Faculty Publications: Political Science

Political Science, Department of

2011

Attitudes

William A. Cunningham
University of Toronto, cunningham@psych.utoronto.ca

Ingrid J. Haas
University of Nebraska-Lincoln, ihaas2@unl.edu

Andrew Jahn
Indiana University,, ajahn@indiana.edu

Attitudes

William A. Cunningham,[1] Ingrid Johnsen Haas,[2] and Andrew Jahn[3]

1. University of Toronto, email cunningham@psych.utoronto.ca
2. University of Nebraska-Lincoln, email ihaas2@unl.edu
3. Indiana University, email ajahn@indiana.edu

Abstract

This chapter reviews social neuroscience research that links social psychological attitudes and evaluative processes to their presumed neural bases. The chapter is organized into four parts. The first section discusses how attitude representations are transformed into evaluative states that can be used to guide thought and action. The next two sections address the related processes of attitude learning and change. The final section discusses applications of these concepts for the study of prejudice and political behavior.

Keywords: attitudes, evaluation, reinforcement, conditioning, amygdala, striatum, orbitofrontal cortex, prefrontal cortex, prejudice

The concepts of attitude and evaluation have been central to social psychology for nearly a century. This should not be surprising, as the ubiquitous act of assigning positive or negative valence is crucial for survival, be it in guiding immediate behavior toward or away from an object, or in anticipation of future rewards or punishments. Through accumulated experience, these evaluative judgments can be consolidated in memory to form a summary attitude, which can be recalled to guide future behavior. These summary attitudes, though imperfect at times, allow for the construction of quick evaluative judgments when similar stimuli or situations occur. For example, once we learn that someone is untrustworthy, we can avoid that person in future situations without needing to re-evaluate all of our previous interactions with the individual. Thus, attitudes enable us to predict the value of objects and the behavior of others, allowing us to adapt to the world through experience and make advantageous decisions.

In this chapter, we review social neuroscience research that links social psychological attitudes and evaluative processes to their presumed neural bases. The chapter is organized into four parts. In the first section, we discuss how attitude representations are transformed into evaluative states that can be used to guide thought and action. In the next two sections, we address the related processes of attitude formation and change. In the last section, we discuss applications of these concepts for the study of prejudice and political behavior.

Attitude Expression

The processes of attitude expression involve the translation of attitudinal representations into an active evaluation that can be used to inform thoughts and behavior (see Cunningham & Zelazo, 2007). Whereas an attitude refers to a relatively stable set of representations (only

some of which may be active at any time), an evaluation reflects the current processing state of the evaluative system (which is determined by the aspects of the attitude that are currently active). Evaluative processes help determine the motivational significance of a stimulus as well as its expected reward or punishment value. In order to do so, these processes draw upon pre-existing attitudes, as well as novel information about the stimulus, contextual information, and current goal states. Evaluative states arise out of dynamic interactions between these elements. Encountered or imagined stimuli (e.g., people, objects, or abstract concepts) elicit relatively automatic evaluations, but these initial "gut reactions" can be modulated by an increasing number of higher-level cognitive and reflective processes (Cunningham & Zelazo, 2007). These higher-order reflective processes send information back to the lower-order processes, allowing for a re-evaluation of the affective response and, if necessary, a different affective interpretation of the same stimuli. This allows for the foregrounding of more relevant and congruent affective representations and the backgrounding of irrelevant or incongruent contextual information in order to achieve a more nuanced evaluation congruent with the current context and/or goals.

When considering the brain regions involved in generating these evaluative predictions, the amygdala has received the most attention. Ever since the classic work of Kluver and Bucy (1937; see also Weiskrantz, 1956) demonstrated that lesions to the temporal lobes led to a decrease in avoidance of potentially threatening stimuli, a common framework for understanding amygdala function has been fear detection and conditioning (see next section). Research using functional magnetic resonance imaging (fMRI) in humans has shown that the amygdala is involved in the detection of threat in many stimulus modalities, including the perception of visual facial expressions of fear (Adolphs et al., 1999; Calder, Keane, Manes, Antoun, & Young, 2000; Morris et al., 1998), cognitive representations of fear (Phelps et al., 2001), threat-related words (Isenberg et al., 1999), and aversive odors (Zald & Pardo, 1997). Given this body of research, Freese and Amaral (2009) have suggested that the amygdala detects danger and then automatically directs behavioral responses.

An examination of the anatomic connections with the amygdala suggests that this region is well suited for automatic vigilance and organized response functions (Davis & Whalen, 2001). Specifically, the amygdala has widespread connections to areas associated with perceptual processing and autonomic/ visceral activation (see Freese & Amaral, 2009, for a review). Thus, following amygdala activation, greater attention can be directed to the stimulus while the body prepares for immediate action. Furthermore, the amygdala has multiple connections to areas of prefrontal cortex (Aggleton, Burton, & Passingham, 1980), receiving from and relaying information to areas of orbitofrontal, insular, and lateral prefrontal cortices (Amaral & Price, 1984; Stefanacci & Amaral, 2000). These connections allow information processed in the amygdala to be used by regions involved in more deliberate forms of decision-making. Through reciprocal connections, amygdala activation can be modulated to take into consideration the entire state of the organism. Thus, following amygdala activation, multiple brain systems are dynamically reorganized (or given the opportunity to reorganize) to appropriately deal with the current environment.

