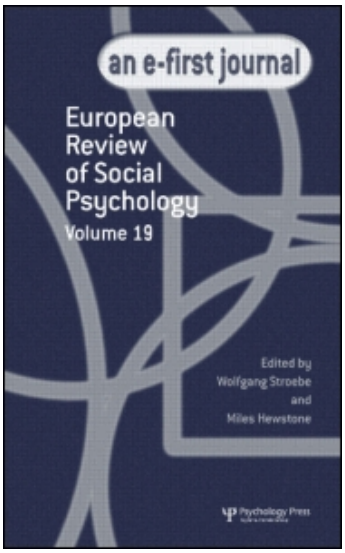

European Review of Social Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713684724

Social neuroscience evidence for dehumanised perception

Lasana T. Harris [a]; Susan T. Fiske [b]
[a] New York University, NY, USA [b] Princeton University, NJ, USA

First Published on: 15 June 2009

To cite this Article Harris, Lasana T. and Fiske, Susan T.(2009)'Social neuroscience evidence for dehumanised perception',European Review of Social Psychology,20:1,192 — 231
To link to this Article: DOI: 10.1080/10463280902954988
URL: http://dx.doi.org/10.1080/10463280902954988

EUROPEAN REVIEW OF SOCIAL PSYCHOLOGY
2009, 20, 192–231

Psychology Press
Taylor & Francis Group

Social neuroscience evidence for dehumanised perception

Lasana T. Harris
New York University, NY, USA

Susan T. Fiske
Princeton University, NJ, USA

Dehumanisation describes perceiving a person as nonhuman in some ways, such as lacking a mind. Social psychology is beginning to understand cognitive and affective causes and mechanisms—the psychological how and why of dehumanisation. Social neuroscience research also can inform these questions. After background on social neural networks and on past dehumanisation research, the article contrasts (a) research on *fully humanised* person perception, reviewing studies on affective and cognitive factors, specifically mentalising (considering another's mind), with (b) *dehumanised perception,* proposing neural systems potentially involved. Finally, the conclusion suggests limitations of social neuroscience, future research directions, and real-world consequences of this all-too-human phenomenon.

Keywords: Dehumanisation; Social neuroscience; Stereotype content model (SCM); Disgust; Morality.

Our perception makes category errors. People sometimes perceive other people as if they are animals or objects, and objects or animals as if they are people. Evolutionary advantages can accrue from anthropomorphism—perceiving animals or objects like people—given the memory benefits of social cognition (Johansson, Mecklinger, & Treese, 2004; Mason, Hood, & Macrae, 2004; Mason & Macrae, 2004; Meiser, 2003; Phelps, 2006a, 2006b; Phelps & LaBar, 2006; Plaks, Grant, & Dweck, 2005; Somerville, Wig, Whalen, & Kelley, 2006; Todd, Lewis, Meusel, & Zelazo, 2008; Todorov,

Correspondence should be addressed to Dr Lasana Harris, Department of Psychology, New York University, New York, USA. E-mail: lasana@nyu.edu

We thank Phil Goff and Jennifer Eberhardt for sharing their stimuli, and Rafael Escobedo and Jack Gelfand for their assistance and instruction in gathering the EEG data.

http://www.psypress.com/ersp DOI: 10.1080/10463280902954988

Gobbini, Evans, & Haxby, 2007; van Knippenberg, van Twuyver, & Pepels, 1994; von Hecker & Dutke, 2004). However, what possible evolutionary argument supports discarding those advantages and viewing a person like an object? Social psychological evidence suggests that all-too-human mechanisms may underlie these misperceptions (Fiske, Harris, & Cuddy, 2004), perhaps facilitating harm normally limited to non-human objects, destructive behaviour otherwise not permissible towards humans. Social neuroscience provides converging evidence for these mechanisms.

Violent, demeaning behaviours directed towards people certainly indicate dehumanisation.[1] This chapter aims to address this important topic by providing a framework that incorporates cognitive-affective neuroscience methodologies. After giving some background on social neural networks and on past dehumanisation research, this article contrasts research on (a) humanised person perception, reviewing studies on affective and cognitive factors, specifically mentalising (considering another's mind) with (b) *dehumanised perception,* proposing neural systems potentially involved. Finally, the conclusion suggests limitations of social neuroscience, future research directions, and real-world consequences of this all-too-human phenomenon.

A NOTE ON SOCIAL NEUROSCIENCE

Before continuing, a word about the social neuroscience approach this chapter employs: For social psychologists also fluent in cognitive-affective neuroscience, social neuroscience allows investigation of behavioural phenomena that are difficult to measure using traditional experimental social-psychological methods and measures. Social neuroscience—an interdisciplinary approach to social psychological questions (Lieberman, 2007; Ochsner, 2007; Ochsner & Lieberman, 2001; Todorov, Harris, & Fiske, 2006)—can incorporate social psychological theory, along with other philosophical, economic, legal, and political sources, then test these unique predictions using human neuroscience methods. As a premise for hypotheses, this chapter uses the folk psychological answer to the philosophical question "What is human?": that is, viewing human beings as entities with internal lives or minds (Adolphs, 2004). Although this approach is susceptible to the problems inherent to folk psychology, the social neuroscience approach adds physiological data, providing converging evidence for self-report.

Social neuroscience, as an emerging field, defines the intersection of social psychology, cognitive-affective psychology, and human neuroscience

[1]We follow an inter-group bias model, so we discuss dehumanisation, a form of extreme inter-group bias, distinguishing among cognition, affect, and behaviour. We use the term *dehumanisation* to describe behaviour, *dehumanising prejudice* to describe affect, and *dehumanised perception* to describe cognitions throughout this article.

(Lieberman, 2007; Ochsner, 2007; Ochsner & Lieberman, 2001; Todorov et al., 2006). It depends on social psychological theories and methods, and measures commonly used in cognitive-affective neuroscience. Considering social psychology alone may not be sufficient to identify all the nuances of a complex phenomenon such as dehumanisation. That being said, the social neuroscience here is much closer to social psychology than other areas of social neuroscience, partially because of its use of traditional social psychological designs. Thus, neuroscience is just one more tool to understand social psychological questions.

BACKGROUND ON SOCIAL NEURAL NETWORKS

To frame the reports of brain activation patterns relevant to social responses, this background section first describes possible social functions of some relevant brain regions. The most common approach employed to study human beings in neuroscience is physiological measurement, particularly neuro-imaging. Functional magnetic resonance imaging (fMRI) measures a correlate of neural activity, cerebral blood flow. This flow occurs at a 4–6-second lag after neural activity, and is not a direct measure of neural activity but of oxygenated blood rushing to clusters of neurons after firing. The spatial resolution, though not nearly as good as electrophysiology in animals, is currently the best available for unobtrusive human brains, but the haemodynamic lag makes it difficult to specify precise timing because it occurs after neural activity. Therefore this measure is approximate at best. However, neuro-imaging studies allow the correlation of questionnaire or behavioural measures with a correlate of neural activation. Thus, several converging strategies can be used to make claims about brain function.

Electroencephalography (EEG) and magneto-encephalography (MEG) more directly measure neural activity in time but not in location. When clusters of neurons fire, an electrical current is conducted to the scalp through tissue, cerebrospinal fluid, and skull. This current is measured at the scalp, giving good temporal resolution about psychological process, but poor spatial resolution of neural structure—it cannot specify the exact location of clusters of neurons. However, reliable event-related potentials (ERPs) do fire to psychological processes, including face perception, cognitive conflict, and cognitive errors, making it a more direct measure of neural firing than fMRI. Neuro-imaging and EEG are the two methods we adopt from neuroscience in this chapter. However, all the other tools of cognitive-affective neuroscience are available to the social neuroscientist to provide converging evidence for the occurrence of psychological phenomena, including facial electromyography, galvanic skin responses, heart rate monitoring, neuropharmacology, and lesion studies.

Neural systems of person perception and social cognition

Several neural regions figure prominently in person perception and social cognition. For social psychologists, one of the most important insights is that the neuroscience literature understands one of the medial prefrontal cortex's (MPFC) functions as serving a person perception network (e.g., Haxby, Gobbini, & Montgomery, 2004). Various aspects of this network activate in forming impressions, understanding another's false beliefs ("Theory of Mind"), attributing dispositions, perceiving close others, and various social cognition tasks (for reviews see Amodio & Frith, 2006; Olsson & Ochsner, 2008).

The subsequent sections will describe relevant neural systems. Person perception involves the pregenual anterior cingulate cortex (pACC) specifically, a sub-region of MPFC, which activates during face perception, along with other neural areas such as the amygdala, insula, superior temporal sulcus (STS), fusiform gyrus of temporal cortex, precuneus, and posterior cingulate (Haxby et al., 2004; see Figure 1). The next subsections summarise relevant research on other aspects of person perception, specifically mental state inference (mentalising), familiarity, and the self.

Mentalising. Mental state inference studies address people's ability to consider someone else's thoughts (see Frith & Frith, 2001). Paradigms often probe this mentalising ability with false belief tasks using cartoons or vignettes describing behaviour. These theory of mind (ToM) studies, along with studies of dispositional attribution (Harris, Todorov, & Fiske, 2005), reliably include the same brain regions, namely the MPFC, superior temporal sulci (STS), and right temporal-parietal junction (RTPJ). These areas implicated in person perception may behave differently in object and animal perception, as well as in dehumanised perception.

Familiarity. Familiarity breeds liking, which links rewards to habituated and fluent social perception. Imaging studies of familiarity in particular and positive social affect more generally illustrate that MPFC activates to both familiar and positive social stimuli. Significantly greater MPFC activity occurs when mothers look at faces of their own child rather than other familiar children, and for familiar children more than unfamiliar children (Leibenluft, Gobbini, Harrison, & Haxby, 2004).

As social psychology shows, familiarity itself links to positivity. Mere exposure demonstrates that simply repeated conscious (Zajonc, 1968) or unconscious (Monahan, Murphy, & Zajonc, 2000) exposure to a neutral stimulus enhances both subsequent liking and subjective familiarity of the

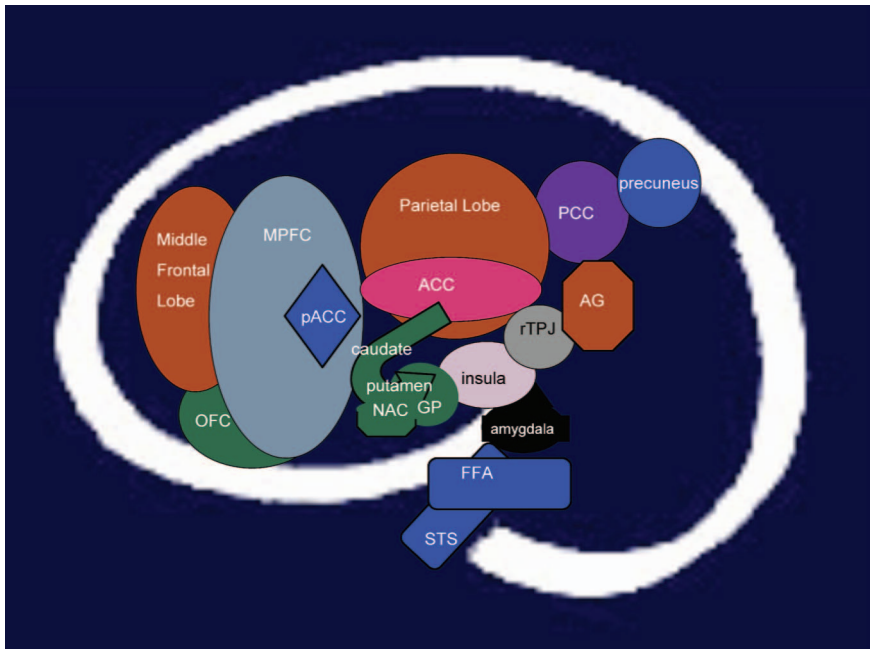

Figure 1. Neural regions involved in social cognition (see online for colour version). Many of these areas are implicated in more than one kind of social cognition, so the colour coding described below is a rough guide. Areas in blue tend to underlie person perception, specifically *medial prefrontal cortex* (MPFC) and *pregenual cingulate* (pACC) *amygdala*, *insula*, *superior temporal sulcus* (STS), *fusiform gyrus of temporal cortex* (FFA), *precuneus, posterior cingulate*, and occasionally *right temporal-parietal junction* (rTPJ). Areas in green tend to underlie social learning, including *orbital* (OFC) and *medial frontal regions* (MPFC), *amygdala*, *insula*, and *striatum*, *including nucleus accumbens* (NAC), *caudate*, *putamen*, and *globus pallidum* (GP). Areas in brown tend to underlie moral judgements, *MPFC*, *posterior cingulate*, *bilateral angular gyri* (AG), *middle frontal gyrus*, *bilateral parietal lobes*, and *insula*. Areas in pink and grey tend to underlie empathy, *amygdala*, *MPFC*, *STS*, *precuneus*, *anterior cingulate cortex* (ACC), and *insula*. This map shows their relative saggital and axial position in the brain, and structures are not drawn to scale and ignore their coronal positions.

stimulus. Familiar names are more popular, and people prefer the familiar letters associated with their own initials, suggesting that familiarity may plausibly explain this name-letter and other implicit egotism effects (Pelham, Mirenberg, & Jones, 2002). Familiar social targets generate positive affect that links in part to the self. Finally, friends have an easier time inferring each other's thoughts and feelings than strangers do (Stinson & Ickes, 1992), illustrating how familiarity moderates mental state inferences. Later sections return to the roles of MPFC and familiarity in dehumanised perception.

