

RUNNING HEAD: NEURAL BASIS OF ATTITUDES

The Neural Bases of Attitudes

Emily B.C. Falk¹, Matthew D. Lieberman¹

¹University of California, Los Angeles

Main Text Word Count: 7,290

Belief and the Neural Basis of Evaluation and Attitudes

Attitudes encompass long-standing evaluations of people, places, and ideas, and may influence a range of behaviors, including those that directly impact political behavior, intergroup relations, and health behaviors among other consequences. Attitudes are central in answering questions such as: Where should we invest community resources? Whom should we vote for in the next election? Where will we spend our paychecks? As such, the study of attitudes has captivated thinkers for centuries, and scientists for decades (Allport, 1935; Aristotle, 1924/1954; Hovland, 1949; Hovland, Janis, & Kelley, 1953). Gordon Allport (1935) called attitudes "the most distinctive and indispensable concept in contemporary American social psychology" (p. 798), and suggested that understanding attitudes would allow us to understand not only the preferences and behaviors of individuals, but would also provide broader insight into the actions of groups and cultures. With this in mind, Allport (1935) defined an attitude as "a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon an individual's response to all objects and situations with which it is related" (p. 810).

Following this early work, research has continued to build our understanding of attitudes and attitude change (Albarracin, Johnson, & Zanna, 2005; Eagly & Chaiken, 1993; Eagly & Chaiken, 2005; Petty & Cacioppo, 1986; Petty, Priester, & Wegener, 1994), however, many questions concerning the nature of attitudes, as well as the underlying mechanisms of attitude formation and attitude change remain unanswered (Eagly & Chaiken, 2005; Gawronski, 2007). For example, what is the role of implicit attitudes in influencing explicit attitudes, behaviors,

interactions with other individuals and groups? How do people internally regulate unpleasant or undesirable attitudes and biases? What are the mechanisms through which attitudes form and change, and what are the mechanisms through which external influences, such as persuasive appeals, influence attitudes? While these are clearly complex questions, they are made even more challenging to tackle by demand characteristics, participants' self-presentational concerns and the fact that participants may not have conscious awareness of the ways in which they are processing information. Together, all of these factors may lead to biased self-reports (Krosnick, Judd, & Wittenbrink, 2005).

As foreshadowed by Allport's (1938) definition of attitudes, which includes a "neural state of readiness, organized through experience," the brain may be able to shed some light on unanswered questions that introspection and self-report data have not (Lieberman, 2007; Ochsner & Lieberman, 2001). More specifically, the recent advance of neuroimaging technologies has opened new possibilities to examine multiple psychological processes in concert, to examine the extent to which different phenomena share common or distinct neural bases, and to link theory developed in social psychology to an extensive neuroscience literature developed primarily in animal models.

For example, a vast literature on fear, conditioning and social behavior in animals has been key in informing existing theories of prejudice, bias and social behavior in humans (Amodio & Lieberman, in press). Furthermore, our evolving understanding of the neural bases of automatic and controlled processes has provided insight into the ways in which implicit and explicit evaluations and

attitudes interact. A body of literature is also beginning to form examining the neural correlates of closely related concepts such as the subjective experience of persuasion, attitude change, behavior change, and message propagation. Lastly, the literature addressing the neural mechanisms that support attitudinally relevant processes has reached a stage where integration can begin to take place (Cunningham & Zelazo, 2007; Cunningham, Zelazo, Packer, & Van Bavel, 2007).

Throughout this chapter, we will work from a broad definition of attitudes as evaluative tendencies that can have cognitive, affective and behavioral antecedents and consequences (Eagly & Chaiken, 2007). We will explore areas in which social neuroscience has provided insights regarding the ways that individuals evaluate attitude objects, express or suppress attitudes, and respond to persuasive appeals.

We will begin with the neural systems that support evaluation and the expression of preferences. In this section, we will focus on research elucidating distinctions between automatic and controlled evaluations, as well as overlap with neurocognitive systems supporting affective processing. We will then discuss what is known about the neural systems that regulate the expression of attitudes and evaluations. In this section, we will discuss automatic and controlled processing of attitudes as related to the study of bias and prejudice, and the regulation thereof. Lastly, we will explore internal and external factors that lead to attitude change. In this section, we will review work that explores the neural bases of persuasion, attitude change, and behavior change.

SCN LAB 10/13/08 10:36 AM

Comment: Do we need more re: definition (why this definition, discussion in the attitudes literature re: how to define an attitude? Are attitudes stable entities, etc.?)

The Neural Bases of Evaluation and Preferences

Evaluation and preference are central in defining the nature of attitudes (Eagly & Chaiken, 2007; Eagly & Chaiken, 1993; Eagly & Chaiken, 2005; Petty, et al., 1994; Zajonc & Markus, 1982), although scholars disagree about the extent to which these evaluations and preferences must remain stable to be considered “real” (Bishop, 1980; Bishop, Hamilton, & McConahay, 1980; Converse, 1970). Regardless of the definition used, however, the relationship between evaluations, preferences and attitudes are intertwined; underlying attitudes may predispose individuals to evaluate objects, situations, people or groups more or less favorably, and depending on the evaluation that is made, individuals may update their underlying attitudes.

However, this process cannot be directly observed. Put another way, current researchers behavioral researchers “do not have an inherent psychological reality that can be verified. In other words, researchers cannot directly observe object-evaluation associations, knowledge structures, or microconcepts” (Eagly & Chaiken, 2005, p.746). This is especially true in the case of implicit and unconscious attitudes; while people sometimes deliberately evaluate the world around them, they also automatically and spontaneously make evaluations that are outside the realm of awareness. Thus, while the resulting attitudes may differ in important ways from consciously and deliberately formed opinions, proxy measures of implicit attitudes such as the implicit association task (IAT) have thus far been one of the few means of inferring what individuals are unable, or unwilling to self-report (Bargh, Chaiken, Gollwitzer, & Pratto, 1992; Draine & Greenwald, 1998; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Greenwald & Banaji, 1995).

Building on our growing understanding of the ways that the brain performs automatic and controlled processing, however, a substantial body of neuroimaging research has enriched our understanding of these processes as well. For example, several studies exploring the neural basis of evaluative judgments and preferences focus on explaining differences between the processing of implicit and explicit judgments and the expression of implicit and explicit attitudes and preferences.

Automatic and Controlled Processing in Evaluation and Preference. Across a range of domains including judgments of beauty (Jacobsen, Schubotz, Hofel, & Cramon, 2005; Vartanian & Goel, 2004), evaluations of places, events, and political figures (Zysset, Huber, Ferstl, & von Cramon, 2002), and brand preferences (Mcclure, 2004), when people make explicit evaluations, regions typically associated with controlled processing, including areas of medial prefrontal cortex (MPFC), medial parietal cortex (MPAC), ventrolateral prefrontal cortex (VLPFC), lateral parietal cortex (LPAC) and anterior cingulate cortex (ACC) are engaged. By contrast, tasks that do not explicitly require evaluative judgments, that involve targets that are presented subliminally or when preferences are measured through implicit behavioral means, increased activity is observed in regions typically associated with automatic processing, such as the amygdala and ventromedial prefrontal cortex (VMPFC), as well as the the insula.