One critical aspect of amygdala function concerns the speed at which it can evaluate the rapid stream of incoming information. Many models of amygdala function suggest that it operates relatively automatically and unconsciously, and current research has provided support for this idea. For example, conscious awareness of a valenced stimulus does not appear to be necessary to produce amygdala activation. In a conceptual replication of previous research on supraliminal face processing (Morris et al., 1996), Whalen and colleagues (1998) demonstrated that subliminal presentations of emotionally fearful faces led to amygdala activation. In addition, Morris, Öhman, and Dolan (1998) found that after participants were classically conditioned to associate particular angry faces with an aversive stimulus, the amygdala showed greater activity to these conditioned faces than to the control faces, using both subliminal and supraliminal presentations. Using depth electrodes, Kawasaki and colleagues (2001) found that the processing of valence (greater neural firing to valenced as opposed to neutral stimuli) occurred 200 milliseconds after stimulus presentation in single-cell recordings of the human amygdala. Taken together, these studies indicate that the human amygdala responds rapidly to valenced stimuli, even when they are presented outside conscious awareness.

However, the suggestion that the amygdala's role in evaluation is valence specific has been called into question. Specifically, several studies have since shown that the amygdala is sensitive not only to fearful or negative information, but also to positive information (Hamann,

Ely, Hoffman, & Kilts, 2002; Hamann & Mao, 2002; Garavan et al., 2001; Liberzon et al., 2003), leading to at least two competing theoretical positions. First, it is possible that the amygdala provides evaluative information about both positive and negative stimuli. According to this view, the amygdala is active whenever generating both positive or negative evaluations. Alternatively, amygdala activation may reflect some process associated with evaluative processing other than valence, such as stimulus intensity or arousal. Hamann and colleagues (2002) replicated the finding that the amygdala responds not only to positive and negative stimuli, but also to unusual or interesting stimuli, suggesting that it serves a more general function than just processing valence. Further, studies that have independently manipulated valence and intensity (Anderson et al., 2003; Small et al., 2003), or used statistical methods to separate the contributions of the two (Cunningham, Raye, & Johnson, 2004), have provided evidence that amygdala activity appears to be associated more with processing affective intensity than with processing any particular valence. Consistent with this idea, patients with bilateral amygdala damage have impaired recognition of emotional arousal, while recognition of valence remains intact (Adolphs, Russell, & Tranel, 1999; see Berntson, Bechara, Damasio, Tranel, & Cacioppo, 2007 for a more nuanced perspective).

One explanation for these findings has been to suggest that amygdala activation may reflect the processing of motivationally relevant stimuli, perhaps recruiting additional resources to facilitate appropriate interactions with the stimulus (e.g., Sander, Grafman, & Zalla, 2003). According to this view, a primary early function of attitudes is to inform us about what is important in any particular situation — then modulate the appropriate second-order perceptual, attentional, autonomic, or cogni- tive/conceptual processes that allow us to deal with the challenges or opportunities that are present. If this is the case, then amygdala activation should vary as a function of the goals of the organism.

To examine the motivational flexibility of and top-down influences on amygdala activation, Cunningham, Raye, and Johnson (2005) presented participants with positively and negatively valenced stimuli during fMRI scanning. After scanning, participants completed an individual differences measure of their prevention— and promotion—focus orientation (i.e., participants indicated whether they were more motivated by negative or positive stimuli, respectively; Higgins, 1997). Consistent with this idea, among participants who were more

promotion focused, greater activation was observed in the amygdala, anterior cingulate gyrus, and extrastriate cortex for positive stimuli. For more prevention-focused participants, greater activation was observed in these same regions for negative stimuli.

In addition, a recent experimental study has provided evidence that situational motives shape amygdala processing in a dynamic fashion (Cunningham, Van Bavel, & Johnsen, 2008). In this study, participants were presented with famous names and asked to focus on either the positive or negative aspects of the person (e.g., ignoring everything bad, how good is this person?). Activity in bilateral amygdala and insula was found to vary as a function of evaluative fit. That is, when focusing on negativity, greater amygdala and insula activity was associated with participants' negativity ratings of the names, but not positivity ratings (recorded after scanning). The opposite pattern was found for the positive-focus condition, such that greater activity was observed in these same regions to ratings of positivity than negativity. Taken together, these studies suggest that chronic and situational motivational concerns can modulate the processing of valenced information to generate situationally appropriate evaluations. These studies demonstrate the power of top-down processes to modulate lower-order processes and provide a new understanding of amygdala function.

Beyond the Amygdala

Although most attention has been directed toward the amygdala, evaluative processes are associated with a much larger circuit involving additional cortical and subcortical regions. Among the more critical subcortical regions associated with evaluation is the ventral striatum, and more specifically the nucleus accumbens (NAcc). Linking NAcc activity to evaluation more closely, studies of economic decision making have shown that NAcc activity is not only correlated with, but is even sometimes a better predictor of, a participant's choice to buy a particular product than is self-report (Knutson et al, 2007). 1 Critically, whereas the amygdala activation appears to be associated with the evaluation of both positive and negative stimuli, the NAcc is primarily involved in the anticipation of positive outcomes and/or receipt of incentives or rewards (Breiter, Aharon, & Kahneman, 2001; Cardinal et al., 2002; Knutson et al., 2001). That is, while the amygdala may not be valence specific, the NAcc may allow for dissociated representations of positive and negative evaluation (see Cacioppo & Berntson,

1994 for a detailed review of evidence of dissociated processing of positive and negative valence).