Self. Social neuroscience also links the MPFC, particularly more ventral (lower) areas of MPFC, to thinking about the self. MPFC activates in paradigms where participants reflect on themselves, access self-knowledge, or compare the self to another (Johnson, Baxter, Wilder, Pipe, Heiserman, & Prigatano, 2002; Kelley, Macrae, Wyland, Caglar, Inati, & Heatherton, 2002; Lieberman, Jarcho, & Satpute, 2004). Self-regulation of affect also activates the MPFC (Ochsner, Knierim, Ludlow, Hanelin, Ramachandran, Glover, et al., 2004). Self-reflection allows people to infer the minds of others (Mitchell, Banaji, & Macrae, 2005), and it does activate the MPFC (along with the posterior cingulate and precuneus).

Research within social psychology on similarity to the self, such as self–other biases, self-referential effect, and self-esteem, has illustrated that the self serves as a positive attitude-object. In essence, things similar to the self become associated with positive affect, and denial of this association leads to less positive affect to self. The endowment effect suggests that people perceive objects belonging to the self as more valuable (Thaler, 1980).

Overall, MPFC activity is implicated in a number of networks involved in fundamental social cognitive processes, including inferring others' minds, preference for familiar others, and links to the most familiar and preferred social target, the self.

Neural systems of (social) reward

Social perception overlaps social reward, as supported by findings that social interaction tends to be intrinsically rewarding. Most people perceive themselves and others positively by default (Fiske, 2004, pp. 23–24; Kwan, John, Kenny, Bond, & Robins, 2004; Sears, 1983; Taylor & Brown, 1988; Taylor & Gollwitzer, 1995). The person-positivity bias is one of the most robust effects in interpersonal ratings; for example, people most often use the top half of rating scales to evaluate others. Of course, not everyone is perceived positively, but positive expectations are clearly the default for social targets.

In converging evidence, neural regions engaged in tracking reward and punishment value especially activate to social stimuli as well (Harris, 2007); these regions include orbital and medial frontal regions, amygdala, insula, and striatum, including nucleus accumbens, caudate, putamen, and globus pallidum (Delgado, 2007, see Figure 1). The MPFC also has a broader role here, primarily as an affective responsive area of the brain especially tuned to social rewards (Harris, McClure, Van den Bos, Cohen, & Fiske, 2007; Van den Bos, McClure, Harris, Fiske, & Cohen, 2007). Although the MPFC may be especially tuned for social perception, its broader function includes reward processing. Participants given an immediate reward for performance exhibit MPFC activity, but especially when a person rather than a computer administers the reward (McClure,

Laibson, Loewenstein, & Cohen, 2004). Thus social neuroscience and neuroeconomic evidence indirectly supports the person-positivity effect, which will prove relevant when comparing person perception to dehumanised perception.

Neural systems of moral judgement

Another aspect of dehumanised perception is moral, as we will argue that dehumanised others allegedly lie beyond normal human moral boundaries. Moral violators elicit disgust (Haidt, Rozin, McCauley, & Imada, 1997) and are often viewed as sub-human (Opotow, 1990). A moral module is viewed as necessary for human behaviour (Hauser, 2006). Moral judgements activate either a more cognitive or a more affective system (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). The more affective system includes MPFC, posterior cingulate, and bilateral angular gyri. The more cognitive system includes the middle frontal gyrus and bilateral parietal lobes. However, because of the role of disgust in marking moral violation, we also include the insula and strial regions in this network (see Figure 1).

Neural systems of empathy

Finally, we will invoke the role of empathy in contrasting person perception and dehumanised perception. Empathy appeared early in helping research (for review, see Batson, 1998), facilitating altruism (Batson, 1991), a process requiring thought about another's internal state. Subsequent work on perspective taking further suggests that empathy reduces prejudice (Galinsky & Moskowitz, 2000). Like social psychology, some neuroscience theory also considers empathy an outcome of emotional contagion (e.g., Levenson, 1996).

Other theories of empathy describe empathy as embodying another's experience (Singer & Fehr, 2005). Embodied social cognition (Niedenthal, Krauth-Gruber, & Ric, 2006) explains emotional reactions by arguing that emotional experiences mentally re-create patterns of neural activity similar to the patterns that occurred during the original experience of the event. Empathy can occur across separate sensory modalities such as pain and touch. Therefore, perceivers may re-create the experienced neural affective pattern of the target. This neural system includes the anterior cingulate cortex (ACC) and the anterior insula during empathy for pain (Singer & Fehr, 2005). The social neuroscience literature shows greater amygdala activity (also implicated in fear conditioning) when participants merely read third-person fearful versus neutral stories about other people (Ruby & Decety, 2004). The MPFC along with STS and precuneus also reliably activates in tasks requiring empathy (Decety & Jackson, 2004, 2006, see

Figure 1). Again, empathy would appear in humanised perception more than dehumanised perception.

One common theme: Medial prefrontal cortex in social cognition

Which neural structures activate most commonly across such complex processes? The relevant neural system must integrate both affective and cognitive information. One candidate focal structure is the MPFC. As previous sections indicate, this neural region activates in detecting valence and value, as well as in mentalising, with reciprocal connection to sub-cortical neural regions implicated in the immediate processing of information for moral decisions and empathy. Therefore our framework takes MPFC activation as an index of humanised perception in the context of our dehumanised perception tasks. This is not the definitive function of the MPFC, but given the literature, activation of the MPFC is one reliable index of humanised perception because of its ubiquity in social cognition, as just described. The next section describes dehumanised perception and its primary mechanism, mentalising, and begins by asking an old but useful philosophical question.

DEHUMANISATION: OUTCASTS, ANIMALS, AND OBJECTS HAVE (HARDLY ANY) MINDS

Before turning to dehumanised perception, let us define *humanised perception*: perceiving a target as possessing an internal life. An internal life is defined as active mental states; that is, thoughts and feelings. These thoughts and feelings presumably cause the target's behaviour, so they have a function and are not epiphenomenal. This definition holds when people perceive most other people. When humanised perception does not occur, and a person is viewed as a non-human target, we consider this phenomenon *dehumanised perception*.

Considering what is human fascinates philosophers, who use thought exercises to argue that the adoption of an "intentional stance" allows a folk psychology of the mind (Dennett, 1987). The *intentional stance* captures the idea that perceivers assume others are agents with plans, goals, and predispositions. When people think about others, they often think about their minds. Inferring what is in someone else's mind—i.e., *mentalising*—helps create one's mental representation of another person as a truly social target. After framing the philosophy of mind approach, other sub-sections will describe the intentional stance as applied or denied to outgroups, animals, and objects.

Philosophy of minds: Intentional and personal (moral) agents

First, entertain the following thought experiment demonstrating the intentional stance. What if, unbeknown to all, someone replaced your immediate family with robots? Consider these no ordinary robots, but robots that resemble and respond just as your family members. How would you distinguish your family from these robots? When lay people consider a similar question, the folk psychological answer is often that these robots will not have the minds or "inner lives" of a person; that is, the robots would not have the thoughts, feelings, and experiences that make up the rich mental lives of people (Appiah, 2003). They may have computer chips and circuits, but they lack the phenomenological experience of what it feels to be human. Similarly, when asked what makes humans unique, people cite intelligence, language, and complex emotions (Leyens et al., 2003), all of which require agency and mind.

To determine that others are human may not be as obvious as simply knowing in theory that they have a mind. The only mind that we know exists is our own because we have phenomenological experience only of our own mental lives. Therefore we infer whether other agents have minds like ours by assuming that a mind underlies their similar behaviour. Perhaps, therefore, strange (dissimilar or unfamiliar) behaviour suggests that the agent is not like us because our minds do not underlie such behaviour. The agent must not be quite human like I am quite human. Thus philosophy of mind plays a vital role in thinking about dehumanising a dissimilar other.

In summary, dehumanisation may result from inference driven by the other's alleged behaviour. This fits social psychology theory on dehumanisation. If behaviour is so heinous that we could not imagine ourselves performing it, then the actors may be dehumanised. Therefore moral judgement elevates or demotes people's human status. The philosophical theories certainly suggest as much—the intentional stance includes the higher-order stance called the *personal stance,* which identifies a person as a moral agent comparable to self (Dennett, 1987).

Assessing intention is not reserved simply for people, suggesting that this additional, personal-moral variable helps differentiate people from objects and animals. Consider another thought experiment: Is it permissible to dismember a computer with which one plays chess? Now consider dismembering a friend with whom one plays chess. Dennett's examples suggest that intentionality is not sufficient to differentiate people and objects because some behaviours are perceived as "right" and others as "wrong" towards the two kinds of intentional agents. The personal stance captures a separate inference reserved for people, a kind of moral filter on behaviour towards intentional agents. Dehumanisation may result from a failure of a person to activate the personal stance, a failure to indicate to the perceiver that the social target is not only an intentional but also a moral agent similar to self.

Dehumanising prejudice may deny intentional and personal agency

Social psychologists have long pondered the mechanisms of dehumanisation. Indeed social psychology itself expanded with the exodus of European Jewish researchers in the wake of the Nazi Holocaust, during which dehumanisation was the daily norm. Early work on prejudice provides a useful point of departure.

Social psychology has described dehumanisation as the worst kind of prejudice (Allport, 1954). Since Allport, the field has theorised about which groups get dehumanised: Outgroups perceived to act outside the prescribed boundaries of moral rules and values (Opotow, 1990; Staub, 1989), as well as social groups considered beyond societal norms (Bar-Tal, 1990), both lead to moral exclusion.

Dehumanisation theory also acknowledges differing forms, distinguishing *uniquely* human characteristics (e.g., language) from *typically* human characteristics (e.g., agency). Lacking either uniquely or typically human characteristics dehumanises one to the equivalent of, respectively, animals or objects (Haslam, 2006; Haslam, Bain, Loughnan, & Kashima, 2008; Loughnan & Haslam, 2007). A model of dehumanisation separates exclusion from unique humanity and typical human nature (Haslam, 2006). The model revisits the idea of essences, once proposed as an account of how people know the difference between computers and people (the answer: computers lacked the "human essence"; Turkle, 2005).

Denying *uniquely human* characteristics disallows aspects that humans share, compared with other species, such as courtesy, culture, intelligence, language, and so on, that elevate humans above animals. Hence, denying uniquely human characteristics reduces people to functional equivalents of animals. Denying *typically human-nature* characteristics disallows aspects central to people's folk definition of what it fundamentally means to be a good example of a human; for example, having complex emotions. Refuting these characteristics makes another appear more like a machine or automaton than a person. These unique and typical characteristics represent two distinct ways of dehumanising people (Loughnan & Haslam, 2007).