Also supporting the role of the VMPFC in implicit evaluation, Koenigs and Tranel (2008) showed that when asked to perform a blind taste test of Coke versus Pepsi, patients with damage in the VMPFC and healthy controls both showed a preference for Pepsi. However, in an open taste test, healthy controls and patients

with lateral brain lesions show a preference for Coke (the so called “Pepsi paradox”), while patients with VMPFC damage maintain their original choices, failing to show typical brand preference effects (Koenigs & Tranel, 2008). Likewise, work by Milne and Grafman (2001) explored gender stereotyping effects in patients with VMPFC damage. In this case, both VMPFC patients and healthy controls demonstrated equal gender stereotyping on an explicit measure, however when performing a gender relevant IAT, patients with VMPFC damage do not show the stereotypic gender associations displayed by healthy control subjects (Milne & Grafman, 2001). This may suggest that elements of stereotyping and prejudice are supported by the same neural mechanisms that support evaluative processing more generally, an issue which will be discussed in more detail later in this chapter. Thus, we see that understanding the basic processes underlying implicit and explicit evaluations could provide insights into scenarios from choosing a soft drink to displaying biased attitudes.

In an early study exploring the differences between implicit and explicit evaluations, Cunningham and colleagues (2003) asked participants to explicitly evaluate famous people on a good/ bad dimension (e.g. Hitler = bad), while on other separate trials participants classified famous people as past/ present (e.g. Hitler = past) (Cunningham, Johnson, Gatenby, Gore, & Banaji, 2003). When comparing brain activity associated with explicit evaluation (Hitler = bad) in contrast to past/present classification (Hitler = past), the researchers observed increased activity in controlled processing regions such as MPFC and VLPFC, as well as the anterior cingulate cortex (ACC). However, regardless of the intention to evaluate (in

both the explicit good/bad evaluative and past/present classification conditions), increased amygdala and insula activity was observed in response to images of famous people who were considered “bad” compared to individuals who were considered “good,” suggesting the likelihood of negative, affective processing (Cunningham, et al., 2003).

Likewise, in a study in which participants evaluated a series of concepts (e.g. murder, happiness, abortion, welfare) on a good/bad dimension as well as categorized concepts on an abstract/concrete dimension, activity in the amygdala was positively correlated with the emotional intensity of the stimulus and insula activity correlated with affective valence across conditions, regardless of intention to evaluate. However, when explicitly evaluating the ideas on a good/bad dimension, to the extent that participants said that they felt ambivalent or reported that they tried to control their evaluation of the topic, increased activity was observed in regions implicated in controlled processing such as the ACC, frontal poles, and VLPFC (Cunningham, Raye, & Johnson, 2004b). These results support the idea that regardless of whether an explicit judgment is made, the brain processes the valence of stimuli, but depending on the demands of a given task or situation, conflicting information and explicit judgments recruit higher level brain networks that may be more sensitive to attitudinal complexity.

Integration of Automatic and Controlled Processing in Evaluation and Preference. While neuroimaging research has allowed both dissociation and simultaneous exploration of automatic and controlled processing, it is not necessarily the case that automatic and controlled aspects of evaluation are

unrelated. In fact, it is likely that evaluations and expressed attitudes at any given time point are the product of interactions between neurocognitive networks that support automatic processing and networks that support controlled processing.

In an effort to integrate the information brought to the fore by neuroimaging regarding the specific ways in which people make evaluations, form attitudes, and change those attitudes, Cunningham, Zelazo and colleagues (2007) have proposed an “iterative reprocessing model” (IR model) of information and affective processing. The IR model posits that at any given time point, evaluations are constructed from an interaction of faster automatic processes (subserved by limbic structures such as the amygdala), and controlled processes (subserved by structures in the PFC and parietal cortex). Cunningham and Zelazo (2007) propose that the way in which these two systems come together depends on time constraints, motivations, and situational factors, and that information is iteratively processed and re-processed to arrive at an evaluation at any given point (Cunningham & Zelazo, 2007; Cunningham, et al., 2007). The authors suggest that we need not conclude (as past research in social cognition has) that implicit and explicit attitudes are fundamentally different entities, but instead that automatic evaluations are important across iterations and are influenced by, as well as influence more controlled processes.

Thus, to summarize, neuroimaging work has contributed to our understanding of the ways in which automatic and controlled processes co-occur to support evaluation and preferences. Regardless of intention to evaluate, the brain seems to register an affective (potentially evaluative) response to target objects in

areas such as the amygdala, VMPFC and insula. Under circumstances in which a more controlled reaction or explicit evaluation is required, areas of the brain that tend to be more involved in controlled processing and conflict monitoring, such as the prefrontal cortex, parietal cortex and anterior cingulate cortex become involved. Researchers have suggested that the two systems interact over a series of iterations in the brain, and that final evaluations are a function of factors such as time constraints, motivations, and the external situation.

The Neural Bases of Responses to Outgroups and the Regulation of Bias

As reviewed above, the study of the neural mechanisms supporting evaluation, preference, and attitude processes spans several different domains. However, the area that has received the most attention under this umbrella is race related attitudes and biases.

In many ways, race-related attitudes are similar to other types of attitudes. For example, they may have affective, cognitive and/or behavioral components, and can be subject to conscious reflection or may reside under the surface. However, strong societal norms surrounding race and prejudice as well as self-presentation concerns on the part of participants create methodological challenges in determining peoples' "real" racial attitudes. Researchers observe discrepancies both between implicit and explicit attitudes surrounding race, and between self-reported attitudes and observed behaviors; most often, individuals simultaneously report unbiased attitudes, but behave in biased ways.

How can we explain this discrepancy? Two decades after writing his seminal chapter on attitudes (Allport, 1935), Allport wrote *The Nature of Prejudice* (1954) in which he reasoned that many white Americans live in a “state of conflict” between explicit opposition to prejudice and an underlying, more subtle form of implicit race bias (Allport, 1954). Over the decades since, research has strongly supported this claim; while old-fashioned racism has decreased in the United States since Allport’s time, a majority of white Americans still exhibit a preference for whites over blacks on implicit evaluation measures (Chen & Bargh, 1997; Devine, 1989; Nosek, Banaji, & Greenwald, 2002), and even individuals who hold explicitly non-racist attitudes and believe in equality may demonstrate biased behaviors towards outgroup members (Amodio, Kubota, Harmon-Jones, & Devine, 2006; Dovidio, Kawakami, & Gaertner, 2002). Many of these behaviors are linked to implicit attitudes, with implicit and explicit attitudes predicting different types of biased behavior, and with implicit racial categorization taking place even when race is irrelevant to task demands (Dickter & Bartholow, 2007; Fazio, Jackson, Dunton, & Williams, 1995).

There are several possible explanations for the discrepancy between implicit and explicit attitudes, and explicit attitudes and behaviors. One is that Americans are still just as prejudiced, but that social norms now preclude the outward expression of racism. A second possible explanation is that even participants who do not hold conscious prejudice have learned cultural associations with different racial groups (e.g. blackness and whiteness). Knowledge of cultural stereotypes may be reflected in response to implicit tasks (e.g. reaction time tasks such as the IAT), which by definition tap into our fast, automatic associations, as well as in more

subtle behaviors that are outside of conscious control (e.g. body language). Amodio and Lieberman (in press), suggest that “the underlying mechanisms for how different forms of bias affect different types of behaviors remain poorly understood, in part because different underlying forms of implicit bias are difficult to parse using behavioral measures” (p. 5), and that this is one area that can especially benefit from the use of neuroimaging.