Whereas activity in the amygdala has been shown to play a role in directing attention towards motivationally significant stimuli and automatically preparing for behavior, little evidence has been found to suggest that this activation leads to the experience of subjective preference. That is, although activation in the amygdala is correlated with objective attitude ratings (Cunningham et al., 2003, 2004), the actual subjective pleasantness associated with receiving (or displeasure associated with not receiving) an expected outcome is correlated with activation in orbitofrontal cortex (OFC; Kringelbach, 2005). Orbitofrontal cortex activity is evident for primary rewards such as food or drink (Kringelbach, O'Doherty, Rolls, & Andrews, 2003; Rolls, 2000), as well as symbolic rewards, such as money (Tom, Fox, Trepel, & Poldrack, 2007; Elliott, Newman, Longe, & William Deakin, 2003; Knutson et al., 2003). Orbitofrontal cortex activity has also been linked to the evaluation of the relative appropriateness of one's responses, activating both to receiving rewards and avoiding punishments (Cunningham, Mowrer, & Kesek, 2009; Kim, Shimojo, & O'Doherty, 2006). Specifically, activity in the medial OFC is typically related to evaluations of positive or rewarding information, whereas activity in the lateral OFC is related to evaluations of negative or punishing information (see Kringelbach & Rolls, 2004 for a review). Thus, while the amygdala and nucleus accumbens may provide information regarding expected outcomes following the perception of a stimulus, the OFC represents the current subjective evaluation.

Because the OFC receives input from multiple sensory modalities, it may play an important role in providing a common metric for representing and comparing disparate aspects of evaluative information (Montague & Berns, 2002; Murray, O'Doherty, & Schoenbaum, 2007; Rolls, 2000; Padoa-Schioppa & Assad, 2006; Wallis & Miller, 2003). Thus, the evaluative connotation of a friendship, a new car, or the ideals of egalitarianism can be reduced to a common evaluative dimension and directly compared across stimulus type. In this sense, whereas relatively more limbic regions may encode and retrieve objective S-R associations, the conversion from attitude to subjective evaluation (which can vary as a function of different contextual factors) may require the OFC. Specifically, activity in medial OFC is typically related to evaluations of positive or rewarding information, whereas activity in lateral OFC is

related to evaluations of negative or punishing information (Anderson et al., 2003; Kringelbach & Rolls, 2004; O'Doherty et al., 2003; but see Northoff et al., 2000).

To the extent that the amygdala and nucleus accumbens provide information regarding expected outcomes following the perception of a stimulus, and the OFC represents the current state of the organism, the dense reciprocal connections between amygdala and OFC allow for a comparison of expected rewards and punishments with current experience. Support for this idea comes from research demonstrating large OFC activations following violations of expectancies (Nobre et al., 1999) and the inability of patients with OFC damage to update representations when predictions and outcomes are incongruent (Rolls et al., 2004; Fellows & Farah, 2003). Thus, whereas subcortical systems provide a low–resolution estimate of likely outcomes, regions of the OFC may be involved in integrating amygdala output with current experience, allowing the current context to play a role in shaping the evaluation (e.g., Blair, 2004; Beer, Heery, Keltner, Scabini, & Knight, 2003; Rolls, 2000; Rolls, Hornak, Wade, & McGrath, 1994).

Constructing More Elaborated Evaluations

In many cases, the evaluation resulting from processing in the amygdala and OFC will be sufficient to produce a behavioral response. In other cases, however, this joint processing may lead to conflict or uncertainty about the stimulus or a predicted outcome. The presence of conflict triggers anterior cingulate cortex activation (see Bush, Luu, & Posner, 2000; Carter et al., 1998), which may then signal the need for additional processing of the stimulus in regions of the lateral prefrontal cortex involved in cognitive control (see Bunge & Zelazo, 2006; MacDonald, Cohen, Stenger, & Carter, 2000; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004). This additional processing in the lateral prefrontal cortex allows for regulation of affect in a top-down fashion by deliberately amplifying or suppressing the processing of certain aspects of the stimulus, changing the input to the system for subsequent processing. Reprocessing will likely modify the current evaluation by modulating activity in lower-order regions (e.g., Cunningham, Johnson et al., 2004; Ochsner, Bunge, Gross, & Gabrieli, 2002; Ochsner et al., 2004).