The field has researched one related mechanism to date: Outgroups allegedly do not feel complex emotions to the same extent as the ingroup (Leyens et al., 2001, 2003). *Infrahumanisation* theory of dehumanisation thus directly involves affect, drawing on the distinction between primary, basic emotions (sadness) and secondary, social, complex emotions (remorse). Infrahumanisation states that we attribute complex secondary emotions less to outgroups than ingroups (Leyens et al., 2001, 2003).

In infrahumanisation paradigms, participants attribute positive and negative primary and secondary emotions to ingroup and outgroup members (Leyens et al., 2001, 2003). The distinction between primary and secondary emotions is inherent in romance languages that distinguish between *émotion* and *sentiment*. The *émotions* are the affective reactions that people and animals both can feel, such as sadness or joy, whereas *sentiments* are complex affective reactions that only people can feel, such as regret or pride. *Émotions* in this sense are basic or primary emotions, whereas *sentiments* are higher-order or secondary emotions, most often social. Participants attribute equal numbers of positive and negative primary *émotions* to both ingroups and outgroups, but fewer positive and negative secondary *sentiments* to outgroups than ingroups. This phenomenon demonstrably occurs both explicitly and implicitly (Demoulin et al., 2005), suggesting that emotional infrahumanisation may result from immediate categorisation (Rodriguez-Torres, Leyens, Rodríguez Pérez, Betancor Rodriguez, Quiles del Castillo, Demoulin, et al., 2005).

People attribute greater humanness to themselves than to others (Haslam, Bain, Douge, Lee, & Bastian, 2005), and a possible mechanism is similarity (Vaes, Paladino, Castelli, Leyens, & Giovanazzi, 2003), but not familiarity (Cortes, Demoulin, Rodriquez, Rodriguez, & Leyens, 2005). Infrahumanisation reduced intergroup helping after a natural disaster (Cuddy, Rock, & Norton, 2007b) and decreased forgiveness following conflict (Tam et al., 2007). Further, when made aware of mass killings, ingroup members infrahumanise outgroup victims (Castano & Giner-Sorolla, 2006); if the ingroup is victim and the perpetrator ambiguous, increased empathy with ingroup victims increases outgroup infrahumanisation (Rodriguez, Coello, Betancor, Rodriguez, & Delgado, 2006), all this suggesting that negative outcomes and uncertainty may moderate infrahumanisation. Children as young as 6 show the infrahumanisation bias (Martin, Bennett, & Murray, 2008).

Relatedly, people associate outgroups and animals, specifically Blacks and apes (Goff, Eberhardt, Williams, & Jackson, 2008; see Figure 2). Across several studies these non-conscious associations appear in early perception and attention. Further, this association increases implicit endorsement of violence against Blacks, activates spontaneously from newspaper articles covering death-eligible Black defendants, and more strongly activates for Black targets ultimately sentenced to death.

Dehumanisation as inter-group cognitive bias and emotional prejudice

Dehumanisation, like other forms of inter-group bias, has both cognitive and affective components. Because group categorisation underlies all forms of dehumanisation, and social norms prescribe which categories get

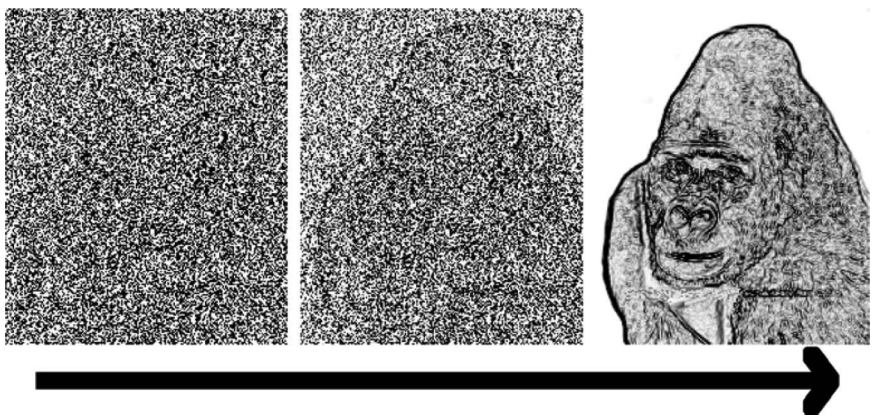

Figure 2. Illustration of the kinds of stimuli used in studies showing faster latent responses in White subjects to associations of Black people and apes. Participants were asked in one paradigm to watch a movie of a picture of an animal slowly clear from noise to image. White participants identified the image much sooner after being primed with Black (first image) than White (second image) faces.

dehumanised, dehumanisation results from an intergroup cognitive phenomenon. As an extreme emotional prejudice, dehumanisation may be motivated by the emotions involved.

Of most relevance here, social targets who elicit disgust are often linked to moral violations and suffer aggression from the perceiver (Haidt et al., 1997). Additionally, as we will see, they activate the disgust-related insula more and the mentalising MPFC much less than other social targets (Harris & Fiske, 2006); in questionnaire studies they are spontaneously mentalised less than other social targets (Harris, 2007). Thus, disgusting social targets both suffer the cognitive bias of dehumanised perception and fail to elicit complex, even slightly positive emotions. Because the cognitive process of mentalising helps detect dehumanised perception of disgusting outgroups, our work on dehumanisation involves both cognitive and affective processes.

Cognition and affect are not mutually exclusive processes, and infrahumanisation like dehumanised perception has elements of both components (an attributional process, even for emotions, must also be cognitive). We differentiate dehumanisation from infrahumanisation processes based on the dominant process involved—attribution of emotions versus attribution of mind—but because both phenomena involve affective as well as cognitive processes, we do not differentiate them based on relative emphases towards cognition or affect. Also, we consider dehumanised perception a denial of typical humanity, not unique humanity, since the phenomenon is concerned with mental state inference, a typically human quality; infrahumanisation theory seems to accommodate both denials.

In contrast: Object perception, animal perception, and anthropomorphism

So far we have seen that outcast humans may allegedly have hardly any mind; now we turn to the opposite case of objects and animals that may apparently have too much mind. People can infer mental states for targets that lack actual minds. People anthropomorphise machines and animals, suspending the belief that they do not have human consciousness (although the debate still contests animals' consciousness, see Gosling, Kwan, & John, 2003).

Perceiving and attributing human-like qualities to non-human targets (i.e., anthropomorphism), is not necessarily the opposite of dehumanised perception. Clearly, specific non-human targets are perceived like people. The most common cases are domestic animals, or pets; many pet owners attribute personalities to their pets (Gosling et al., 2003). This is not surprising, considering the great social services these animals provide, particularly companionship; studies show that after a traumatic life event, depressed elderly people with pets fare better than depressed elderly people without (NIH, 1987). Regardless of the extent to which animals have consciousness, along the dimension of humanised perception these animals may be perceived as more human than actual human beings who happen to be at the bottom of the social hierarchy. For this reason we consider dehumanised perception a cognitive error, and suggest that perceiving other people is also a mental, not strictly visual phenomenon. We will elaborate on object perception and anthropomorphism below.

As an everyday example, children commonly create imaginary people (imaginary friends) and effortlessly interact with them as if they exist. Even adults may hold or suspend their beliefs about whether a target possesses a mind, allowing either dehumanisation or anthropomorphism (for relevant articles, see Kwan & Fiske, 2008). Comparing object perception to person perception, and exploring the cognitive and neural mechanisms that underlie anthropomorphism provides subsequent hypotheses that address how people decide which targets have mental lives and which do not. The next two sections elaborate some psychological and neural mechanisms.

Object versus person perception. Studies that contrast person and object perception (Harris et al., 2007; Mitchell, Heatherton, & Macrae, 2004) clearly implicate the MPFC in social perception. In one study participants saw pictures of positive and negative people and objects while fMRI recorded their neural responses (Harris et al., 2007). Pictures of people came from our Stereotype Content Model (SCM) social-groups database (Harris , 2007, Fig. 3), and pictures of objects came from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 1996). "Positive"

Figure 3. Stereotype content model (see online for colour version). The interaction of perceived trait warmth and competence predicts the type of emotion that different social groups elicit. Groups perceived as low on both dimensions are dehumanised groups who elicit *disgust* (blue) and include drug addicts and homeless people. Groups high on both dimensions elicit *pride* (red) and include college students and American heroes. Groups high on warmth, but low on competence elicit *pity* (yellow) and include elderly and disabled people. Groups low on warmth and high on competence elicit *envy* (green) and include business and rich people. All social groups come from US samples.

pictures of people (defined as pictures that activated MPFC) had pretested as eliciting pity, envy, or pride (all emotions that entail at least some positive valence); examples included, respectively, older and disabled people, rich and business people, and ingroup heroes. Wholly negative pictures were dehumanised people who elicit disgust (homeless people, drug addicts). Positive pictures of objects (e.g., flowers) pretested high (6 or more) on a 1–7 Likert scale of valence, while negative pictures of objects (e.g., a collapsed building) pretested low on the valence measure (2 or less). The objects data came directly from published ratings of the IAPS (see Lang et al., 1996). In the scanner, participants simply viewed each picture for 500 milliseconds while indicating either that the picture was positive or negative, or depicted a person or object.

Although no area of the MPFC activated more to people than objects overall, the neuroimaging results reveal an area of MPFC recruited

exclusively for *valence differences among people*, but not for valence differences among objects. Also, a different area of MPFC activated more to positive than to negative stimuli in general. This combination of results suggests that positive and negative valence differentiates in a unique region for people, not objects, in the MPFC.

Valence extends to social and nonsocial feedback in critical learning experiences that guide future behaviour. Is receiving reward or punishment from a person the same experience as receiving reward or punishment from a computer program? Again contrasting person and object perception, we addressed this question with the following experiment. Participants in another imaging study performed a time-estimation task in which they had to guess how long different time-periods took to elapse (Van den Bos, McClure, Harris, Fiske, & Cohen, 2007). On each trial, participants first saw how much time they had to estimate (e.g., 9.5. seconds) before they then pressed a button to start the estimation, and another to signal when the specified time had elapsed. These participants received either a reward (juice) or a punishment (quinine, a bitter liquid), allegedly from either an experimenter or a computer program. In fact, a computer program randomly determined both types of feedback.

These data also show a region of MPFC (similar to the one identified in the above experiment) that differentiates reward and punishment for social stimuli (the experimenter), but not non-social targets (the computer program). Consistent with the Harris et al. (2007) data, but using a completely different paradigm, this study demonstrates neural differences between person and object perception; areas of MPFC integrate information and represent value that differentiates good and bad feedback from people, but not objects. These data hint at the possibility that receiving reward and punishment (learning cues for future behaviour) may be a separable psychological process for people and objects; a part of the brain distinguishes the source as person or object.

Similar evidence is provided by another imaging study where participants played an Ultimatum Game with either another person or a computer (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). Unfair offers from people activated regions of the insula, dorsal lateral prefrontal cortex, and MPFC extending into the ACC. These areas responded more when receiving unfair offers from people than when unfair offers were received from the computer, suggesting that simply perceiving intention (communicated by fair and unfair offers) is not sufficient for humanised perception.

This latter study illustrates another psychological process separable for people and objects. However, this study suggests separable intentions from people and objects, a trait unique to targets with mental lives. It also implies that people can infer intentions (or something like them) from objects, hinting that separable neural processes may underlie intentional inference

from human and non-human agents. Possibly, people must suspend their belief that the object does not have a mental life, even if ever so temporarily, to imbue that object with a quality of mental life: intention. Are there cases beyond the context of economic games where people imbue non-human targets with mental life?

Anthropomorphism. Given that people infer fairness from objects, a trait implying mental life, but the neural processes underlying this inference are separable, it suggests that separable neural structures may underlie other mental inference processes for people and objects. We tested this hypothesis with a series of neuroimaging studies that extended previous work in social psychology on another form of mental life inference: dispositional attribution.