To this point, several research teams have harnessed neuroimaging as a method for exploring responses to racial outgroups,¹ focusing heavily on the amygdala as a key correlate of race bias. This stems from the amygdala’s role in fear conditioning (Davis, 1992), and the hypothesized relationship between fear, threat, and prejudice (Smith, 1993).²

For example, the first study to explore the relationship between implicit and explicit racial attitudes in the brain was conducted by Phelps and colleagues (Phelps, O'Connor, Cunningham, & Funayama, 2000). In this study, white participants viewed photos of black and white male faces as part of a task that was unrelated to social evaluation. The researchers then had participants complete both an explicit measure of modern racism (the Modern Racism Scale, McConahay, 1986), and two implicit measures of race bias (the IAT and a startle eye blink task). While most participants did not show any bias on the explicit racism measure, many did show bias according to the implicit measures. Interestingly, while there was no main effect of black versus white faces on brain activity across participants, the

amount of bias expressed through implicit measures was positively correlated with amygdala activity.

Subsequent research efforts have also demonstrated relationships between implicit bias and amygdala activity for both racial and non-racial outgroups (Amodio, Harmon-Jones, Devine, & Curtin, 2004; Cunningham, Johnson, Raye, Chris Gatenby, Gore, & Banaji, 2004a; Eberhardt, 2005; Hart, Whalen, Shin, McInerney, Fischer, & Rauch, 2000; Kaplan, Freedman, & Iacoboni, 2007; Phelps, 2001; Phelps, Cannistraci, & Cunningham, 2003; Phelps & Thomas, 2003). It should be noted, however, that while this work is typically characterized as exploring responses to outgroups (and is referred to as such in this chapter), most of the studies reviewed examine responses of white participants to black faces.

Breaking this pattern, work by Lieberman and colleagues (2005) using both African American and Caucasian American participants demonstrated that both African American and White participants showed increased amygdala activity in response to Black faces, suggesting that cultural learning, and not the familiarity of one's own race may be responsible for the responses observed (Lieberman, Hariri, Jarcho, Eisenberger, & Bookheimer, 2005). This is also consistent with behavioral work demonstrating that black participants often hold implicit biases against black targets (Ashburn-Nardo, Knowles, & Monteith, 2003; Livingston, 2002), and with neuroimaging findings that increased amygdala activity is observed when white people respond to darker skinned photographs white people as compared to lighter skinned photographs of white people (Ronquillo, Denson, Lickel, Lu, Nandy, &

Maddox, 2007). By contrast, in a recent study examining attention biases to race targets revealed by event related potentials, white and black participants each showed ingroup/outgroup biases when viewing faces that are of their own or the opposite race, suggesting that while in some cases black participants show bias against black targets, it is also true that ingroup/outgroup effects take precedence under other circumstances (Dickter & Bartholow, 2007).

However, also consistent with behavioral work regarding implicit attitudes, it is not always the case that stigmatized or outgroup faces elicit the same degree of bias, or corresponding amygdala responses (Cunningham, et al., 2004a; Lieberman, et al., 2005). Thus, in the next section, we turn our attention to the neuroimaging literature that addresses when and how individuals are likely to regulate automatic, biased responses.

Regulation of Race Bias. Given that regions associated with arousal and negative affect (e.g. amygdala, insula) are sometimes, but not always observed in response to outgroup faces, researchers have asked questions such as: Is fear/arousal always an automatic response to outgroup faces? How do individuals regulate automatic, and potentially biased responding? Is bias always a first response that then needs to be controlled? Under what circumstances do automatic biases predominate, and under what circumstances should we observe more controlled processing? How are different regulation strategies related to different behavioral outcomes?

In considering the question of whether outgroup faces always elicit amygdala activity (or put another way, whether we spontaneously regulate this response under some circumstances), Phelps and colleagues showed that while unfamiliar black faces elicited more amygdala activity than unfamiliar white faces, this effect disappeared when both white and black faces were positively perceived, familiar faces (Phelps, et al., 2000). Similarly, Wheeler and Fiske (2005) observed that while white participants showed increased amygdala activity in response to black versus white faces when asked to categorize the race of the person presented (race salient condition), this effect disappeared when participants were asked to personalize the individuals depicted by guessing information about the target, such as whether the he or she liked various vegetables (Wheeler & Fiske, 2005). Likewise, in a recent study in which participants made superficial ratings (regarding age) or personal ratings (regarding food preferences) of stigmatized outgroup members, when making superficial judgments, increased activity was observed in regions associated with negative affective states (e.g. insula), while increased activity in social-cognitive/self-processing regions (e.g. MPFC) was associated with making individuating, personal ratings (Harris & Fiske, 2007). Results of this kind support the idea that familiarity and personalization of outgroup targets may reduce the automatic tendency toward bias.

Conversely, Hart and colleagues (2000) observed that amygdala activity increased in response to outgroup faces, but only after time had passed (Hart, et al., 2000). This finding suggested the opposite of a familiarity effect. Thus, Hart and colleagues concluded that the amygdala response to an unfamiliar face habituates to

same race faces but not to outgroup faces, challenging Phelps and colleagues' (2000) conclusion that amygdala responses are modulated by familiarity.

A study by Cunningham and colleagues (2004) may help resolve the seeming discrepancy described above (Cunningham, et al., 2004a). In this study, white participants viewed photos of abstract pictures and of Black and White human faces for short (30 ms) or longer (525ms) time periods while in an fMRI scanner. The shorter duration stimuli were not accessible to conscious awareness (participants did not report seeing them). However, when comparing brain activity as participants viewed black faces (compared to white faces), participants showed increased amygdala activity in response to black faces when the stimuli were presented outside of conscious awareness. However when the participants had the opportunity to consciously process the stimuli (when the face was on the screen for 525 ms) the difference in the amount of amygdala activity to black versus white faces was reduced, and activity in areas of controlled processing (right VLPFC, right DLPFC, ACC) increased. Furthermore, activity in controlled processing regions such as the DLPFC and ACC was correlated with change in amygdala activity, suggesting that these areas may be recruited to downregulate the initial amygdala response. Therefore, the authors suggest that when viewing members of an outgroup, our initial response tends to be automatic and affective, but that this response is soon regulated by more controlled processing in the PFC and ACC (Cunningham, et al., 2004a). The authors also suggest that the impact of controlled processing in regulating the automatic response may be somewhat volitional. Given that all participants reported low levels of prejudice on an explicit measure, it is likely that

participants are motivated to present themselves as non-prejudiced, and quickly regulate the initial automatic, affective response. This pattern of results has also been observed in response to other stigmatized outgroups (e.g. obese, transsexual, unattractive and facially pierced individuals), with increases in negative affective processing regions such as the amygdala and insula prompting greater responses in regulatory regions such as the anterior cingulate, prefrontal cortex (Krendl, Macrae, Kelley, Fugelsang, & Heatherton, 2006).