More complex networks of processing allow for more complex construals of a stimulus. This occurs because more information about a stimulus can be integrated

into the construal and because these networks support the formulation and use of higher-order rules for deliberately selecting certain aspects of a stimulus or context to which to attend (Bunge & Zelazo, 2006; Cunningham & Zelazo, 2007). The selection function of prefrontal cortex may foreground specific aspects of information (and background others). Further, prefrontal cortex may also play a role in keeping current goals and contextual demands/constraints in mind, which is important for fulfilling the competing goals of minimizing error while minimizing processing load (e.g., Cunningham, Zelazo, Packer, & Van Bavel, 2007). This characterization of the prefrontal cortex is consistent with its hypothesized role in allowing for higher levels of reflective consciousness via reprocessing (Zelazo, 2004) and in the monitoring and control of cognition and behavior (e.g., Carver & Scheier, 2001; Shallice, 1982; Stuss & Benson, 1986). Taken together, the dynamic interactions among different brain regions support a flexible and complex process of evaluation that unfolds over time and exists on a continuum from relatively automatic (and simple) to relatively reflective (and complex).

Attitude Formation

The goal of an adaptive learning system is to develop appropriate and accurate estimates of the future value of certain stimuli. As such, attitudes reflect our previous experience with the environment and the learning that accompanies our ongoing positive and negative experiences. To the extent that positive experiences accompany the presence of particular stimuli or following a particular behavior, a more positive attitude will develop (and negative attitudes will develop for negative experiences). The next time that we encounter that same (or similar) situation, our evaluative system can infer that a similar positive (or negative) experience will occur. In general, attitudes that follow direct experience are likely to develop more quickly and result in stronger predictions (Fazio & Zanna, 1981).

The assumptions of reinforcement conditioning can be best summarized by Thorndike's law of effect (1911), which states that an organism will be more likely to repeat actions associated with a reward, and less likely to repeat actions associated with a punishment or negative stimulus. Whereas classical conditioning establishes an evaluative association between one's previous experiences and the attitude object, instrumental conditioning forms associations between hedonic outcomes and the

behavior that produced them. Furthermore, instrumental conditioning involves a fundamentally different set of interactions between the cortical structures and neurotransmitters comprising the instrumental response, which are discussed in detail below.

The most well-known example of both classical and instrumental conditioning is that of fear conditioning (Watson & Rayner, 1920), and among the brain structures believed to be involved in fear conditioning, the amygdala has received the most attention (Davis, 1992; Adolphs, Tranel, Damasio, & Damasio, 1995; LeDoux, Cicchetti, Xagoraris, & Romanski, 1990). For example, during fMRI, the amygdala is active while learning evaluative contingencies (LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998), and patients with amygdala lesions do not show fear conditioning (LaBar, LeDoux, Spencer, & Phelps, 1995). Although much of the learning literature assigns the amygdala a critical role in fear conditioning, recent research has also demonstrated an important association between the amygdala and the processing of reward-related and novel information (see Murray, 2007 for a review). For example, studies of appetitive conditioning on rats with amygdala lesions have shown that the amygdala is necessary for learning some approach behaviors (Cardinal, Parkinson, Lachenal, et al., 2002; Everitt et al., 1999). The human amygdala exhibits a similar pattern of reward processing. Among hungry participants conditioned to associate a visual cue with the odor of a pleasant-smelling food, the amygdala is more active in response to seeing this conditioned visual stimulus than in participants whose hunger has been sated (Gottfried, O'Doherty, & Dolan, 2003). Taken together, these studies suggest that the amygdala plays an active role in conditioned responses regardless of valence (Everitt, Cardinal, Hall, Parkinson, & Robbins, 2000).

Beyond the Amygdala

As in the attitude expression literature, it is a mistake to ascribe too much function to the amygdala. In particular, areas of the striatum and prefrontal cortex also play important roles in reinforcement learning. One prominent theory of striatal contributions to reinforcement learning is that of the actor/critic model (Sutton & Barto, 1998; but see Khamassi et al., 2005), in which the ventral and dorsal striatum serve distinct functions in generating evaluative representations. According to this model, the ventral striatum plays the role of the "critic," generating predictions about the likely reward value of a particular

stimulus. The ventral striatum updates representations to the extent that the prediction and the actual outcome differ (the temporal difference error), which is relayed to the dorsal striatum. This dorsal region (the "actor") then processes this information and makes certain stimulus-response associations more or less likely, depending on the hedonic outcome previously experienced.

To empirically test these roles for dorsal and ventral striatum in humans, O'Doherty and colleagues (2004) hypothesized that these regions would also show differences in activation in response to the amount of control participants had in choosing a particular reward. Two groups were formed, one involving instrumental learning and the other classical learning. In the instrumental learning condition, participants were able to choose which stimuli they preferred based on expected value (i.e., the hedonic value of the reward multiplied by the probability of receiving it), while in the classical learning condition a computer chose the stimuli and the participants had to guess which stimulus was chosen. Ventral striatum activity was observed across both conditions, consistent with its role in appraising predicted value. Dorsal striatum activity, on the other hand, was positively correlated with the prediction error signal during instrumental conditioning. This result supports the hypothesis that this region serves in an "actor" role in exhibiting greater prediction-error-related activity when rewards are chosen during instrumental conditioning, as opposed to classical conditioning.

