model of guilt: implications for motivation and self-regulation in the context of prejudice. *Psychol. Sci.* **18**, 524–530 (2007).
138. Gozzi, M., Raymond, V., Solomon, J., Koenigs, M. & Grafman, J. Dissociable effects of prefrontal and anterior temporal cortical lesions on stereotypical gender attitudes. *Neuropsychologia* **47**, 2125–2132 (2009).
139. Milad, M. R. & Quirk, G. J. Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* **420**, 70–74 (2002).
140. Sloman, S. A. The empirical case for two systems of reasoning. *Psychol. Bull.* **119**, 3–22 (1996).
141. Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S. & Russin, A. Just say no (to stereotyping): effects of training on the negation of stereotypic associations on stereotype activation. *J. Pers. Soc. Psychol.* **78**, 871–888 (2000).
142. Mallan, K. M., Sax, J. & Lipp, O. V. Verbal instruction abolishes fear conditioned to racial out-group faces. *J. Exp. Soc. Psychol.* **45**, 1303–1307 (2009).
143. Bouton, M. E. Conditioning, remembering, and forgetting. *J. Exp. Psychol. Anim. B* **20**, 219–231 (1994).
144. Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E. & Vance, S. L. The regulation of explicit and implicit racial bias: the role of motivations to respond without prejudice. *J. Pers. Soc. Psychol.* **82**, 835–848 (2002).
145. Kleiman, T., Hassin, R. R. & Trope, Y. The control-freak mind: stereotypical biases are eliminated following conflict-activated cognitive control. *J. Exp. Psychol. Gen.* **143**, 498–505 (2014).
146. Monteith, M. J., Ashburn-Nardo, L., Voils, C. I. & Czopp, A. M. Putting the brakes on prejudice: on the development and operation of cues for control. *J. Pers. Soc. Psychol.* **83**, 1029–1050 (2002).
147. Mendoza, S. A., Gollwitzer, P. M. & Amodio, D. M. Reducing the expression of implicit stereotypes: reflexive control through implementation intentions. *Pers. Soc. Psychol. Bull.* **36**, 512–525 (2010).
148. Bar, M. et al. Top-down facilitation of visual recognition. *Proc. Natl Acad. Sci. USA* **103**, 449–454 (2006).
149. Van Bavel, J. J., Packer, D. J. & Cunningham, W. A. Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: evidence of in-group enhancement (not out-group disregard). *J. Cogn. Neurosci.* **23**, 3343–3354 (2011).
150. Ratner, K. G., Dotsch, R., Wigboldus, D., van Knippenberg, A. & Amodio, D. M. Visualizing minimal ingroup and outgroup faces: implications for impressions, attitudes, and behavior. *J. Pers. Soc. Psychol.* **106**, 897–911 (2014).
151. Krosch, A. K. & Amodio, D. M. Economic scarcity alters the perception of race. *Proc. Natl Acad. Sci. USA* **111**, 9079–9084 (2014).
152. Ofan, R. H., Rubin, N. & Amodio, D. M. Seeing race: N170 responses to race and their relation to automatic racial attitudes and controlled processing. *J. Cogn. Neurosci.* **23**, 3152–3161 (2011).
153. Brebner, J. L., Krigolson, O., Handy, T. C., Quadflieg, S. & Turk, D. J. The importance of skin color and facial structure in perceiving and remembering others: an electro-physiological study. *Brain Res.* **1388**, 123–133 (2011).
154. Caldara, R., Rossion, B., Bovet, P. & Hauert, C. A. Event-related potentials and time course of the 'other-race' face classification advantage. *Neuroreport* **15**, 905–910 (2004).
155. Caldara, R. et al. Face versus non-face object perception and the 'other-race' effect: a spatiotemporal event-related potential study. *Clin. Neurophysiol.* **114**, 515–528 (2003).
156. Ito, T. A. & Urland, G. R. The influence of processing objectives on the perception of faces: an ERP study of race and gender perception. *Cogn. Affect. Behav. Neurosci.* **5**, 21–36 (2005).
157. Walker, P. M., Silvert, L., Hewstone, M. & Nobre, A. C. Social contact and other-race face processing in the human brain. *Soc. Cogn. Affect. Neurosci.* **3**, 16–25 (2008).
158. Wiese, H., Stahl, J. & Schweinberger, S. R. Configural processing of other-race faces is delayed but not decreased. *Biol. Psychol.* **81**, 103–109 (2009).
159. Contreras, J. M., Banaji, M. R. & Mitchell, J. P. Multivoxel patterns in fusiform face area differentiate faces by sex and race. *PLoS ONE* **8**, e69684 (2013).
160. Ratner, K. G., Kaul, C. & Van Bavel, J. J. Is race erased? Decoding race from patterns of neural activity when skin color is not diagnostic of group boundaries. *Soc. Cogn. Affect. Neurosci.* **8**, 750–755 (2013).
161. Greenwald, A. G., Poehlman, T. A., Uhlmann, E. & Banaji, M. R. Understanding and using the Implicit Association Test: III. meta-analysis of predictive validity. *J. Pers. Soc. Psychol.* **97**, 17–41 (2009).
162. Payne, B. K. Prejudice and perception: the role of automatic and controlled processes in misperceiving a weapon. *J. Pers. Soc. Psychol.* **81**, 181–192 (2001).
163. Greenwald, A. G., McGhee, D. E. & Schwartz, J. K. L. Measuring individual differences in implicit cognition: the Implicit Association Test. *J. Pers. Soc. Psychol.* **74**, 1464–1480 (1998).

Acknowledgements

Work on this article was supported by a National Science Foundation grant (BCS 0847350). The author thanks members of the NYU Social Neuroscience Laboratory, J. Freeman, K. Ratner and the three anonymous reviewers for their helpful feedback on earlier versions of this article.

Competing interests statement

The author declares no competing interests.