Other deliberate factors can influence the use of controlled processing to regulate bias as well. For example, Lieberman and colleagues (2005) required participants to either match images on the basis of race or to label the race of faces presented (Lieberman, et al., 2005). The authors reasoned that the top-down nature of the verbal labeling task would require more controlled processing, and indeed in this study, the verbal labeling task showed increased activity in right VLPFC. Furthermore, although the authors observed increased amygdala activity when participants visually matched photos of people according to race, the effect disappeared when participants were required to verbally label the images as belonging to a given race, and the amount of increased activity in RVL PFC correlated with decreases in amygdala activity.³ The authors conclude, therefore, that while automatic responses are likely when individuals are confronted with images, the process of labeling these evaluatively-laden stimuli has a top-down regulatory effect (Lieberman, et al., 2005). Follow up work addressing methodological limitations in previous affect labeling studies also supports the regulatory role RVL PFC in diminishing amygdala activity, via a pathway through medial prefrontal cortex

(Lieberman, Eisenberger, Crockett, Tom, Pfeifer, & Way, 2007). This further supports the idea that changes in the way that an individual processes information (about an outgroup target, for example) can influence the ultimate type of evaluative process engaged.

Consequences of the need to regulate. The strength of the relationship between automatic neural responses in the amygdala and their regulation by prefrontal networks prompted Richeson and colleagues (2003) to hypothesize that for people who have a strong, automatic tendency towards implicitly biased attitudes, regulation might become more difficult under conditions of cognitive load or when controlled processing resources are otherwise depleted. Thus, even for people who hold explicitly unbiased attitudes, and who are likely to be motivated to regulate that bias (due to societal norms, or for other reasons), Richeson and colleagues hypothesized that following a demanding cognitive task these participants might show increased bias. Likewise, following an interaction with an outgroup individual, participants might show evidence of depleted cognitive resources. Indeed, in a series of behavioral and fMRI studies, this is exactly what they found (Richeson, Baird, Gordon, Heatherton, Wyland, Trawalter, & Shelton, 2003; Shelton, Richeson, Salvatore, & Trawalter, 2005).

Furthermore, in the fMRI investigation, the extent of controlled processing (as indexed by activity in DLPFC) engaged by the presentation of black faces was correlated with implicit racial bias. This suggests that for individuals who held greater implicit bias, more prefrontal resources were recruited when confronted

with a situation that warranted regulation. Furthermore, the amount of activity in prefrontal cortex engaged by presentation of black faces mediated the relationship between implicit bias and interference on the Stroop color naming task following interaction with a black person. These results provide additional support for the idea that activity in regions that are typically associated with controlled processing can regulate automatic race bias, however, prolonged need to regulate in one area may spillover to produce decreased regulatory ability in other tasks (Richeson, et al., 2003).

To summarize, longer presentation times and tasks requiring cognitive control are less likely to allow PFC availability, while tasks that allow a medium length of presentation and require less cognitive control are more likely to find effects (Knutson, Wood, Spampinato, & Grafman, 2006). Thus, while early studies examining amygdala responses to black faces produced conflicting results, we now recognize the importance of understanding the time course of the stimulus presented, and the sensitivity of the experimental design to pick up on changes in key brain regions over time. We now also know that greater demands on controlled processing resources may deplete the ability for participants to regulate bias. The results described above suggest that the amount of implicit bias observed is likely a function of interplay between the strength of automatic responses (indexed by activity in affective processing regions such as the amygdala), and the strength of controlled processing (indexed primarily by activity in networks involved in controlled processing in prefrontal cortex). Processing of outgroups and other stigmatized categories is influenced both by the timecourse of the stimulus and

response measured, by the demands of the task at hand (Cunningham, et al., 2007), and by factors such as prior contact with the outgroup (Walker, Silvert, Hewstone, & Nobre, 2008). Furthermore, specific regulation strategies can be employed to reduce implicit biases that would otherwise be present (Harris & Fiske, 2007; Krendl, et al., 2006; Lieberman, et al., 2005; Wheeler & Fiske, 2005).

Stereotypes, bias and non-racial outgroups

While race is one of the most salient characteristics that distinguish people in groups, and while a large corpus of work relating neurocognitive activity to bias and the regulation of bias has focused on race (Lieberman, 2007), other work has explored the extent to which these same processes apply in other intergroup situations (e.g. political outgroups, gender outgroups, etc.).

One particular area of interest has been neural responses to political outgroups. Consistent with the literature on responses to racial outgroup faces, several researchers have examined the interplay between automatic, affective responses and more controlled, deliberate responses to partisan outgroup faces. For example, Knutson and colleagues (2006) observed that at a neural level, the activation of political attitudes (operationalized as performing an IAT using images of politicians who belonged to the participant's ingroup and outgroup) produced neural activations in both emotion processing regions (VMPFC), and more controlled processing regions of prefrontal cortex. However, participants who reported stronger party affiliation on an explicit measure of political preferences, showed less activity in controlled processing regions (lateral PFC) while completing

the IAT. These findings support the idea that political attitudes may be processed along stereotypic, or symbolic lines in cases when individuals hold stronger prior attitudes (Knutson, et al., 2006).⁴

However, as in the literature describing race bias, explicit beliefs did not exclusively predict responses; some types of information processing are more likely to lead to automatic, affective responses than others. For example, consistent with the work of Lieberman and colleagues (2005), Knutson and colleagues (2006) observed increased amygdala activity when participants viewed images of outgroup politicians' faces during the IAT, but not when responding to written names.

A study by Kaplan and colleagues (2007) also showed activity in both automatic and controlled processing regions in response to the presentation of outgroup political faces, but these authors came to different conclusions regarding the meaning of the activated networks. Activity in both affective processing regions (e.g. insula), as well as more control oriented regions of prefrontal cortex were observed when viewing the faces of political candidates from an opposing political party as compared to viewing faces of the participant's own political candidate. However, unlike previous work on race bias, in which prefrontal areas are interpreted as being engaged to down-regulate negative affective responses, the authors of this study suggest that the presentation of political outgroup faces actually engages controlled processing networks that *upregulate* negative affective responses (Kaplan, et al., 2007). Unfortunately, the temporal resolution of the study did not allow causal inference about the direction of the effect, but it is interesting to

consider this interpretation in light of the differing motivations inherent to the stimuli; people are motivated to appear less racist, but it may be more societally acceptable to be a strong partisan.

In line with this hypothesis, research on motivated reasoning suggests that when reasoning about counter attitudinal political information, increased activity is observed in regions associated with affective evaluation (VMPFC, amygdala, insula), but not traditional cognitive control areas such as DLPFC and ACC (Westen, Blagov, Harenski, Kilts, & Hamann, 2006).⁵ By contrast, in a sample of low prejudice individuals, activity in neural regions linked to conflict detection (dACC) predicted stereotype inhibition in both private and public settings, while activity in neural regions associated with error-perception (rACC) predicted behavioral control of bias for individuals who reported high sensitivity to societal non-prejudice norms in public settings (Amodio, et al., 2006).

Thus, consistent with the literature on evaluation, preferences, and reactions to racial outgroups, when responding to non-racial outgroups, individuals tend to have automatic and often affective responses, represented by increased activity in areas such as the VMPFC, amygdala and insula. Also consistent with the race literature, in many cases, increases in affective processing regions is accompanied by corresponding increases in regulatory regions of lateral prefrontal cortex. Unlike the responses to racial outgroups, however, there may be less societal pressure to appear unbiased toward non-racial outgroups (e.g. partisan groups), and hence different regulatory strategies may be employed.