Although the computations provided by the basal ganglia may appear sufficient for outcome reinforcement learning, current neurobiologically constrained computational models of reinforcement learning and decision-making indicate that the orbitofrontal cortex (OFC) plays a complementary role to the basal ganglia system (see Frank & Claus, 2006). Specifically, the OFC appears to be necessary for representing current reward states (Knutson, Fong, Bennett, Adams, & Hommer, 2003; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001) and updating representations when stimulus contingencies change (Chudasama & Robbins, 2003; Fellows & Farah, 2003). As noted earlier, the OFC may provide a representation of reward magnitude currently being received, which can then be used to flexibly guide behavior and update representations in basal ganglia. That is, the computations of hedonic state from the OFC may provide a powerful input to the comparison between predicted and received outcomes to generate prediction error signals.

While conditioning forms a critical part of the formation of attitudes, there are several instances where attitudes can form in the absence of direct experience. For example, people can learn the evaluative association of a stimulus without any direct experience with the aversive stimulus per se (e.g., Bandura, 1977). People can learn by observing others' responses, or by simply being told that something is good or bad, helpful or harmful. To illustrate this point, Phelps et al. (2001) verbally instructed participants that they would receive at least one shock during the course of an fMRI experiment when a particular stimulus was presented. Although no shocks were ever received during scanning, the authors observed left amygdala activation when the threatening stimulus was presented. The verbally created association between the stimulus and the potentially aversive outcome was therefore sufficient to elicit an amygdala response. Evidence for the amygdala's potential role in observational learning was provided in a follow-up study. Participants that simply watched someone else receiving shocks during a conditioning task had greater amygdala activated to the conditioned stimuli (Olsson, Nearing, & Phelps, 2007). Lastly, the social status of a person has been shown to alter brain activity without direct experience. When playing a trust game, participants has less striatum activity to partners who were portrayed to be trustworthy or untrustworthy when compared to neutral partners (Delgado, Frank, & Phelps, 2005). This suggests that social knowledge may have the ability to supersede conditioned knowledge (and the conditioning process).

Overall, these findings highlight an exciting array of possibilities for learning and attitude formation beyond simple classical conditioning. The ability to mentally empathize and learn from observation produces neural effects similar to those produced through classical conditioning alone. We can dislike groups of people who we have never met, have an aversion to new situations, or have abstract opinions for which we would be willing to die. For such attitudes, conditioning may only play a small part in the development and maintenance of attitude representations.

Attitude Change

Changing circumstances can necessitate the reversal of previously learned associations between stimuli and their hedonic value. This attitude reversal can occur when current evaluations are no longer useful, or when

they turn out to be harmful; for example, a positive evaluation toward a seemingly delicious piece of food must update rapidly if it induces vomiting once ingested. This will result in a more negative evaluation the next time that the food is encountered, suggesting that the attitude has changed.

In exploring how this reversal learning occurs, however, it is useful to make a distinction between ingrained attitudes that are activated more automatically, and more flexible evaluations that can be reversed relatively quickly. This distinction has been categorized as an interaction between the temporary, or short-term, and the permanent, or long-term, aspects of attitude reversal (Frank & Claus, 2006). Both are relative to the number of interactions an organism has had with a certain stimulus or situation. In the case of temporary attitude change, for example, updating and potential reversal of an attitude can occur from situation to situation, while in permanent or long-term attitude change, the reversal occurs over several interactions that are contrary to a previously held attitude.

Although we have focused thus far on the role of the OFC in representing received rewards and helping to provide a reinforcement signal for learning, the OFC is also involved in the updating of evaluative representations either because the representation is no longer appropriate (e.g., extinction) or because the context or situation requires a different response (e.g., Blair, 2004; Beer, Heery, Keltner, Scabini, & Knight, 2003; Rolls, 2000; Rolls, Hornak, Wade, & McGrath, 1994). One way that these processes can be achieved is by providing a complementary predictive reward signal to the basal ganglia that updates more rapidly. Whereas the basal ganglia learning systems slowly incorporate changes in reward contingencies across time and situation, the OFC can update more quickly, using recent experiences to guide an evaluative signal. While the OFC rapidly evaluates and processes sensory information (Kringelbach & Rolls, 2003, 2004), it is argued that this region is also critical in evaluating the associations between environmental stimuli and reinforcement.

Due to these properties, it is possible the OFC plays a key role in reversing or extinguishing previously learned behavior-reward associations (Rolls, 2000). Evidence supporting this hypothesis is illustrated by experiments looking at the effects of learning and reversal in clinical patients with lesions in these prefrontal areas. In one such study by Fellows and Farah (2003), evaluative reversal was examined in patients with OFC lesions.

Patients with either ventromedial OFC (vmOFC) or dorsolateral OFC (dlOFC) lesions performed a simple decision-making task where the reward and loss contingencies of two decks of cards were reversed in the middle of the experiment. While dlOFC lesion patients performed equally to controls on the card game, vmOFC patients performed significantly worse following reversal of the contingencies of the decks. For the worst performers, maximal lesion overlap was observed in the left posteromedial orbitofrontal cortex. These results suggest that attitude reversal, which requires rapid updating, recruits areas of the vmOFC to make these behavioral changes.