Social psychology has established that given the right combination of information about a social target's past behaviour and other people's behaviour, a perceiver will make a dispositional attribution to that social target. For instance, being told Jane shops in a health-food store, and then given additional low consensus information (Hardly anyone shops in that health-food store), low distinctiveness information (Jane shops in every other kind of health-food store), and high consistency information (In the past, Jane has always shopped in that health-food store), perceivers are likely to infer that there is something about Jane that caused her behaviour (Jane likes to shop in health-food stores). This covariance model of dispositional attribution (Kelley, 1972) has been well established in the field and predicts when people are likely to attribute the cause of behaviour to a social target, as apart from circumstances, particular entities, or ambiguous causes (McArthur, 1972).

Participants read high and low consensus, consistency, and distinctiveness information in a direct replication of the McArthur (1972) attribution paradigm, while we collected neural data using fMRI (Harris et al., 2005). Their task was to indicate who caused the behaviour, whether the human subject of the sentence, the entity, circumstance, or some combination. We replicated previous behavioural results of more dispositional attributions, given low consensus, low distinctiveness, and high consistency information. Activation in the superior temporal sulcus (STS; a region that tracks motion that appears to be made by a biological agent, implying intentions) exactly mimicked this pattern, namely activating in only one out of eight combinations. In addition, participants also showed increased MPFC activity to low distinctiveness and high consistency information, paired with either low consensus or high consensus information (in line with the literature using questionnaire data, showing that people neglect consensus information, relative to consistency and distinctiveness; Fiske & Taylor, 2008). Thus the MPFC activated in attributions to a specific person or to people in general.

To compare person and object perception, we then replicated this paradigm using objects as the targets performing behaviour. Because we did not want to demand anthropomorphism, we chose action sentences instead of emotion, accomplishment, and opinion sentences. As just described, participants first saw the action (The pen fell off the table) before information combinations that suggested a causal property residing in the pen (Hardly any other pen falls off the table; the pen falls off almost every other table; in the past the pen almost always fell off the table). Neural activity for these "dispositional" attributions for objects overlapped with the STS activity previously engaged for people, but also activated bilateral amygdala instead of MPFC. This suggests separable neural systems encoding dispositional inferences for objects, a phenomenon close to anthropomorphism, but distinct from (de)humanised perception.

Summary

After the previous section providing background on social neural networks, this section has focused on dehumanisation in various forms, starting with people normally understanding other people as having intents and agency; dehumanisation denies them an agentic mind. Dehumanised groups are extreme outgroups viewed as disgusting. In contrast, overly humanised objects and animals may acquire agency, but the neural mechanisms evidently differ.

PUTTING TOGETHER DEHUMANISATION AND SOCIAL NEUROSCIENCE

Previous sections have introduced first the social neuroscience of person perception and social cognition, social rewards, moral judgement, and empathy—and second, dehumanisation as a general form of prejudice and intergroup bias, noting its relevance to anthropomorphism and object perception. This section pulls together affect and cognition in social cognition, citing social psychology and social neuroscience evidence.

Affect in person perception: The stereotype content model

The social psychological literature provides a number of models that describe the affective process during person perception. This section begins with a review of a model predicting intergroup emotion, the stereotype content model (SCM; Fiske, Cuddy, & Glick, 2007; Fiske, Cuddy, Glick, & Xu, 2002), which we use as a framework to test dehumanisation linked to reported disgust. Then the section will discuss broader social-group

approach–avoidance reactions, linked to amygdala activation, an area implicated in emotional vigilance.

The inter-group emotions literature suggests that affective reactions to social targets emerge from their perceived social categories and associated stereotypes (Alexander, Brewer, & Hermann, 1999; Fiske et al., 2002; Mackie & Smith, 2002; Neuberg & Cottrell, 2002; Stephan & Stephan, 2000). The stereotype content model (see Figure 3) begins with two basic dimensions of person perception: warmth and competence (Fiske et al., 2002, 2007; Peeters, 1983; Rosenberg, Nelson, & Vivekananthan, 1968; Tausch, Kenworthy, & Hewstone, 2007; Wojciszke, Bazinska, & Jaworski, 1998). People rapidly assess a social target's intention (good or ill) towards them. The warmth trait dimension captures the social target's assessed benevolent or malevolent intention, while the competence dimension captures the target's perceived ability to enact those intentions (Cuddy, Fiske, & Glick, 2007a; Fiske et al., 2002; Harris & Fiske, 2006). The resulting two-dimensional space predicts characteristic affective responses to distinct social categories, based on the interaction of perceived warmth and competence. Location of different social categories on the trait dimensions comes from their perceived status and competition, social structure variables associated with the inferred social category. Then, specific emotions are elicited by the social category (see also Eagly & Mladinic, 1989; Hamilton, 1981; Jackson et al., 1996; Stangor, Sullivan, & Ford, 1991; Zanna & Rempel, 1988).

The SCM rests on general social psychological principles and is well supported (Fiske et al., 2007). SCM data have come from US representative surveys (Cuddy et al., 2007a), US college samples (Fiske et al., 2002), as well as Asian and European samples (Cuddy et al., 2009). In these studies an initial sample within the population first lists the most relevant social categories in that context. The social categories with some consensus (above 15%) are then rated by a separate sample on the trait dimensions warmth and competence, and in some studies on the four predicted emotions (pride, envy, pity, disgust), four predicted behavioural intentions (active or passive harm or help), as well as the social structural predictors of warmth and competence (respectively, competition and status). A cluster analysis on the categories' scores on the warmth and competence trait ratings then generally results in the four quadrants towards the corners of the 2×2 space.

Social categories perceived as high on both warmth and competence elicit the ingroup, complex, positive, social emotion *pride*; in US samples these cultural prototype groups include middle-class people, American heroes (firefighters, police officers, astronauts), and college students (in college samples). These social targets are admired and respected. The high-high groups receive both active and passive help.

Social categories perceived as low in warmth but high on competence elicit the complex, ambivalent, social emotion *envy*: business people and rich

people in all samples. These groups are perceived as not nice, but well respected; high-status outgroups in all samples tend to fall into this space. These groups receive the volatile combination of passive help (go-along-to-get-along) but also active attack, under social breakdown.

Social categories perceived as high on warmth but low on competence elicit the complex, ambivalent, social emotion *pity*: older people in all samples and disabled people in US samples. These likable social targets are perceived as inept and needing active help, being cared for, but often neglected (passive harm).

Finally, social categories perceived as low on both warmth and competence elicit the negative, basic emotion *disgust*. These extreme outgroups include homeless people and drug addicts in US samples (and poor people in all samples). These groups tend to be dehumanised targets (Harris & Fiske, 2006), a point that forms the basis of this chapter. As discussed previously, social targets who elicit disgust are often linked to moral violations and suffer aggression from the perceiver (Haidt et al., 1997). They receive both active and passive harm (neglect). We will come back to these groups in particular.

Cognition in person perception

Background on mentalising. Mentalising is a cognitive process that involves inferring a target's intention (Frith & Frith, 2001). Attributing intention is not simply reserved for people who presumably do have minds; perceivers also attribute intentions to objects that objectively do not have cognitions and emotions. Participants likewise infer the intentions of dots in biological motion (Heberlein, Adolphs, Tranel, & Damasio, 2004) and of shapes in non-random motion (Heider & Simmel, 1944). Inferring intention in these instances may rely on the assumption that the targets have goals that reside in their minds. Although obviously people realise that these targets are not human and do not have internal lives of any kind, nevertheless people continue to talk to cars and computers (though always knowing that no answer is forthcoming). Therefore, as discussed above, mentalising is a higher-order cognitive process that people quickly recruit even to non-human targets, and it *seems* to require the belief, if only temporarily, that the target has "an internal life," or conscious cognitive and emotional experiences.

However, mentalising with human targets has certain additional trademarks. Consider mentalising as a cognitive process with affective correlates; that is, it may often involve certain kinds of emotion. For instance, we can easily report how trustworthy a face seems, in fact, reliably after 100ms of face presentation (Willis & Todorov, 2006). A judgement of trustworthiness assesses the social target's good or ill intentions, which represent an

immediate primitive affective (or at least evaluative) appraisal. This appraisal relies on the same neural architecture, the amygdala, as other immediate approach–avoid affective judgements such as fear (Engell, Haxby, & Todorov, 2007; Winston, Strange, O'Doherty, & Dolan, 2002). Yet this complex evaluation of intention occurs so quickly that it is difficult to make a case for the role of higher-order cognitive processes in this judgement. In fact, the amygdala often responds before information has even reached neo-cortical structures (LeDoux, 1998). Therefore, perhaps thinking about the mind of a person relies on both higher-order cognitive processes and sub-cortical affective structures that appraise intention and process basic emotion. Mentalising thus may be a cognitive response with affective correlates.

Mentalising as a cognitive process. Social psychology has addressed two major aspects of mentalising: thinking about a target's mind by perspective taking and thinking about a target's mind by making dispositional inferences. In both cases, the *target*'s external features (e.g., facial appearance; see Todorov, Mandisodza, Goren, & Hall, 2005; Zebrowitz, 1999) perhaps most influence initial thoughts about that other's mind. These external features convey clues about the social target's internal state, but also activate stereotypes about the perceived social category (Fiske, 1998). Stereotypes guide inferences about the target's mind and assist in predicting the target's behaviour. Sometimes, however, perspective taking and considering the target's mind can short-circuit these default stereotyping processes.

First, perspective taking de-biases social thought on both the conscious and unconscious levels (Galinsky & Moskowitz, 2000). For example, perspective taking deactivates stereotypes by increasing cognitive overlap between the self and the other, allowing the other to be perceived as similar (Galinsky, Ku, & Wang, 2005). These effects occur both in real and minimal groups (Galinsky & Moskowitz, 2000). Prior to this work, social psychology had a long history of implicating perspective taking in elevating moral reasoning (Kohlberg & Hersh, 1977), encouraging altruism (Batson, 1991, 1998), and lowering aggression (Richardson, Hammock, Smith, & Gardner, 1994).

Second, people may consider another's mind and use those imagined mental contents as a marker of a stable attribute about that person. Because of their ability to deduce personality, people have famously been considered naïve scientists (Heider, 1958). Inferring a social target's disposition is a type of mentalising that makes a stable attribution to the target traits, goals, preferences, and the like, often in the service of predicting behaviour. Activating a disposition is similar to activating a stereotype because it includes a mental script of that social target's expected behavioural repertoire, as well as the perceiver's potential behaviour towards the social

target. A friendly person who is reserved versus a friendly person who is outgoing each have different behavioural tendencies towards greeting with either a shy glance or a fierce hug. People may make "behavioural" predictions even about non-human agents, and anthropomorphising animals and objects may reflect the disposition-to-behaviour feature of mentalising processes (Gosling et al., 2003).

Attribution theory in social psychology has focused almost exclusively on the process of making dispositional attributions to others (Jones, 1979), including the dimensions that predict these attributions (Kelley, 1972; McArthur, 1972). Dispositional attributions are so crucial to social perception that people automatically infer dispositions from even thin slices of behaviour (Ambady & Rosenthal, 1993; Dunning, Meyerowitz, & Holzberg, 1989). In addition, these first impressions often influence global judgements of the individual (see Asch, 1946). This judgement anticipates affectively based good or bad evaluations of the social target, even when only the face, not detailed information itself is available (Todorov & Uleman, 2002). As noted, dispositional attribution activates the neural structures associated with mentalising. This suggests that implicit information about the social target, possibly correlated with affect, is relevant to thinking about their minds. We next explore social psychological data that provide converging evidence for the social neuroscience data on dehumanised perception.

Social psychological data on dehumanised perception

The neuroscience studies mentioned above generated a cognitive hypothesis testable with experimental social psychological paradigms: Perceivers less often infer the mental state of dehumanised people. This hypothesis is tested with a social cognitive paradigm that examines whether perceivers report mentalising dehumanised targets, and how they rate these targets on a number of dimensions drawn from the MPFC literature.

Cognitive responses to dehumanised people. Because social targets who elicit disgust may not be perceived as human to the same extent as other social targets, the number of thoughts about the mental contents of dehumanised targets may differ from other social targets.