Neural Bases of Persuasion and Attitude Change

Having considered the ways that the brain supports our evaluations of objects, concepts, brands, people and groups, it makes sense to consider the ways in which these moment to moment evaluations, represented through brain activity, influence the ultimate structure of our attitudes and behaviors, and the ways that outside influences may shape these evaluations. Here we briefly explore both an example of internally driven attitude change (cognitive dissonance), as well as an example of externally driven attitude change (persuasion), and suggest areas that may benefit especially from future social neuroscience research.

Dissonance based attitude change. The earliest work in social neuroscience to explore attitude change was conducted by Lieberman and colleagues (Lieberman, Ochsner, Gilbert, & Schacter, 2001). This work explored the phenomenon of cognitive dissonance, in which conflicting initial attitudes and behaviors presumably produce discomfort that leads to subsequent attitude change (Festinger, 1957). This team explored the effects of cognitive dissonance-induced attitude change in both anterograde amnesia patients and healthy controls. The researchers observed that while the amnesia patients had no memory of performing a behavior that conflicted with their prior attitudes, as with healthy controls, the patients did change their attitudes to be more in line with the performed behavior. Thus, in contrast to previous explanations of cognitive dissonance effects involving conscious rationalization, the researchers suggested that even when individuals have no memory of inconsistent prior attitudes and behaviors, implicit processes are likely at work that still result in attitude change.

Subsequent imaging work in this area has suggested that increased activity in controlled processing regions (right inferior frontal gyrus, MPFC), as well as positive affective areas (nucleus accumbens) is correlated with increased post-decisional, dissonance induced attitude change. Conversely, activity in areas associated with unpleasant emotions, such as the insula is inversely correlated with attitude change (the more attitudes change following a forced choice paradigm, the less activity is observed in the insula) (Jarcho, Berkman, & Lieberman, in preparation). Thus, we can think of dissonance based attitude change within a similar framework to other types of attitude (or bias) regulation discussed above, wherein automatic responses in affective processing regions are regulated by higher-order processes in prefrontal cortex.

The Neural Bases of Persuasion. In considering phenomena such as evaluations of outgroup faces, partisan group symbols, and post-decisional attitude change, we have largely ignored the potential influence of outside sources intended to shape or change people's attitudes (that is, all of the research discussed to this point has been primarily internally driven). However, in reality, many factors including societal norms, group norms and explicit persuasive appeals influences individuals' attitudes and behaviors.

Work by Falk and colleagues and by Klucharev and colleagues has taken preliminary steps towards uncovering the neural bases of the experience of persuasion by an external source (Falk, Berkman, Harrison, Mann, & Lieberman, in

preparation; Falk & Lieberman, in preparation; Falk, Rameson, Kang, Liao, Inagaki, & Lieberman, under review; Klucharev, Smidts, & Fernandez, 2008).

Falk and colleagues demonstrated that across two diverse cultural/linguistic groups (Americans and Koreans) and using two different types of media (plain text and video based messages), activity in the dorsomedial prefrontal cortex (DMPFC), bilateral posterior superior temporal sulcus (pSTS), and bilateral temporal poles (TP) is associated with the experience of persuasion (Falk, et al., under review). These researchers have also shown that in some situations, medial temporal lobes, left VLPFC, VMPFC and visual cortex are correlated with the experience of persuasion. Likewise, Klucharev and colleagues observed that expert power in presenting arguments resulted in increased activity in left prefrontal and parietal cortices, as well as the medial temporal lobes, which they attribute to increased semantic processing and memory encoding when information comes from an expert source (Klucharev, et al., 2008).

Interestingly, the constellation of regions observed most consistently in the persuasion studies conducted by Falk and colleagues (DMPFC, pSTS, TP), has previously been observed in response to tasks related to theory of mind processing (Frith & Frith, 2003), and not in response to other types of tasks (Cabeza & Nyberg, 2000). The notion that social cognition and persuasion rely on similar neural mechanisms is consistent with Emerson's proposal that the goal of persuasion "is to bring another out of his bad sense into your good sense" (Emerson, 1880). Prior behavioral research has also touched on the relationship between perspective

taking and persuasion (Aaker & Williams, 1998; Campbell & Babrow, 2004; Eagly & Chaiken, 1993; Eagly & Chaiken, 2005; Jarcho, et al., in preparation), but there has been less direct emphasis on social cognition as a key factor in persuasion research; this may be a fruitful area of exploration and dialogue between neuroscience and social psychology, in addition to existing knowledge about the role of cognition and affect in the process of persuasion.

The current results also contribute to our understanding of the “theory of mind network.” Given the results described, it is possible that theory of mind processes may be a special case of considering beliefs, desires, and intentions more generally, whether tied to another’s mind or merely indicated within an argument without being linked to a particular mind.⁶

Attitude Change. While the work described above is the first to explore externally driven evaluations in the brain, it is also interesting to consider the interplay between external forces (such as persuasive arguments) and internal forces that actually result in attitude change.

In order to explore attitude change following externally provided information, Falk and Lieberman explore both objective attitude change (calculated as the difference between attitudes at two time points) and subjective attitude change (measured following the attitude change manipulation) using a paradigm in which an attitude is induced about a fictional animal and then with subsequent information the attitude is changed (Crites, Fabrigar, & Petty, 1994; Fabrigar & Petty, 1999; Falk & Lieberman, in preparation). Falk and Lieberman report that

SCN LAB 10/19/08 6:23 PM

Comment: Not sure who will be on this paper... will add people.

increased objective attitude change in this paradigm (calculated as final attitude minus initial attitude) is positively associated with neurocognitive networks traditionally observed in social cognitive and mentalizing processes (DMPFC, TPJ) and self-referential processing (MPFC, posterior cingulate). Similarly, but not identically, subjective attitude change (measured as amount participants believed their attitudes had changed) was correlated with increased activity in networks traditionally observed in social cognitive and mentalizing processes (DMPFC, pSTS, TPJ, TP), and self-referential processing (MPFC, posterior cingulate), as well as networks previously implicated in memory (medial temporal lobe (MTL)/hippocampus), reward processing (ventral striatum), and emotion regulation (VLPFC) (Falk & Lieberman, in preparation).

Consistent with the role of social information and perspective taking in attitude change and decision making (Aaker & Williams, 1998; Campbell & Babrow, 2004; Eagly & Chaiken, 1993; Eagly & Chaiken, 2005; Salganik, Dodds, & Watts, 2006), the reported findings suggest that both objective and subjective attitude change are supported by neurocognitive networks involved in social cognitive processing (DMPFC, pSTS, TPJ), and overlap with the regions associated with the experience of persuasion described above (Falk, et al., under review; Klucharev, et al., 2008). Given that much information that forms the basis of our attitudes and judgments comes from other people, adopting the perspective of the message communicator would be key to both taking on the initial attitude presented and to changing that attitude upon receipt of new information. Also consistent with prior work suggesting the importance of self-related cognitions in the formation and

change of attitudes (Bem, 1967; Burger, Messian, Patel, del Prado, & Anderson, 2004; Meyers-Levy & Peracchio, 1995), as well as with prior work on dissonance-induced attitude change (Jarcho, et al., in preparation), calculated attitude change was positively related to activity in self-processing areas (MPFC, precuneus, posterior cingulate).