Neuroimaging investigations examining these prefrontal regions have corroborated these results in healthy subjects, and have revealed an important distinction between the medial and lateral regions of the OFC in influencing the behavioral response to rewarding and punishing outcomes. In an fMRI investigation by O'Doherty and colleagues (2001), two abstract fractal representations were presented in which participants had to discover through trial and error which fractal was associated with greater gains over multiple trials. Selecting the correct fractal would display a monetary reward on the screen, while an incorrect choice would display the amount lost. The values of the fractals were then switched partway through each trial block, and blood oxygen-level dependent (BOLD) activity was examined during this switch. Immediately after the reversal, medial OFC activity was significantly associated with the BOLD signal following a rewarding stimulus (i.e., the actual receipt of the reward), while lateral OFC activity showed a similar correlation following a punishing stimulus. In both cases, activity in these OFC regions was correlated with the magnitude of the reward or punishment. Based on previous findings, the authors suggested that there are limited and weak connections between the medial and lateral OFC, while within each of these regions there exist several strong and robust neural connections. Therefore, it is useful to categorize the medial and lateral OFC as separate structures contributing to the updating of the relative hedonic values of rewarding and punishing stimuli, respectively.

While the OFC has been shown to play an important role in rapidly updating and reversing attitudes, the basal ganglia, due to its role in forming habits and unconscious motor responses (Jog, Kubota, Connolly, Hillegaart, & Graybiel, 1999; Packard & McGaugh, 1996), is thought to be involved in updating attitudes and

evaluations over the long term, and is strongly modulated by dopamine (DA; Gerfen, 2000). For example, increased DA during positive reinforcement leads to an increased probability for the action that produced the reinforcement (Frank & Claus, 2006; Houk, Adams, & Barto, 1995). The structure of the dopamine pathway itself supports this inference, as it projects to the basal ganglia and medial region of the frontal cortex, experiencing phasic increases in DA when events are better than expected, and phasic decreases in DA when events are worse than expected (Schultz, Dayan, & Montague, 1997). Evidence for the basal ganglia's role in habit formation also comes from lesion studies showing that lesions to a specific part of the striatum affect the learning of habitual responses. For example, lesions to the dorsolateral striatum in rats led to decreased habitual responding, and an increased sensitivity to reward cues when learning stimulus-response associations (Yin, Ostlund, Knowlton, & Belleine, 2005).

Although the basal ganglia and OFC appear to represent two different modes of learning and forming attitudes, we hypothesize that both are influenced by the same information and differ as a function of their learning rates. The basal ganglia is influenced by all positive and negative events providing information about the relative hedonic value of a particular stimulus or behavior, leading to predictions about the reward outcome of a behavior over a relatively long period of time (Jog et al., 1999). When these predictions are violated, as posited by the actor/critic model (O'Doherty et al., 2004), an updating of representations takes place in order to reflect this new information. However, this process takes relatively longer than that of the OFC, which is able to update its representations rapidly on a trial-to-trial basis. Eventually, then, the evaluations stored by the basal ganglia could become the basis for the relatively automatic representations held in working memory by the OFC, leading to these representations becoming the same in both areas, until additional information is encountered and processed by the basal ganglia. This model also possibly explains why, under cognitive load, the representation stored by the basal ganglia takes prominence. The OFC, which contains many of the higher-order cognitive areas required for nuanced and critical thinking, also requires significant cognitive resources to manipulate any attitudes or representations in working memory. Thus, the more automatic and readily accessible attitudes of the basal ganglia will tend to predominate in these cognitively demanding situations.

Yet, not all attitudes can be represented as simple stimulus-evaluation or stimulus-response associations. As noted earlier, evaluations can be highly context dependent and can sometimes lead to conflict, as when both positive and negative characteristics are associated with the attitude object, resulting in a state known as ambivalence (Cacioppo & Berntson, 1994). Thus, although in most situations the OFC can help generate an unambiguous evaluation constrained by a situation, this sometimes fails, and more elaborate "higher-order" attitudes need to be developed to organize attitude representations. That is, by deliberately weighting some information more than others, individuals can form a more integrated evaluation. Unlike inhibition, which drives inconvenient information out of mind, these integrated evaluations yield more complex activations, and may represent and account for inconsistencies. The evaluations that result from this type of processing are similar to what Petty and Cacioppo (1984) call an "elaborated attitude," which is known to be relatively stable and resistant to change. In this sense, these evaluations can be thought of as "resolved ambivalence."

In terms of the underlying brain systems for such a process, the anterior cingulate cortex (ACC) has been shown to play a key role in modulating the activity of both of these areas in attitude reversal. In particular, the ACC is involved in conflict monitoring and evaluating the appropriateness of behavior in specific situations. Largely guided by contextual cues, the ACC assists in modifying behavior to maximize hedonic reward (Bush, Luu, & Posner, 2000; Carter et al., 1998; Holroyd & Coles, 2002). The ACC is thought to signal the need for additional processing in areas of the lateral prefrontal cortex that can reorganize representations in a more abstracted form (Hazy, Frank, & O'Reilly, 2007; Rougier, Noelle, Braver, Cohen, & O'Reilly, 2005). Through repeated re-organization of the same information, a second-level representation can be generated that can be used for subsequent judgments and behavior.