As one indicator, people describe other people using verbs that differentially abstract behaviour. *Mental state* verbs describe behaviour by implying the mental content of the agent, whereas *descriptive action* verbs and *interpretive action* verbs do not require inferring mental content (Semin & Fiedler, 1988). Therefore, the amount of mental state verbs participants generate in response to different social targets may help reveal the amount of mentalising to dehumanised targets. These targets should

generate fewer mental state verbs when participants describe a day in that social target's life.

To demonstrate this effect, Princeton University undergraduates saw one of eight pictured social targets and were asked to image what their day is like. Pictures came from our larger picture database pretested on warmth and competence. Two pictures represented each SCM quadrant as follows: female college student and male American firefighter (pride), business woman and rich man (envy), elderly man and disabled woman (pity), and female homeless person and male drug addict (disgust). No ratings differentiated the social targets within each SCM quadrant, so the average ratings for each pair generated the measure for that part of the space.

Participants first described a day in the life of one pictured social target, using up to 15 lines. We used a coding scheme (Semin & Fiedler, 1988) to test the hypothesis. Specifically: while adjectives represent the most abstract level (e.g., thirsty), verbs are more concrete but, as noted earlier, can differentiate levels of abstractness. To review, the most abstract are *mental state verbs* that describe actions in terms of the target's internal state (e.g., quench). These are followed by *interpretive action verbs* that interpret the target's action (e.g., guzzle), and finally by *descriptive action verbs* (e.g., drink) that simply describe the action. Note what the examples in parentheses suggest about the cognition behind the use of each type of descriptive term: Adjectives describe the person, mental state verbs describe the mind of the person, interpretive action verbs interpret the behaviour of the person, while descriptive action verbs describe the behaviour in terms of the object being acted upon. This suggests that people will use fewer mental state verbs to describe a day in the life of dehumanised targets, if perceivers fail to think about the contents of the target's mind.

A pair of independent raters jointly classified verbs into the mental state, interpretive action, and descriptive action categories. The number of extracted verbs in each category and adjectives in the descriptions were averaged across the two raters, creating a single, reliable score per participant on each type of verb and adjective. No significant differences distinguished the total number of words used to describe the different social targets. Also, no significant differences appeared in adjectives, interpretive action verbs, or descriptive actions verbs used in the descriptions of any type of social target. However, as hypothesised, participants did use significantly fewer mental state verbs to describe a day in the life of dehumanised targets, compared with all the other social targets, suggesting that participants were spontaneously inferring the contents of the dehumanised targets' minds *less* than other social targets (Harris, 2007). These results suggest that perceivers do not infer the contents of dehumanised targets' minds or take their perspective to the same extent as they do for other social targets.

What processes could be associated with describing all the social targets? Participants, when instructed to imagine a day in the life of a social target, could have engaged in a variety of processes. For example, participants may use thoughts about their own social life in an attempt to imagine what another person's might be. Participants may also have thought about the social life of people familiar to them, people who might be similar to the social target. Participants may have attempted to empathise with the social targets, perhaps remembering experiences in their own life that correlate with the possible daily experiences of the social target. These plausible explanations may describe the strategy employed by participants when mentalising.

We next asked participants to rate the pictured social target on a number of dimensions, including their own subjective experience regarding the social target. These dimensions included ones chosen to assess potential mentalising: warmth, competence, similarity, familiarity, perceiver's ability to mentalise the target, perceiver's ability to infer target dispositions, and perceiver's empathy for the target. Other rating dimensions derived from the social psychological literature on dehumanisation: responsibility of target for own situation, control of target over own situation, being articulate, being intelligent, having complex emotionality, target being self-aware, ups and downs in target's life, and target's typical humanity.

A three-factor solution described the rating dimensions: a *competence-autonomy* dimension consisted of competent, articulate, intelligent, similar, responsible for situation, and control over situation; a *warmth-mentalising* dimension consisted of warm, mentalise, infer dispositions, and familiar; and an *emotional-connection* dimension consisted of typically human, empathise, complex emotions, and self-awareness. Significantly *less* competence-control and warmth-mentalising were reported for the dehumanised targets, consistent with the SCM. Dehumanised targets were rated significantly lower on rating dimensions derived from tasks that activate the MPFC. Participants think less about the contents of dehumanised targets' minds. They also rated these targets lower on personality and subjective-experience dimensions. No differences emerged on the emotional-connection dimension, except on the item typically human, on which the dehumanised targets were rated significantly lower than the other three (Harris, 2007).

Affective responses to dehumanised people. Clearly, people have affective reactions when perceiving others. According to our research these reactions are more than simple positive or negative valence, and they depend on the perceived social group of the target. Consistent with complex affect responses to people is the notion that a person can evoke more than one affective response. Indeed, if asked, participants will report a variety of

emotions in response to any one person. We documented this complexity by asking Princeton University students to rate pictured people on the four emotions predicted by the SCM.

All pictures came from our social groups database. All people represented one of eight pretested SCM groups. Two social categories depicted each quadrant as follows:

(a) *High warmth, high competence*—American heroes (e.g., firefighters, police officers, astronauts), college students;
(b) *Low warmth, high competence*—business people, rich people;
(c) *High warmth, low competence*—disabled people, elderly people;
(d) *Low warmth, low competence (dehumanised people)*—homeless people, drug addicts.

Demographic variables (race and gender) varied evenly across all social targets, and age varied evenly across all except the elderly people and college students.

Nearly 300 standardised pictures appeared in roughly numerically equivalent subsets, each presented as a separate sample, so participants rated only one individual subset in an online study. Each social target was rated on the following: *emotions* (pride, envy, pity, disgust), *warmth* (warm, trustworthy, friendly), *competence* (competent, capable, skilled), *social interaction* (similarity, familiarity, likelihood of interacting), and how *typically human* they appeared. The first two scales resulted from synonyms used in previous SCM studies to measure warmth and competence (see Fiske et al., 2002). Each scale had a Cronbach's alpha of at least .89.

Each type of social target elicited some degree of each emotion (Harris, 2007). Also, for the dehumanised people, 92% of their pictures elicited more disgust that the other three emotions. As would be expected, based on status and competence, the social interaction measure correlated positively with pride and envy, but negatively with pity and disgust (the low-competence half of the space). This pattern shows the most likelihood of interaction, similarity, and familiarity when the social target elicits pride, and the least when the target elicits disgust. The two ambivalent emotions, although the effect sizes are smaller, suggest closer social interaction ties to envy than to pity. These findings suggest that the emotion disgust may drive the dehumanised perception.

Social neuroscience data on dehumanised perception

Building on social psychology, the reciprocal nature of social neuroscience appears in a series of studies aimed at demonstrating dehumanised

perception. The social psychological theory described thus far makes predictions about the neuroscience data, as discussed next.

Neural indicator of dehumanised perception. To this point, we have not presented direct neuroimaging evidence of dehumanised perception. We have spent a significant part of the article describing a possible neural indicator of the phenomenon, reduced MPFC activation. We tested the hypothesis that MPFC activation to social groups that elicit disgust may be reduced, compared to other social groups that elicit a more ambivalent social emotion, and to a fixation baseline. In several studies, participants reported how they felt about pictures of social targets representing each quadrant of SCM space, while recording neural activity (Harris & Fiske, 2006, 2007; Harris et al., 2007). For instance, in one study (Harris & Fiske, 2006), participants saw pictures of social targets, faces expressing neutral emotion, landscapes, and carnival/Venetian-styled face-masks. Participants were asked to indicate via button press while viewing the images in the scanner which of four emotions (pride, envy, pity, and disgust) they felt towards each picture.

Behavioural data indicated that the social targets each elicited the predicted emotion above chance level. Accordingly, social targets that elicited pride, envy and pity, all activated MPFC above fixation and landscape baselines, and were more active in three versus one deviant cell contrast analyses. But not all targets activated this mentalising region: No significant MPFC activity emerged above fixation baseline or in the deviant cell contrast analysis in response to pictures of social targets who elicited disgust. Instead, these targets activated the amygdala and insula, areas associated with vigilance and visceral disgust. Moreover, a prior study showed a similar pattern of neural activity for disgusting objects. There was not a complete lack of MPFC activity to disgust-inducing social targets; a subsequent more focused region of interest (ROI) analysis within the overlapping area of MPFC that had activated in response to the other social targets above fixation baseline did reveal a much smaller effect size (roughly half) for the dehumanised targets, compared with the other three. We take these results as a neural indicator of dehumanised perception because a brain region involved in social cognition, including mentalising and person perception as discussed above, is less active to these social targets.

Background on the amygdala and insula. Dehumanised perception describes a failure to infer the contents of a social target's mind. If participants in our studies are not thinking about the mind of homeless people and drug addicts, are they simply avoiding thoughts about the target, suppressing inferences about the social target's mind? Or are they just very unfamiliar with the social target, and therefore have no relevant mental substance for spontaneous mentalising? What other brain areas activate

while the MPFC deactivates to dehumanised targets? The amygdala and the insula both are active in our studies, and both are affective neural regions of the person perception neural network.

Perhaps an over-simplified analogy of a security alarm approximately describes the function of the amygdala; it becomes active whenever emotionally significant stimuli are present (Whalen, 1998). This vigilance idea expands on its role in fear conditioning (Phelps, 2006a, 2006b); its specific function in approach–avoid affect will continually refine as the field advances. At a minimum, the amygdala receives direct input from the thalamus (a relay station in the brain), and from its seat in temporal cortex, projects to almost all of the brain directly or indirectly, and tracks trustworthiness judgements (Engell et al., 2007; Winston et al., 1998) and implicit bias (Phelps et al., 2000), for example. Untrustworthy and outgroup people both require vigilance.

The amygdala allows social interaction and perhaps triggers mentalising. Patients with amygdala lesions show ToM impairments (Fine, Lumsden, & Blair, 2001; Stone, Baron-Cohen, Calder, Keane, & Young, 2003), cannot attribute social meaning to animated geometrical shapes (Heberlein et al., 1998), cannot reason about social exchange (Stone, Cosmides, Tooby, Kroll, & Knight, 2002), and cannot process emotion from faces (Young et al., 1995). Similarly, primates decrease social functioning following amygdala lesions (Brothers, Ring, & Kling, 1990; Dicks, Myers, & Kling, 1969; Emery et al., 2001; Kling, Lancaster, & Benitone, 1970; Kling & Steklis, 1976). This suggests that the amygdala plays a key role in mental state inference and general social information processing.

The insula, like the amygdala, is a primary affective area, but the insula reliably attunes to disgust (Phan, Wager, Taylor, & Liberzon, 2002). It also activates to a range of affective stimuli and sometimes links to the basic negative emotion anger (Phan et al., 2002). Moreover, patients with damage to insula and striatal regions do not recognise disgust facial expressions, and do not report experiencing disgust, suggesting that this region is necessary for disgust (Calder, Keane, Manes, Antoun, & Young, 2000). More recently, the insula has been considered a key region for interoception as part of a punishment neural network (see Seymour, Singer, & Dolan, 2007). As a summary, the insula reliably indicates (negative) arousal. Unlike the amygdala, this brain region is not a bundle of sub-cortical nuclei, but is housed in the cortex, although close to sub-cortical structures.

The insula also allows empathising, as described above. Research has demonstrated that the insula is more active when participants witness a close other experience a negative event, specifically a mild shock (Singer et al., 2006). Although psychologists have considered empathy akin to mentalising, neuroscientists view mentalising as the more cognitive component of the more affective empathy system, which involves a number of areas in addition to the insula (Singer, 2006).

In racial prejudice, these areas of the brain associate with approach–avoidance affect, and are both more active when White participants look at unfamiliar Black than White male faces (Hart et al., 2000; Lieberman, Hariri, Jarcho, Eisenberger, & Bookheimer, 2005; Phelps et al., 2000; Wheeler & Fiske, 2005). Moreover, these activations correlate with implicit measures of bias and fear such as the implicit association test (IAT) and startle-eye blink response (Phelps et al., 2000). Notably, the race-amygdala effect attenuates for familiar faces (Phelps et al., 2000), repeated faces (Hart et al., 2000), and individuated faces (Wheeler & Fiske, 2005). Therefore these affective areas of the person perception neural network may be crucial indicators of dehumanised perception.