Behavior Change. A last area of interest in considering how the brain supports evaluation, preference and attitude change is the extent to which the regions of the brain that are associated with making evaluations and changing attitudes actually map on to the areas of the brain that predict relevant behavior changes as well. Preliminary evidence suggests that there is overlap in the neural systems that support persuasion, attitude change, intention change and behavior change (Falk, et al., in preparation).

Building on behavioral work by Mann and colleagues (2004), Falk and colleagues (in preparation) had participants complete measures of prior sunscreen use and flossing behavior, as well as attitudes towards the objects and intentions to perform the associated behaviors (floss, use sunscreen) over a one week period (Falk, et al., in preparation; Mann, Sherman, & Updegraff, 2004). Following completion of these prior attitude and prior behavior measures, participants underwent fMRI brain scans while viewing persuasive messages related to each behavior. Following the scanning session, participants again completed attitude and intention measures, and were given floss and sunscreen as a "gift" for participating in the experiment. The researchers followed up with each participant one week

later to find out how much of the floss and sunscreen were used and how often participants had completed the behaviors during the week following the scan.

In examining areas of the brain that were more active while viewing persuasive messages in individuals who changed their behavioral intentions, increased activity was observed in parts of the medial parietal network (previously observed in the attitude change studies described above), as well as in the posterior superior temporal sulcus. Likewise, in individuals who changed their behavior following the experiment, increased activity was observed in these same regions, compared to individuals who did not change their behavior (Falk, et al., in preparation). Thus, it seems that while not exclusively parallel, there is some degree of overlap in the neural systems that support our evaluations of persuasive messages, as well as the subsequent change of attitudes, intentions, and potentially behaviors. Future social neuroscience work is needed to determine the extent to which these results generalize to other types of attitudes and behaviors.

Summary and Future Directions

Over the past decade, our understanding of the neural systems that support evaluation, preferences, attitudes and persuasion has grown into a base that will support ongoing investigations. Future investigations will continue to explore the ways that the brain generates evaluations of the social environment, makes judgments, forms preferences, and acts upon these attitudes and preferences under various circumstances.

While the research reviewed above suggests that there is some degree of convergence in the neural systems that support evaluations of issues and people, as well as in response to racial and non-racial outgroups, behavioral research has clearly demonstrated that the dynamic process of evaluation and attitude change differs depending on factors such as the initial strength of attitudes, and factors related to the cause of potential attitude change (Eagly & Chaiken, 1993; Petty & Cacioppo, 1986). Future work is needed to explore moderators and boundary condition of the neural bases of each attitudinal process described in this chapter. Research is also needed to explore the relationship between the neurocognitive predictors of attitudes, intentions and behaviors under different circumstances. For example, while initial steps have been taken to explore neural predictors of behavior change following a persuasive message, it will be interesting to understand how this pattern differs depending on the behavior in question, prior attitudes related to that behavior, and factors related to the message and message delivery. Given the relative novelty of social neuroscience as a tool to understand the processes underlying evaluation and attitudes, there are clearly several areas that merit further investigation, which are beyond the scope of this chapter.⁷

Lastly, as a final caveat, while neuroimaging allows examination of multiple processes in concert, and may allow us to link our understanding of human psychological processes to a vast neuroscience literature in animal models, it is also subject to inherent weaknesses. For example, the scanner environment is likely to reduce our ability to simulate real life situations, and may also introduce confounds related to the novelty of the situation and/ or the conditions under which

information is delivered. Thus, just as behavioral research informs the questions that are asked in fMRI, it will be useful to consider novel hypotheses generated by the work that can be tested outside of the scanner in a more naturalistic environment. By employing an iterative process in which behavioral and neuroimaging research continue to inform one another, both disciplines will benefit.

References

- Aaker, J. L., & Williams, P. (1998). Empathy versus pride: The influence of emotional appeals across cultures. *Journal of Consumer Research*, 25(3), 241-261.
- Albarracín, D., Johnson, B. T., & Zanna, M. P. (Eds.). (2005). *The Handbook of Attitudes*. Mahwah, NJ: Lawrence Erlbaum.
- Allport, G. W. (1935). Attitudes. In C. M. Murchison (Ed.), *Handbook of Social Psychology*. Winchester, MA: Clark University Press.
- Allport, G. W. (1954). *The nature of prejudice*: Reading, MA: Addison-Wesley.
- Amodio, D., & Lieberman, M. (in press). Pictures in our heads: Contributions of fMRI to the study of prejudice and stereotyping. In T. Nelson (Ed.), *Handbook of Prejudice, Stereotyping, and Discrimination*: Earlbaum Press.
- Amodio, D. M., Harmon-Jones, E., Devine, P. G., & Curtin, J. J. (2004). Neural Signals for the Detection of Unintentional Race Bias. *Psychological Science*.
- Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (2006). Alternative mechanisms for regulating racial responses according to internal vs external cues. *Soc Cogn Affect Neurosci*, 1(1), 26-36.
- Anders, S., Eippert, F., Weiskopf, N., & Veit, R. (2008). The human amygdala is sensitive to the valence of pictures and sounds irrespective of arousal: an fMRI study. *Soc Cogn Affect Neurosci*, 3(3), 233-243.
- Aristotle (1924/1954). *The Rhetoric of Aristotle*. , *Aristotle, with an English translation: the "Art" of rhetoric, by John Henry Freese*. Oxford: Clarendon Press.

- Ashburn-Nardo, L., Knowles, M. L., & Monteith, M. J. (2003). Black americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition, 21*(1), 61-87.
- Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *J Pers Soc Psychol, 62*(6), 893-912.
- Bem, D. J. (1967). Self-Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena. *Psychological Review, 74*(3), 183-200.
- Bishop, G. (1980). Pseudo-opinions on public affairs. *Public Opinion Quarterly, 44*(2), 198.
- Bishop, G. D., Hamilton, D. L., & McConahay, J. B. (1980). Attitudes and nonattitudes in the belief systems of mass publics. *The Journal of Social Psychology, 110*, 53-64.
- Bonilla-Silva, E., & Glover, K. (2004). "We are all Americans": The Latin Americanization of Race Relations in the United States. In M. Krysan & A. Lewis (Eds.), *The Changing Terrain of Race and Ethnicity*. New York: Russel Sage Foundation.
- Burger, J. M., Messian, N., Patel, S., del Prado, A., & Anderson, C. (2004). What a coincidence! The effects of incidental similarity on compliance. *Personality and Social Psychology Bulletin, 30*(1), 35-43.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J Cogn Neurosci, 12*(1), 1-47.