Prejudice

An early focus for the social neuroscience study of attitudes has been the domain of prejudice. Initial studies demonstrated a role for the amygdala in the processing of other-race faces. Hart and colleagues (2000) demonstrated that, for White participants, amygdala

activation to supraliminal Black faces habituated more slowly than White faces; the reverse pattern was found for Black participants. They concluded that all faces are processed immediately for their threat value, but that ingroup faces are deemed safe more quickly than outgroup faces. The role of the amygdala in intergroup perception was further expanded upon by Phelps and colleagues (2000), who showed that greater amygdala activation to Black than White faces was correlated with an indirect measure of race bias that reflects a preference for one race over the other — the Implicit Association Test (IAT; Greenwald et al., 1998). Interestingly, for White participants, neither of these studies showed greater overall amygdala activation to Black faces relative to White faces.

One potential explanation for not finding the expected greater amygdala activation to Black than White faces is that control processes may inhibit or reconstrue an activated emotional response. That is, higher-level cognitive functions may moderate automatically activated attitudes. Thus, for participants viewing long blocks of Black or White faces (as in Hart et al., 2000 and Phelps et al., 2000), there is greater opportunity for control processes to dampen or attenuate any automatic effects that would otherwise be observed. Consistent with this hypothesis, Cunningham and colleagues (2004) found that the majority of White participants had greater amygdala activation to Black than to White faces (which were randomly intermixed), but only when the faces were presented briefly and masked such that participants did not report seeing the faces. For faces that could be clearly seen and thus consciously processed, decreased amygdala activation for Black relative to White faces was accompanied by activation in areas of the PFC and the anterior cingulate gyrus — areas associated with cognitive control.

Interestingly, it appears that mental activities that counteract prejudiced thoughts may diminish control in other situations. According to Baumeister, Bratslavsky, Muraven, and Tice (1998), self-regulation is a limited resource, and any act of control not only uses up resources at the time of control, but also for some time afterward while the system recuperates. Richeson and Shelton (2003) found that after non-prejudiced White participants interacted with a Black individual — a task that may require cognitive control for participants who harbor prejudice but want to act or appear egalitarian — they subsequently performed worse on the Stroop task, a task that requires cognitive control

for incompatible trials (e.g., reporting that the word green is in a red print color). In a follow-up fMRI study, Richeson and colleagues (2003) scanned White participants while they viewed Black and White faces. Afterward, participants performed the Stroop task. As in Cunningham and colleagues (2004), greater activation was observed in the right lateral PFC while participants viewed Black compared with White faces. Furthermore, the degree of right PFC activity while viewing Black faces during the fMRI task predicted subsequent Stroop performance, with those with the most right PFC activity during fMRI performing the worst on the Stroop task. Presumably, the cognitive cost of control was manifested in the subsequent cognitive task. This pattern of findings provides support for the idea that nonprejudiced participants attempt to regulate their emotional responses to Black faces.

In light of these findings (and the work discussed earlier), the amygdala should not be considered a source of prejudice, but as a component in a larger framework of competing automatic and controlled processes which modulate the expression of prejudice. For example, Phelps, Cannistraci, and Cunningham (2003) reported on a patient who, despite bilateral amygdala damage, still showed evidence of automatic race biases on an indirect measure of automatic associations, suggesting that automatic evaluative responses are possible without an amygdala. This is consistent with the position that the amygdala is better characterized as a responsive component to motivationally relevant stimuli, instead of strictly negative, fearful, or threatening stimuli (Canli et al., 2005; Cunningham et al., 2004, 2008; Mather et al., 2004).

In the processing of social groups, and people or objects in general, other areas are associated with processing emotional intensity and valence, notably, the right prefrontal and orbitofrontal cortex. While explicit acts of hate and more overt forms of discrimination are the most salient features of prejudice, they often overshadow a complementary form of prejudice in the form of positive associations toward ingroup members. Indeed, the history of intergroup conflict provides strong evidence that "ingroup love" is a more common root of discrimination than "outgroup hate" (Brewer, 1999). Moreover, in contexts where discrimination arises as a result of differential evaluations of two groups, ingroup bias can lead to the same patterns of discrimination as outgroup derogation. Take, for example, racial discrimination in the context of a hiring decision: ingroup bias

and outgroup derogation would both lead a White candidate to be hired over a Black candidate. Although these decisions are the result of different affective processes, the result is identical — the candidates receive unequal levels of treatment and discrimination ensues.

Recent research has begun to examine the neural processes involved in these ingroup and outgroup biases. In one study, participants were randomly assigned to a novel mixed-race team without a history of contact or conflict with an outgroup team (Van Bavel, Packer, & Cunningham, 2008). Following assignment, participants spent three minutes memorizing the team membership of 24 faces, and these faces were presented during fMRI scanning. Unlike previous research that has found greater amygdala activation to Black than White faces, this study found greater amygdala activation to team ingroup than outgroup faces, regardless of race. Although this may seem counterintuitive if one takes the view that the amygdala responds only (or primarily) to threatening stimuli, it is consistent with the view that amygdala activation reflects the processing of the affective nature of motivationally significant stimuli (Anderson & Phelps, 2001; Cunningham et al., 2008). In many cases, negative or threatening stimuli take this role; however, in some situations positive stimuli can have greater motivational significance. The results from this study imply that, in the absence of intergroup conflict and outgroup derogation, the ingroup may be motivationally primary (see also, Allport, 1954). This suggestion was bolstered by finding additional activations in reward processing regions, such as the striatum and orbitofrontal cortex, that show greater activation to ingroup than outgroup faces and that significantly correlated with self-reported preferences for ingroup (vs. outgroup) members.