Neural malleability in dehumanised perception. Previous social neuroscience shows that neural activation during intergroup bias, specifically racial bias paradigms, is malleable (Hart et al., 2000; Phelps et al., 2000; Wheeler & Fiske, 2005). Moreover, the neural areas that changed their activation patterns were areas more active to our dehumanised targets, further evidence that this dehumanisation bias could be influenced. One such paradigm (Wheeler & Fiske, 2005; for replication, see Harris & Fiske, unpublished; see Figure 4) shows that target-race difference in the amygdala and insula depends on the task: It reduces when participants individuate the social target by inferring the vegetable preference of pictured Black people (Harris & Fiske, unpublished; Wheeler & Fiske, 2005). This paradigm also successfully changed neural responses to dehumanised targets. Participants made vegetable-preference judgements or categorical judgements about dehumanised and other targets. A picture of a vegetable flashed before each pictured target. In the preference condition, participants indicated with a button press whether the targets would like the vegetable. To make this judgement, participants had to infer the target's preference without using the information conveyed by the social category (the vegetables were not associated with stereotypes). In the categorical condition, participants indicated whether the pictured social target was over or under middle age. To make this judgement, participants had to rely on external features of the target without getting into their minds, so this more superficial judgement should simply activate the social category.

Comparing neural responses to the dehumanised targets in the preference condition to activation in the categorical condition resulted in an area of MPFC being activated more for the preference judgements (Harris & Fiske, 2007). At least temporarily, perceivers could mentalise the dehumanised targets. In this instance participants often report using either their own preferences, or those of familiar others who resemble the targets, to make their response. The activated MPFC area was similar to the same contrast for the other social targets (Harris & Fiske, 2007).

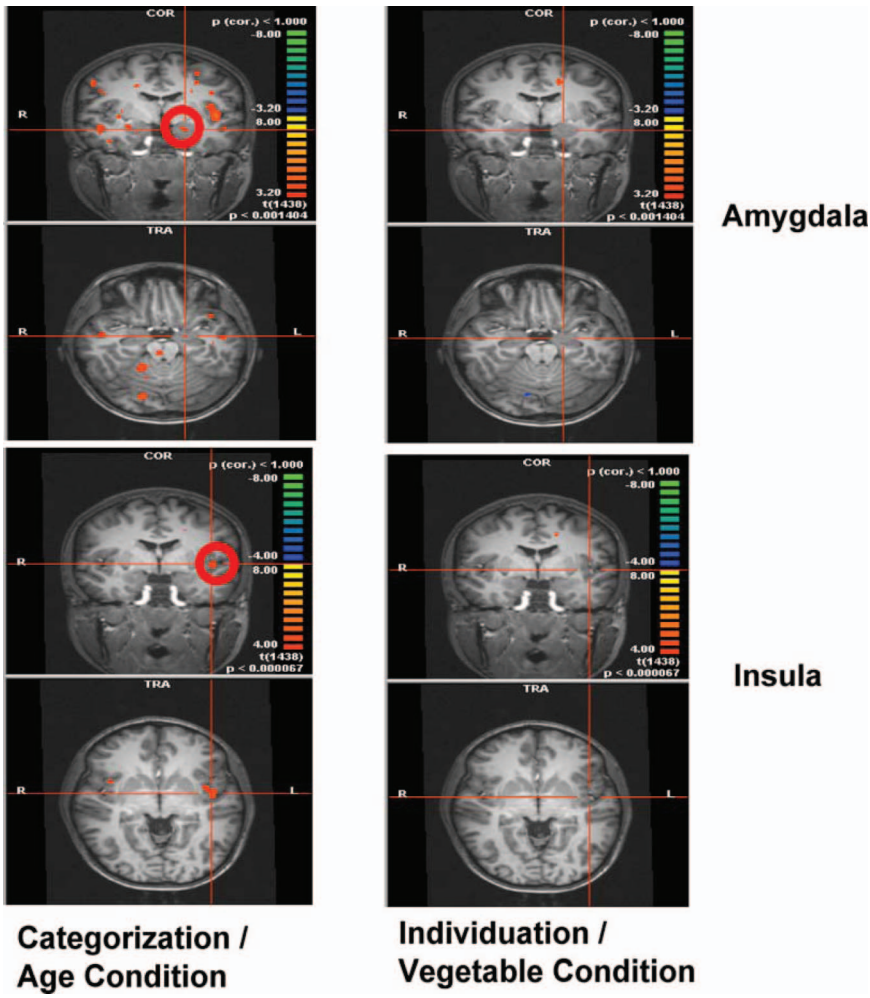

Figure 4. Manipulating amygdala and insula (see online for colour version). These contrast maps are the result of comparing black to white faces, with areas in orange showing more activation to black than white faces. The areas in the crosshairs show amygdala and insula activity change resulting from having participant categorise or individuate black and white faces. Participants made categorical over/under high school age judgements and individuated like/dislike vegetables judgements. There is significantly more activation in both brain regions to black faces than white faces after categorical versus individuated judgements. Each image also shows the *p*-value.

Electroencephalography (EEG)

Participants reported how they felt about pictures of social targets representing each quadrant of SCM space, while recording neural activity

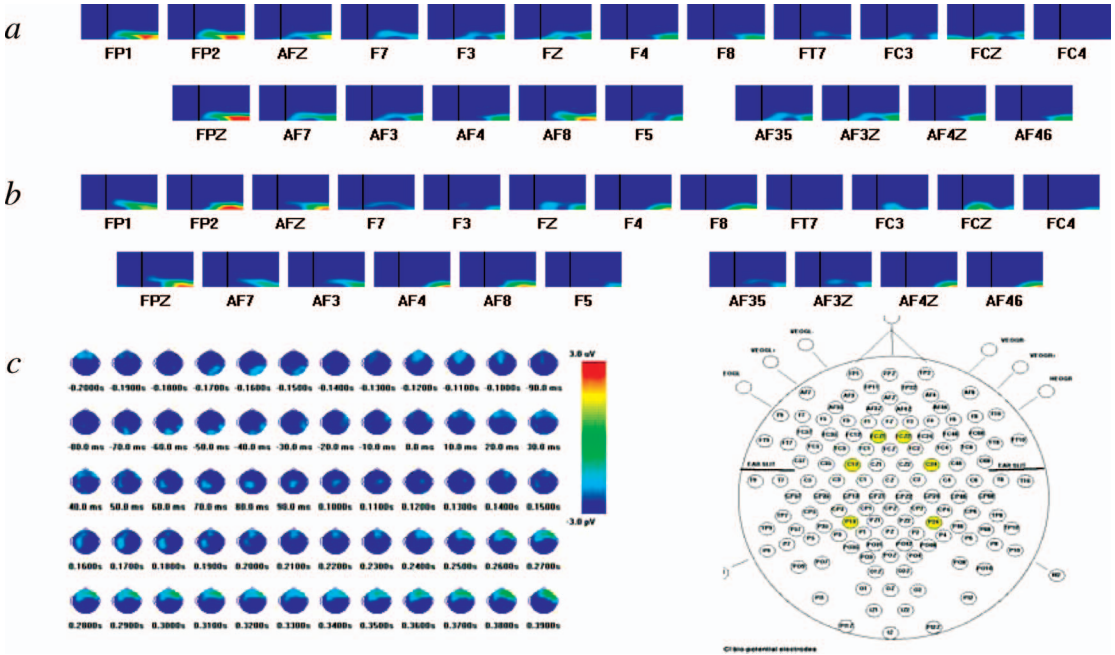

Figure 5. EEG activity to dehumanised groups (see online for colour version). Participants in the EEG study simply looked at pictures of social targets and reported the evoked emotion. (a) Average α waves in response to pride, pity, and envy targets, recorded from the frontal regions identified on the map in the bottom right corner. (b) Average α waves in response to disgust targets, recorded from the frontal regions identified on the map in the bottom right corner. (c) Heat maps showing areas of recorded signal 2000 milliseconds before through 3900 milliseconds after stimulus presentation. In all maps, brighter colours indicate more activity: a and b show less activity to dehumanised targets recorded in electrodes over frontal cortex, notably FPZ, FP1, and FP2; c shows more activity to social targets that are not dehumanised. Activity to all targets begins around 80 ms, over visual cortex, c shows differential activity to dehumanised targets beginning around 1600 ms over temporal cortex, then around 2400 ms over frontal cortex. Yellow electrodes in the array were not operationable.

with EEG (Harris, Gelfand, Escobedo, & Fiske, unpublished). The EEG data showed essentially compatible results with the previously described fMRI studies, and demonstrate a rapid time-frame for these responses. There is a differentiation in electrodes over frontal cortex approximately 100 ms after presentation of the social target. A first negative component responded maximally to those who elicited the complex social emotions of pride, envy, and pity, followed by a second negative component at 300 ms spread over most of the scalp. This latter component responded maximally to the dehumanised targets who elicited the simple basic negative emotion of disgust (see Figure 5 for scalp maps and intensity graphs). But does this evidence together with all the other evidence we described really indicate that these social targets are perceived as less human? Perhaps.

CONCLUSIONS

Thus far we have presented arguments and data supporting the claim that some social targets are not perceived as typically human to the same extent as other social targets. Specifically, people do not imagine the mental states of these dehumanised targets. This phenomenon, dehumanised perception, responds to social targets belonging to social groups perceived as low on trait warmth and competence that elicit the basic emotion disgust. Participants use fewer mental state verbs when describing these targets, and rate them lower on humanising dimensions. Furthermore, there is reduced activation in the MPFC, an area of the brain necessary for social cognition, when people look at these dehumanised targets, compared to other social targets and resting baselines. This difference occurs about 150 milliseconds after stimulus presentation. However, the MPFC can be reactivated to these dehumanised targets if perceivers spontaneously mentalise the targets by inferring their preference. This suggests that the phenomenon is not ''hard-wired'', and provides converging evidence for mentalising as a mediating mechanism necessary for perception of typical humans.

But why would a perceiver not infer the mental state of a target with mental states, given the predictive benefits of doing so? Recent research suggests that dehumanised perception may be a spontaneous regulation strategy. When participants are not able to escape looking at these dehumanised targets, as is the case in all the experiments in this line of research described thus far, participants show neural indicators of dehumanised perception. However, if participants are given the opportunity to escape looking at these targets, participants do not show neural activation patterns indicative of dehumanised perception. Also, physiological markers of disgust tend to be lower when participants engage in dehumanised perception compared to escape, providing further evidence that not perceiving another as human may be a regulation strategy, perhaps aimed at avoiding empathy exhaution (see Batson, 1991).

Limitations of social neuroscience approaches

Reverse inferences—inferring psychological function from brain data—plague any cognitive neuroscience correlational approach (see Poldrack, 2006). This problem can be more severe if the researcher treats the neuroscience data as the sole dependent variable, providing support for the occurrence of a mental process, because these data may mistakenly be viewed as sufficient evidence of the occurrence of that process. There are two ways around this problem. The first is to show via different statistical methods (for example modelling) that the neural region is indeed correlated with the specific variance of the underlying mechanism, satisfying multiple axioms. The other method is to provide converging evidence, and view the neural data as only part of the evidence, but not the whole. Because the brain region may compute additional functions, then researchers can test these alternatives using traditional behavioural methods, and a reciprocal research program is generated. Dissociations within the same paradigm provide strong support for one process over another. The latter has been our preferred strategy.

Having said this, we do not consider social neuroscience a sufficient social psychology, but it does address social psychological theory. Therefore, social psychologists can utilise this technique. It is becoming a theoretically distinct field because the reciprocal nature of the approach generates new theory, theory unique enough to the social neuroscience perspective that it is becoming distinct. The question remains whether social neuroscience in its infancy has developed enough theory to have an identifiably distinct body of work. That question will be answered as the field continues to generate research.

Another interesting aspect of social neuroscience is the parallel emergence of neuroeconomics (see Glimcher, Camerer, Poldrack, & Fehr, 2008, for review of the field). This field addresses questions specific to judgement and decision making, a perspective adjacent to social psychology, although focused on economic theory. Therefore, the technique differs, relying more on regressing economic models on brain data, and holding different requirements for behaviour (if the behaviour has no real-life economic consequences, then it is not a valid reflection of a phenomenon). Is social neuroscience distinct from other emergent techniques such as neuroeconomics? This question remains open.