- Campbell, R. G., & Babrow, A. S. (2004). The role of empathy in responses to persuasive risk communication: overcoming resistance to HIV prevention messages. *Health Commun, 16*(2), 159-182.
- Chen, M., & Bargh, J. A. (1997). Nonconscious Behavioral Confirmation Processes: The Self-Fulfilling Consequences of Automatic Stereotype Activation. *Journal of Experimental Social Psychology, 33*(5), 541-560.
- Converse, P. E. (1970). Attitudes and non-attitudes: continuation of a dialogue. In E. R. Tuft (Ed.), *The quantitative analysis of social problems*. Reading, MA: Addison Wesley.
- Crites, S. L., Fabrigar, L. R., & Petty, R. E. (1994). Measuring the affective and cognitive properties of attitudes: Conceptual and methodological issues. *Personality and Social Psychology Bulletin, 20*(6), 619-634.
- Cunningham, W., Johnson, M., Gatenby, J., Gore, J., & Banaji, M. (2003). Neural components of social evaluation. *J. Personal. Soc. Psychol., 85*, 639-649.
- Cunningham, W., Johnson, M., Raye, C., Chris Gatenby, J., Gore, J., & Banaji, M. (2004a). Separable Neural Components in the Processing of Black and White Faces. *Psychological Science, 15*(12), 806-813.
- Cunningham, W., Raye, C. L., & Johnson, M. K. (2004b). Implicit and Explicit Evaluation: fMRI Correlates of Valence, Emotional Intensity, and Control in *Journal of Cognitive Neuroscience*.
- Cunningham, W., & Zelazo, P. (2007). Attitudes and evaluations: a social cognitive neuroscience perspective. *Trends Cogn Sci, 11*(3), 97-104.

- Cunningham, W., Zelazo, P., Packer, D. J., & Van Bavel, J. J. (2007). The Iterative Reprocessing Model: A Multilevel Framework for Attitudes and Evaluation. *Social Cognition*.
- Davis, M. (1992). The Role of the Amygdala in Fear and Anxiety. *Annual Reviews in Neuroscience*, 15(1), 353-375.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), 5-18.
- Dickter, C. L., & Bartholow, B. D. (2007). Racial ingroup and outgroup attention biases revealed by event-related brain potentials. *Soc Cogn Affect Neurosci*, 2(3), 189-198.
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *J Pers Soc Psychol*, 82(1), 62-68.
- Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General*, 127(3), 286-303.
- Eagly, A., & Chaiken, S. (2007). The Advantages of an Inclusive Definition of Attitude. *Social Cognition*, 25(5), 582-602.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*: Harcourt Brace Jovanovich College Publishers.
- Eagly, A. H., & Chaiken, S. (2005). *Attitude research in the 21st century: The current state of knowledge*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Eberhardt, J. (2005). Imaging Race. *American Psychologist*, 60(2), 181-190.

- Emerson, R. W. (1880). *Letters and Social Aims* (Vol. IV). Cambridge MA: Riverside Press.
- Fabrigar, L., & Petty, R. (1999). The Role of the Affective and Cognitive Bases of Attitudes in Susceptibility to Affectively and Cognitively Based Persuasion. *Personality and Social Psychology Bulletin*, 25(3), 363-381.
- Falk, E., Berkman, E., Harrison, B., Mann, T., & Lieberman, M. D. (in preparation). Neural Correlates of Health Behavior Change.
- Falk, E., & Lieberman, M. (in preparation). The Neural Bases of Attitude Change. UCLA.
- Falk, E., Rameson, L., Kang, Y., Liao, B., Inagaki, T., & Lieberman, M. D. (under review). The Neural Correlates of Persuasion.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *J Pers Soc Psychol*, 69(6), 1013-1027.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229•238-229•238.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson & Company.
- Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philos Trans R Soc Lond B Biol Sci*, 358(1431), 459-473.
- Gawronski, B. (2007). Editorial: Attitudes can be Measured! But What is an Attitude? *Social Cognition*, 25(5), 573-581.

- Greenwald, A. G., & Banaji, M. R. (1995). Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes. *Psychological Review*, *102*(1), 4-27.
- Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Soc Cogn Affect Neurosci*, *2*(1), 45-51.
- Hart, A., Whalen, P., Shin, L., McInerney, S., Fischer, H., & Rauch, S. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *Neuroreport*, *11*(11), 2351.
- Herbert, C., Ethofer, T., Anders, S., Junghofer, M., Wildgruber, D., Grodd, W., et al. (2008). Amygdala activation during reading of emotional adjectives--an advantage for pleasant content. *Soc Cogn Affect Neurosci*, nsn027.
- Hovland, C. I. (1949). Reconciling conflicting results derived from experimental and survey studies of attitude change. *American Psychologist*, *14*, 8-17.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion: Psychological studies of opinion change*. New Haven, CT: Yale University Press.
- Jacobsen, T., Schubotz, R., Hofel, L., & Cramon, D. (2005). Brain correlates of aesthetic judgment of beauty. *NeuroImage*, *10*.
- Jarcho, J., Berkman, E., & Lieberman, M. D. (in preparation). Neural Correlates of Post-Decisional Attitude Change. UCLA.
- Kaplan, J., Freedman, J., & Iacoboni, M. (2007). Us versus them: Political attitudes and party affiliation influence neural response to faces of presidential candidates. *Neuropsychologia*, *45*(1), 55-64.

- Klucharev, V., Smidts, A., & Fernandez, G. (2008). Brain mechanisms of persuasion: how 'expert power' modulates memory and attitudes. *Soc Cogn Affect Neurosci*, nsn022.
- Knutson, K., Wood, J., Spampinato, M., & Grafman, J. (2006). Politics on the brain: An fMRI investigation. *PSNS*, 1(1), 25-40.
- Koenigs, M., & Tranel, D. (2008). Prefrontal cortex damage abolishes brand-cued changes in cola preference. *Soc Cogn Affect Neurosci*, 3(1), 1-6.
- Krendl, A. C., Macrae, C. N., Kelley, W. M., Fugelsang, J. A., & Heatherton, T. F. (2006). The good, the bad, and the ugly: An fMRI investigation of the functional anatomic correlates of stigma. *Social Neuroscience*, 1(1), 5 - 15.
- Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). The Measurement of Attitudes. In D. Albarracin, B. T. Johnson & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. xii, 826). Mahwah, N.J.: Lawrence Erlbaum Associates Publishers.
- Lieberman, M. (2007). Social cognitive neuroscience: a review of core processes. *Annual Review of Psychology*, 58(18.11).
- Lieberman, M., Hariri, A., Jarcho, J., Eisenberger, N., & Bookheimer, S. (2005). An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nat Neurosci*, 8(6), 720-722.
- Lieberman, M., Ochsner, K., Gilbert, D., & Schacter, D. (2001). Do Amnesics Exhibit Cognitive Dissonance Reduction?: The Role of Explicit Memory and Attention in Attitude Change. *Psychological Science*, 12(2), 135-140.

- Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting feelings into words: affect labeling disrupts amygdala activity in response to affective stimuli. *Psychol Sci, 18*(5), 421-428.
- Livingston, R. W. (2002). The role of perceived negativity in the moderation of African Americans' implicit and explicit racial attitudes. *Journal of Experimental Social Psychology, 38*, 405-413.
- Mann, T., Sherman, D., & Updegraff, J. (2004). Dispositional Motivations and Message Framing: A Test of the Congruency Hypothesis in College Students. *Health Psychology, 23*(3), 330-334.
- McClure, S. (2004). Neural Correlates of Behavioral Preference for Culturally Familiar Drinks. *Neuron, 44*(2), 379-387.
- Meyers-Levy, J., & Peracchio, L. A. (1995). Moderators of the impact of self-reference on persuasion. *Journal of Consumer Research, 22*(4), 408-423.
- Milne, E., & Grafman, J. (2001). Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *J Neurosci, 21*(12), RC150.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice, 6*(1), 101-115.
- Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist, 56*(9), 717-734.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

- Petty, R. E., Priester, J. R., & Wegener, D. T. (1994). Cognitive processes in attitude change. In R. S. J. Wyer (Ed.), *Handbook of social cognition* (Vol. 1, pp. 69-142). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Phelps, E. (2001). Faces and races in the brain. *Nat Neurosci*.
- Phelps, E., Cannistraci, C. J., & Cunningham, W. (2003). Intact performance on an indirect measure of race bias following amygdala damage. *Neuropsychologia*.
- Phelps, E., O'Connor, K. J., Cunningham, W., & Funayama, E. S. (2000). Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation. *J Cogn Neurosci*.
- Phelps, E., & Thomas, L. (2003). Race, Behavior, and the Brain: The Role of Neuroimaging in Understanding Complex Social Behaviors. *Political Psychology*, 24(4), 747-758.
- Richeson, J., Baird, A., Gordon, H., Heatherton, T., Wyland, C., Trawalter, S., et al. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nat Neurosci*, 6(12), 1323-1328.
- Ronquillo, J., Denson, T. F., Lickel, B., Lu, Z.-L., Nandy, A., & Maddox, K. B. (2007). The effects of skin tone on race-related amygdala activity: an fMRI investigation. *Soc Cogn Affect Neurosci*, 2(1), 39-44.
- Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, 311(5762), 854-856.
- Sears, D. O., & Savalei, V. (2006). The Political Color Line in America: Many "Peoples of Color" or Black Exceptionalism? *Political Psychology*, 27(6), 895-924.

- Sears, D. O., & Valentino, N. A. (1997). Political Matters: Political Events as Catalysts for Preadult Socialization. *American Political Science Review*, 91, 45-65.
- Shelton, J. N., Richeson, J. A., Salvatore, J., & Trawalter, S. (2005). Ironic effects of racial bias during interracial interactions. *Psychol Sci*, 16(5), 397-402.
- Smith, E. R. (1993). Social identity and social emotions: Toward new conceptualizations of prejudice. *Affect, cognition, and stereotyping: Interactive processes in group perception*, 297—315-297—315.
- Van Bavel, J. J. (2008). *The neural substrates of ingroup bias: An fMRI investigation*. Paper presented at the Social and Affective Neuroscience.
- Vartanian, O., & Goel, V. (2004). Neuroanatomical correlates of aesthetic preference for paintings. *NeuroReport*, 15(5), 893.
- Walker, P. M., Silvert, L., Hewstone, M., & Nobre, A. C. (2008). Social contact and other-race face processing in the human brain. *Soc Cogn Affect Neurosci*, 3(1), 16-25.
- Westen, D., Blagov, P. S., Harenski, K., Kilts, C., & Hamann, S. (2006). Neural bases of motivated reasoning: an FMRI study of emotional constraints on partisan political judgment in the 2004 U.S. Presidential election. *J Cogn Neurosci*, 18(11), 1947-1958.
- Wheeler, M., & Fiske, S. (2005). Social-Cognitive Goals Affect Amygdala and Stereotype Activation. *Psychological Science*, 16(1), 56-63.
- Zajonc, R. B., & Markus, H. (1982). Affective and cognitive factors in preferences. *Journal of Consumer Research*, 9(2), 123-131.

Zysset, S., Huber, O., Ferstl, E., & von Cramon, D. Y. (2002). The Anterior Frontomedian Cortex and Evaluative Judgment: An fMRI Study. *NeuroImage*, 15(4), 983-991.

Notes

¹ For a review of neuroscience as applied to social psychology of race, see (Eberhardt, 2005). Eberhardt discusses social implications of using imaging for these purposes, and reviews literature on the neuroscience of face processing and race, which is beyond the scope of this chapter.

² Researchers have now moved from conceiving of the amygdala as responding exclusively to negative affect and/or fear stimuli, to conceptualizing this structure as responding to emotional intensity of a stimulus, and arousal regardless of valence (Anders, Eippert, Weiskopf, & Veit, 2008; Herbert, Ethofer, Anders, Junghofer, Wildgruber, Grodd, & Kissler, 2008; Van Bavel, 2008).

³ Neuroimaging has also allowed the distinction of concepts such as implicit prejudice from implicit stereotyping. In one study, Potanina, Pfeifer, Lieberman, and Amodio (2006) presented participants with a task that required the participants to either guess whether target Black and White individuals were either athletic (a more cognitive, stereotypic distinction) or whether the person would be a good friend (a more affective, potentially prejudice based distinction). The researchers found that implicit stereotyping was associated with neural activity in controlled processing regions such as the inferior temporal lobe (ITL) and PFC (but not the amygdala), while making affective (potentially prejudice related judgements) was associated with activity in automatic processing regions such as the amygdala (but not ITL or PFC).

⁴ Knutson and colleagues (2006) also lay out several areas that may be common to lower level attitude processing in general. In response to implicit priming of political attitudes, the researchers observed activity in premotor cortex, L inferior frontal gyrus, cingulate gyrus, and fusiform. Given that this constellation of regions appears across several conditions that all involved low level attitude processing, the authors suggest that these regions may be involved in lower-level attitude processes in general.

⁵ See also Sears's work on "symbolic processing" and Marcus' work on "affective intelligence."

⁶ Left VLPFC was the only other region observed in all three studies. In conjunction with MTL in the first two studies reported, it appeared that VLPFC might be contributing to memory encoding processes. However, when persuasiveness judgments were made at the time of viewing the videos in the third study, VLPFC activity was still present but MTL activity was no longer associated with feeling persuaded. Given these results, the authors conclude that left VLPFC may be engaged in reappraisal processes during passages subsequently rated as persuasive, or may be related to the theory of mind network observed, as VLPFC has also been observed in a significant number of theory of mind studies.

⁷ Several other areas also merit further investigation. For example, given that many important attitudes are formed in childhood and adolescence (Sears & Valentino, 1997) it would be interesting to compare the neural systems engaged when children

and adolescents make evaluations and/or express attitudes related to concepts about which they have already formed attitudes, versus about which they are in the process of forming attitudes. Obvious questions in this regard include whether children and adolescents show similar patterns to adults, and if not whether systematic differences can help inform our understanding of how, for example, race bias and political attitudes are formed and influence later attitudes and preferences. Other interesting questions in this line of thinking might include an exploration of when automatic biases are first present at a neural level, and how regulation of automatic evaluations differs between children and adults, or among children showing more or less bias on a given issue.

Lastly, while research suggests that the experiences of Black Americans and other non-white groups differ in important ways (Sears & Savalei, 2006), at present, discrimination exists against many of these groups. Neuroimaging work addressing bias has focused exclusively on black-white comparisons, and has primarily focused on white participants' perceptions of black targets (although there are exceptions, e.g. Lieberman et al., 2005). It would be interesting to explore the neural responses of white participants to other outgroups, e.g. "honorary whites" (Bonilla-Silva & Glover, 2004), and also to continue to explore the similarities and differences in the ways that members of dominant and less dominant groups respond to outgroup faces and scenarios.