Societal consequences of dehumanisation

If some affective experience correlates with failure to mentalise, while other affective experiences correlate with mentalising, then affect may be able to adjust dehumanised perception. Disgust as an emotion is associated among other things with the unfamiliar, and a social interaction measure has been inversely associated with disgust, suggesting that increasing familiarity with

these targets may change how they are perceived. This relationship also suggests that affective responses to social targets are learned.

An investigation of social learning may reveal how these associations are learned (possibly from the media, socialisation processes, associations in culture), and how they may be extinguished. A possible societal consequence of this dehumanisation may be reduced helping and neglect of dehumanised groups. Disgust predicts both passive and active harm (Cuddy et al., 2008), and people report less willingness to help dehumanised groups (Cuddy et al., 2007). This suggests that problems like homelessness represent a challenge because people are not ordinarily motivated to help when thinking about members of this category.

To return to a variant of our earlier robot example, imagine a computer that plays chess and has an occasional conversation with its human competitor. Would it be egregious to dismember the computer? Imagine a friend with whom the competitor engages in the same activities; now consider dismembering the friend. The second consideration seems unpleasant because the friend is a person, and the computer is not. But what separates the two? A human essence is arguably the distinguishing factor, but what exactly is a human essence? Consider that the computer has intentions, just as the friend. But also consider that the computer cannot suffer, as the friend can suffer. Because people are aware of phenomenological states such as anger, pain, sadness, pleasure, and so on, they can identify when another person experiences these states. Furthermore, if the perceiver has experienced these phenomenological states resulting in a kind of spontaneous empathy, then people enjoy witnessing others' appetitive states such as pleasure, except for instances of schadenfreude triggered by envy from social comparison where these preferences are reversed, but dislike witnessing others' aversive states such as pain. Spontaneous vicarious reactions suggest empathy that distinguishes between immediate "hot", affective reactions that occur without intent, and intentional cognitions that can be "cold" and devoid of affect (Frith, 2003). Because people have hot empathic reactions to phenomenal states, perhaps this underlies their moral intuitions—their sense of right and wrong—a sense of morality (Haidt, 2007).

Final thoughts

The nature of perceived humanity is a fundamental question in a number of fields of scientific investigation, but often lies dormant and partially ignored because of the magnitude of the question. A social neuroscience approach, because of its reciprocal interdisciplinary nature, begins to strip away some of the mystique from large questions like this, making it possible to ask subsequent research questions. Therefore we also demonstrate a method that can tackle this difficult question concerning people. Emotions,

intentionality, moral psychology, prejudice and inter-group relations, person perception, economic value, all relate to this research and require consideration in the research process. This makes social neuroscience a valuable technique for exploring social psychological questions because it generates new data and theory due to its interdisciplinary nature. By using philosophy as a mediator between social psychology and neuroscience, future generations of scientists may continue to unlock the fields' most complicated and important questions.

REFERENCES

Adolphs, R. (2004). Emotion, social cognition, and the human brain. In J. Cacioppo & G. Berntson (Eds.), *Essays in social neuroscience* (pp. 121–132). Cambridge, MA: MIT Press.

Alexander, M. G., Brewer, M. B., & Hermann, R. K. (1999). Images and affect: A functional analysis of out-group stereotypes. *Journal of Personality and Social Psychology*, *77*, 78–93.

Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behaviour and physical attractiveness. *Journal of Personality and Social Psychology*, *64*, 431–441.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews, Neuroscience*, *7*, 268–277.

Appiah, K. A. (2003). *Thinking it through: An introduction to contemporary philosophy*. New York: Oxford University Press.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, *41*, 258–290.

Bar-Tal, D. (1990). Delegitimization: The extreme case of stereotyping and prejudice. In D. Bar-Tal, C. Graumann, A. Kruglanski & W. Stroebe (Eds.), *Stereotyping and prejudice: Changing conceptions*. New York: Spinger-Verlag.

Batson, C. D. (1991). *The altruism question: Towards a social-psychological answer*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Batson, C. D. (1998). Altruism and prosocial behaviour. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.) *The handbook of social psychology* (Vol. 2, pp. 282–316). New York: McGraw-Hill.

Brothers, L. A., Ring, B., & Kling, A. S. (1990). Response of neurons in the macaque amygdala to complex social stimuli. *Behavioural Brain Research*, *41*, 199–213.

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, *3*, 1077–1078.

Castano, E., & Giner-Sorolla, R. (2006). Not Quite Human: Infrahumanisation in response to collective responsibility for intergroup killing. *Journal of Personality and Social Psychology*, *90*, 804–818.

Cortes, B. P., Demoulin, S., Rodriguez, R. T., Rodriguez, A. P., & Leyens, J. P. (2005). Infra-humanisation or familiarity? Attribution of uniquely human emotions to the self, the ingroup, and the outgroup. *Personality and Social Psychology Bulletin*, *31*, 243–253.

Cuddy, A. J., Fiske, S. T., Glick, P. (2007a). The BIAS Map: Behaviours from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*, 631–648.

Cuddy, A. J., Fiske, S. T., Kwan V. S., Glick, P., Demoulin, S., Leyens, J. P., et al. (2009). Stereotype content model holds across cultures: Universal similarities and some differences. *British Journal of Social Psychology*, *48*, 1–33.

Cuddy, A. J. C., Rock, M. S., & Norton, M. I. (2007b). Aid in the aftermath of Hurricane Katrina: Inferences of secondary emotions and intergroup Helping. *Group Processes and Intergroup Relations*, *10*, 107–118.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioural and Cognitive Neuroscience Reviews*, *3*, 406–412.

Decety, L., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science*, *15*, 54–58.

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Science*, *1104*, 70–88.

Demoulin, S., Torres, R. R., Perez, A. R., Vaes, J., Paladino, M. P., Gaunt, R., et al. (2005). Emotional prejudice can lead to infrahumanisation. *European Review of Social Psychology*, *15*, 259–296.

Dennett, D. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Dicks, D., Myers, R., & Kling, A. S. (1969). Uncus and amygdala lesions: Effects on social behaviour in the free-ranging rhesus monkey. *Science*, *165*, 69–71.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*, 1082–1090.

Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes towards women and men. *Personality and Social Psychology Bulletin*, *15*, 543–558.

Emery, N. J., Capitanio, J. P., Mason, W. A., Machado, C. J., Mendoza, S. P., & Amaral, D. G. (2001). The effects of bilateral lesions of the amygdala on dyadic social interactions in rhesus monkeys (*Macaca mulatta*). *Behavioural Neuroscience*, *115*, 515–544.

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, *19*, 1508–1519.

Fine, C., Lumsden, J., & Blair, R. J. (2001). Dissociation between 'theory of mind' and executive functions in a patient with early left amygdala damage. *Brain*, *124*, 287–298.

Fiske, S. T. (2004). *Social beings: A core motives approach to social psychology*. New York: Wiley.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Science*, *11*, 77–83.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878–902.

Fiske, S. T., Harris, L. T., & Cuddy, A. J. (2004). Why ordinary people torture enemy prisoners. *Science*, *306*, 1421–1632.

Fiske, S. T., & Taylor, S. E. (2008). *Social cognition: From brains to culture*. Boston: McGraw-Hill.

Frith, C. (2003). The scientific study of consciousness. In M. A. Ron & T. W. Robbins (Eds.), *Disorders of brain and mind* (2nd ed., pp. 197–222). Cambridge, UK: Cambridge University Press.

Frith, U., & Frith, C. (2001). The biological basis of social interaction. *Current Directions in Psychological Science*, *10*, 151–155.

Galinsky, A. D., Ku, G., Wang, C. S. (2005). Perspective-taking and the self–other overlap: Fostering social bonds and facilitation social coordination. *Group Processes and Intergroup Relations*, *8*, 109–124.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*, 708–724.

Glimcher, P., Camerer, C., Poldrack, R., & Fehr, E. (2008). *Neuroeconomics: Decision making and the brain*. Oxford, UK: Elsevier.

Goff, P. A., Eberhardt, J. L., Williams, M., & Jackson, M. C. (2008). Not yet human: Implicit knowledge, historical dehumanisation, and contemporary consequences. *Journal of Personality and Social Psychology*, *94*, 292–306.

Gosling, S. D., Kwan, V. S., & John, O. P. (2003). A dog's got personality: A cross-species comparative approach to personality judgements in dogs and humans. *Journal of Personality and Social Psychology*, *85*, 1161–1169.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(14), 2105–2108.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*, 998–1002.

Haidt, J., Rozin, P., McCauley, C., & Imada, S. (1997). Body, psyche, and culture: The relationship between disgust and morality. *Psychology and Developing Societies*, *9*, 107–131.

Hamilton, D. L. (1981). *Cognitive processes in stereotyping and intergroup behaviour*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Harris, L. T. (2007). *Dehumanized perception fails to represent the contents of a social target's mind*. Dissertation thesis, Princeton University, NJ, USA.

Harris, L. T., & Fiske, L. T. (2006). Dehumanising the lowest of the low: Neuroimaging responses to extreme outgroups. *Psychological Science*, *17*, 847–853.

Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Social Cognitive and Affective Neuroscience*, *2*, 45–51.

Harris, L. T., McClure, S., Van den Bos, W., Cohen, J. D., & Fiske, S. T. (2007). Regions of MPFC differentially tuned to social and nonsocial affective stimuli. *Cognitive and Behavioural Neuroscience*, *7*, 309–316.

Harris, L. T., Todorov, A., & Fiske, S. T. (2005). Attributions on the brain: Neuro-imaging dispositional inferences beyond theory of mind. *NeuroImage*, *28*, 763–769.

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs. ingroup face stimuli. *Brain Imaging*, *11*, 2351–2355.

Haslam, N. (2006). Dehumanisation: An integrative review. *Personality and Social Psychology Review*, *10*, 252–264.

Haslam, N., Bain, P., Douge, L., Lee, M., & Bastian, B. (2005). More human than you: Attributing humanness to self and others. *Journal of Personality and Social Psychology*, *89*, 937–950.

Haslam, N., Bain, P., Loughnan, S., & Kashima, Y. (2008). Attributing and denying humanness to others. *European Review of Social Psychology*, *19*, 55–85.

Hauser, M. D. (2006). The liver and the moral organ. *Social Cognitive and Affective Neuroscience*, *1*, 214–220.

Haxby, J. V., Gobbini, M. I., & Montgomery, K. (2004). Spatial and temporal distribution of face and object representations in the human brain. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 889–904). Cambridge, MA: MIT Press.

Heberlein, A. S., Adolphs, R., Tranel, D., & Damasio, H. (2004). Cortical regions for judgements of emotions and personality traits from point-light walkers. *Journal of Cognitive Neuroscience*, *16*, 1143–1158.

Heberlein, S. A., Adolphs, R., Tranel, D., Kemmerer, D., Anderson, S., & Damasio, A. R. (1998). Impaired attribution of social meanings to abstract dynamic geometric patterns following damage to the amygdala. *Society for Neuroscience Abstracts*, *24*, 1176.

Heider, F. (1958). *The psychology of interpersonal relations*. Wiley, New York.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behaviour. *American Journal of Psychology*, *57*, 243–259.

Jackson, L. A., Hodge, C. N., Gerard, D. A., Ingram, J. M., Ervin, K. S., & Sheppard, L. A. (1996). Cognition, affect, and behaviour in the prediction of group attitudes. *Personality and Social Psychology Bulletin*, *22*, 306–316.

Johansson, M., Mecklinger, A., & Treese, A. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience*, *16*, 1840–1853.

Johnson, S. C., Baxter, L. C., Wilder, L. S., Pipe, J. G., Heiserman, J. E., & Prigatano, G. P. (2002). Neural correlates of self-reflection. *Brain*, *125*, 1808–1814.

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, *34*, 107–117.

Kelley, H. H. (1972). Attribution in social interaction. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the cause of behaviour* (pp. 1–26). Hillsdale, NJ: Lawrence Erlbaum & Associates Inc.

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*, 785–794.

Kling, A. S., Lancaster, J., & Benitone, J. (1970). Amygdalectomy in the free-ranging vervet. *Journal of Psychiatric Research*, *7*, 191–199.

Kling, A. S., & Steklis, H. D. (1976). A neural substrate for affiliative behaviour in non-human primates. *Brain Behaviour and Evolution*, *13*, 216–238.

Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. *Theory Into Practice*, *16*, 53–59.

Kwan, V. S. Y., & Fiske, S. T. (2008). Missing links in social cognition: The continuum from nonhuman agents to dehumanised humans. *Social Cognition*, *26*, 125–128.

Kwan, V. S. Y., John, O. P., Kenny, D. A., Bond, M. H., Robins, R. W. (2004). Reconceptualizing individual differences in self-enhancement bias: An interpersonal approach. *Psychological Review*, *111*, 94–110.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-6*. Gainesville, FL: University of Florida.

LeDoux, J. (1998). Fear and the brain: Where have we been, and where are we going? *Biological Psychiatry*, *44*, 1229–1238.

Leibenluft, E., Gobbini, M. I., Harrison, T., & Haxby, J. V. (2004). Mothers' neural activation in response to pictures of their children and other children. *Biological Psychiatry*, *56*, 225–232.

Levenson, R. W. (1996). Biological substrates of empathy and facial modulation of emotion: Two facets of the scientific legacy of John Lanzetta. *Motivation and Emotion*, *20*, 185–204.

Leyens, J. P., Cortes, B. P., Demoulin, S., Dovidio, J., Fiske, S. T., Gaunt, R., et al. (2003). Emotional prejudice, essentialism, and nationalism. *European Journal of Social Psychology*, *33*, 703–718.

Leyens, J. P., Rodriguez-Perez, A., Rodriguez-Torres, R., Gaunt, R., Paladino, M. P., Vaes, J., et al. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology*, *31*, 395–411.

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, *58*, 259–289.

Lieberman, M. D., Hariri, A., Jarcho, J. M., Eisenberger, N. I., & Bookheimer, S. Y. (2005). An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nature Neuroscience*, *8*, 720–722.

Lieberman, M. D., Jarcho, J. M., & Satpute, A. B. (2004). Evidence-based and intuition-based self-knowledge: An fMRI study. *Journal of Personality and Social Psychology*, *87*, 421–435.

Loughnan, S., & Haslam, N. (2007). Animals and androids: Implicit associations between social categories and nonhumans. *Psychological Science*, *18*, 116–121.

Mackie, D. M., & Smith, E. R. (2002). Intergroup emotions: Prejudice reconceptualized as differentiated reactions to out-groups. In J. P. Forgas & K. D. Williams (Eds.), *The social self: Cognitive, interpersonal, and intergroup perspectives* (pp. 309–326). Philadelphia: Psychology Press.

Martin, J., Bennett, M., & Murray, W. S. (2008). A developmental study of the infrahumanisation hypothesis. *British Journal of Developmental Psychology*, *26*, 153–161.

Mason, M. F., Hood, B. M., & Macrae, C. N. (2004). Look into my eyes: Gaze direction and person memory. *Memory*, *12*, 637–643.

Mason, M. F., & Macrae, C. N. (2004). Categorising and individuating others: The neural substrates of person perception. *Journal of Cognitive Neuroscience*, *16*, 1785–1795.

McArthur, L. A. (1972). The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, *72*, 171–193.

McClure, S. M., Laibson D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate & delayed monetary rewards. *Science*, *306*(5695), 503–507.

Meiser, T. (2003). Effects of processing strategy on episodic memory and contingency learning in group stereotype formation. *Social Cognition*, *21*, 121–156.

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*, 1306–1315.

Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2004). Distinct neural systems subserve person and object knowledge. In J. Cacioppo & G. Berntson (Eds.), *Essays in social neuroscience*. Cambridge, MA: MIT Press.

Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science*, *11*, 462–466.

Neuberg, S. L., & Cottrell, C. A. (2002). Intergroup emotions: A socio-functional approach. In D. M. Mackie & E. R. Smith (Eds.), *From prejudice to intergroup relations: Differentiated reactions to social groups* (pp. 265–283). New York: Psychology Press.

Niedenthal, P., Krauth-Gruber, S., & Ric, F. (2006). *Psychology of emotion: Interpersonal, experiential and cognitive approaches*. New York: Psychology Press.

NIH. (1987). *The health benefits of pets. Workshop summary; September 10–11*. Bethesda, MD: National Institutes of Health, Office of Medical Applications of Research.

Ochsner, K. N. (2007). Social cognitive neuroscience: Historical development, core principles, and future promise. In A. Kruglanksi, & E. T. Higgins (Eds.), *Social psychology: A handbook of basic principles* (2nd ed., pp. 39–66). New York: Guilford Press.

Ochsner, K. N., Knierim, K., Ludlow, D. H., Hanelin, J., Ramachandran, T., Glover, G. et al. (2004). Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience*, *16*, 1746–1772.

Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist*, *56*, 717–734.

Olsson, A., & Ochsner, K. N. (2008). The role of social cognition in emotion. *Trends in Cognitive Sciences*, *12*, 65–71.

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, *46*, 1–20.

Peeters, G. (1983). Relational and informational patterns in social cognition. In W. Doise & S. Moscovici (Eds.), *Current issues in European social psychology* (Vol. 1, pp. 201–237). Cambridge, UK: Cambridge University Press.

Pelham, B. W., Mirenberg, M. C., & Jones, J. T. (2002). Why Susie sells seashells by the seashore: Implicit egotism and major life decisions. *Journal of Personality and Social Psychology*, *82*, 469–487.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, *16*, 331–348.

Phelps, E. A. (2006a). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology*, *57*, 27–53.

Phelps, E. A. (2006b). Emotion, learning, and the brain: From classical conditioning to cultural bias. In P. Baltes, P. Reuter-Lorenz, & F. Rosler (Eds.), *Lifespan development and the brain: The perspective of biocultural co-constructivism* (pp. 200–216). New York: Cambridge University Press.

Phelps, E. A., & LaBar, K. S. (2006). Functional neuroimaging of emotion and social cognition. In R. Cabeza & A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition* (pp. 421–453). Cambridge, MA: MIT Press.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., et al. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, *12*, 729–738.

Plaks, J. E., Grant, H., & Dweck, C. S. (2005). Violations of implicit theories and the sense of prediction and control: Implications for motivated person perception. *Journal of Personality and Social Psychology*, *88*, 245–262.

Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*, 59–63.

Richardson, D. R., Hammock, G. S., Smith, S. M., & Gardner, W. (1994). Empathy as a cognitive inhibitor of interpersonal aggression. *Aggressive Behaviour*, *20*, 275–289.

Rodríguez, A., Coello, E., Betancor, V., Rodríguez, R., & Delgado, N. (2006). When is the infrahumanisation to a threatening outgroup higher? Effects of the ambiguity of situation and the empathy with ingroup. *Psicothema*, *18*, 73–77.

Rodríguez-Torres, R., Leyens, J. P., Rodríguez Pérez, A., Betancor Rodriguez, V., Quiles del Castillo, M., Demoulin, S. et al. (2005). The lay distinction between primary and secondary emotions: A spontaneous categorization. *International Journal of Psychology*, *40*(2), 100–107.

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*, 283–294.

Ruby, P., & Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience*, *16*, 988–999.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision making in the ultimatum game. *Science*, *300*, 1755–1757.

Sears, D. O. (1983). The person-positivity bias. *Journal of Personality and Social Psychology*, *44*, 233–250.

Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, *54*, 558–568.

Seymour, B., Singer, T., & Dolan, R. (2007). The neurobiology of punishment. *Nature Reviews Neuroscience*, *8*, 300–311.

Singer, T. (2006). The neural basis and ontogeny of empathy and mind reading: Review of literature and implication for future research. *Neuroscience and Biobehavioral Reviews*, *30*, 855–863.

Singer, T., & Fehr, E. (2005). The neuroeconomics of mind reading and empathy. *American Economic Review*, *95*, 340–345.

Singer, T., Seymour, B., O'Doherty, J., Klaas, E. S., Dolan, J. D., Frith, C. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, *439*, 466–469.

Somerville, L. H., Wig, G. S., Whalen, P. J., & Kelley, W. M. (2006). Dissociable medial temporal lobe contributions to social memory. *Journal of Cognitive Neuroscience*, *18*, 1253–1265.

Stangor, C., Sullivan, L. A., & Ford, T. E. (1991). Affective and cognitive determinants of prejudice. *Social Cognition*, *9*, 359–380.

Staub, E. (1989). *The roots of evil: The origins of genocide and other group violence*. New York: Cambridge University Press.

Stephan, W. G., & Stephan, C. W. (2000). An integrated theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23–45). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Stinson, L., & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology*, *62*, 787–797.

Stone, V. E., Baron-Cohen, S., Calder, A., Keane, J., &Young, A. (2003). Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia*, *41*, 209–220.

Stone, V. E., Cosmides, L., Tooby, J., Kroll, N., & Knight, R. T. (2002). Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage. *Proceedings of the National Academy of Sciences*, *99*, 17, 11531–11536.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.

Tam, T., Hewstone, M., Cairns, E., Tausch, N., Maio, G., & Kenworthy, J. (2007). The impact of intergroup emotions on forgiveness in Northern Ireland. *Group Processes and Intergroup Relations*, *10*, 119–136.

Tausch, N., Kenworthy, J., & Hewstone, M. (2007). The confirmability and disconfirmability of trait concepts revisited: Does content matter? *Journal of Personality and Social Psychology*, *92*, 554–556.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.

Taylor, S. E., & Gollwitzer, P. M. (1995). Effects of mindset on positive illusions. *Journal of Personality and Social Psychology*, *69*, 213–226.

Thaler, R. (1980). Towards a positive theory of consumer choice. *Journal of Economic Behaviour and Organization*, *1*, 39–60.

Todd, R. M., Lewis, M. D., Meusel, L., & Zelazo, P. D. (2008). The time course of social-emotional processing in early childhood: ERP responses to facial affect and personal familiarity in a go-nogo task. *Neuropsychologia*, *46*, 595–613.

Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, *45*, 163–173.

Todorov, A., Harris, L. T., & Fiske, S. T. (2006). Towards behaviourally inspired social neuroscience. *Brain Research*, *1079*, 76–85.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*, 1623–1626.

Todorov, A., & Uleman, J. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, *83*, 1051–1065.

Turkle, S. (2005). *The second self: Computers and the human spirit*. Cambridge, MA: MIT Press.

Vaes, J., Paladino, M. P., Castelli, L., Leyens, J. Ph., Giovanazzi, A. (2003). On the behavioural consequences of infra-humanisation: The implicit role of uniquely human emotions in intergroup relations. *Journal of Personality and Social Psychology*, *85*, 1016–1034.

Van den Bos, W., McClure, S., Harris, L T., Fiske, S. T., Nystrom, L., & Cohen, J. D. (2007). Dissociating affective evaluation and social cognitive processes in ventral medial prefrontal cortex. *Cognitive and Behavioral Neuroscience*, *7*, 337–346.

van Knippenberg, A., van Twuyver, M., & Pepels, J. (1994). Factors affecting social categorisation processes in memory. *British Journal of Social Psychology*, *33*, 419–431.

von Hecker, U., & Dutke, S. (2004). Integrative social perception: Individuals in low working memory benefit from external representations. *Social Cognition*, *22*, 336–365.

Whalen, P. J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current Direction in Psychological Science*, *7*, 177–188.

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice: Social-cognitive goals affect amygdala and stereotype activation. *Psychological Science*, *16*(1), 56–63.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*, 592–598.

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, *5*, 277–283.

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin*, *24*, 1245–1257.

Young, A. W., Aggleton, J. P., Hellawell, D. J., Johnson, M., Broks, P., & Hanley, J. R. (1995). Face processing impairments after amygdalotomy. *Brain*, *118*, 15–24.

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315–334). Cambridge, UK: Cambridge University Press.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*, 1–27.

Zebrowitz, L. A. (1999). *Reading faces: Window to the soul?* Boulder, CO: Westview Press.