Political Attitudes

As social neuroscience investigations have begun to inform our understanding of basic social psychological processes, researchers have started to apply these findings to improve our understanding of political behavior. Although research in this area has typically focused on replicating previous evaluative effects in the domain of political judgment (e.g., Kaplan, Freedman, & Iacoboni, 2007), other research has taken into consideration that fact that the evaluation of political candidates differs from other evaluative categories in that people rarely have direct exposure to the candidate, and the evaluation is often influenced by group membership and

political ideology (Westen, Blagov, Harenski, Kilts, & Hamann, 2006). Though undecided voters may be relatively more open to new information, once people have made a decision about which political candidate to support in an election they have a tendency to minimize the influence of any new information that may be inconsistent. For example, once one has decided to support candidate X, he or she may not want to learn about any inconsistencies in candidate X's statements. On this view, motivated reasoning helps to maintain attitudes in the face of inconsistent information (e.g., Rahn, 1993), and as such, the study of political attitudes may allow for a better understanding of the motivational components of attitude acquisition, expression, and change.

Although motivated reasoning often prevents attitude change, there are situations in which attitudes about political candidates may, and perhaps should, change. Voters who are uncertain, have weak or ambivalent attitudes, or those who are low on identification with a political party may be more likely to show attitude change in response to new information about a political candidate (e.g., Lavine, 2001; Lodge & Taber, 2005; McGraw, Hasecke, & Conger, 2003). Unlike attitude change resulting from direct experience, attitude change here likely involves the integration of new semantic information or inferred group membership to generate new attitudes. Thus, attitude change should be expected to require more prefrontal components that may serve to shape attitude representations. Consistent with this idea, when exposed to negative political advertising, participants who showed greater activation in the dorsolateral prefrontal cortex also showed greater negative attitude change (Kato, et al., 2009). This pattern of results supports the hypothesis that, unlike reversal learning in simple conditioning, attitude change for more abstracted information like political attitudes may require a reorganization of information mediated by the prefrontal cortex (O'Reilly, Noelle, Braver, & Cohen, 2002).

Political evaluation is often considered to be a deliberate act, where people weigh various options and develop a coherent political ideology. Yet, political ideology can also operate relatively automatically, coloring and shaping our perceptions of people and the information that they present. For instance, while people with low political expertise may be more open to new information, people with high political expertise and identification are more likely to interpret new information in terms of automatically accessible schemas. Knutson, Wood, Spampinato, and Grafman (2006) demonstrated that, in contrast

to people who were less politically involved, highly politically identified participants had less lateral prefrontal cortex activity when responding to politically relevant stimuli. These data suggest that highly identified people engaged with particular information in a less deliberate fashion. As such, political experts may be able to rely on already established schemas to make quick evaluations, which may allow them to automatically interpret and incorporate consistent information into existing schemas.

Political ideology can also function as a social identity and guide our perceptions of others. For example, in one study participants were asked to think about the opinions and preferences of a person who had a similar or dissimilar political affiliation (Mitchell, Macrae, & Banaji, 2006). It was assumed that participants who more strongly identified with a certain political party would process a similar-minded person as an ingroup member, and therefore activate brain areas that have been linked to self-referential processing. For example, liberals were expected to be able to understand the mental states of another liberal more than a conservative, and the converse was expected for conservatives. Considering the mental state of a similar other led to activity in ventral areas of the medial prefrontal cortex (PFC), whereas considering the mental state of a dissimilar other led to activity in more dorsal areas of the medial PFC. Interestingly, individuals who strongly self-categorized with a political group, as measured by the IAT, had greater ventral medial PFC activity to politically similar others and less dorsal medial PFC activity to dissimilar others. Because regions of the medial PFC have previously been implicated in building mental models of other minds and simulating the thoughts and feelings of other people (a process called mentalizing; Mitchell, 2006), with more ventral areas being more involved in the processing of self-relevant information (Kelley et al., 2002), the authors concluded that although similar and dissimilar others both recruit regions involved in understanding others, similar others were more likely to be processed like the self.

Summary

Although understanding attitudes presents a considerable challenge for scientists, there is a rich history of relevant theoretical ideas and findings from social and cognitive psychology and intriguing new findings from social cognitive neuroscience. Borrowing from the literatures on reinforcement learning and affective processing, we can appreciate the intricate neural systems that

attitudes operate upon. Interestingly, many of the same brain regions that are involved in expression, formation, and change appear to be identical, suggesting a unified set of dynamic processes that can give rise to multiple attitudinal phenomena. As such, evaluation may reflect the current processing of an integrated information processing system at any given time. The particular ways in which information is constrained, weighed, and integrated as evaluations are constructed online from attitudes and the exact computations of the various brain regions involved in these processes will require more in-depth investigation.

More importantly, much of this research has examined relatively simple learning paradigms, or responses to relatively simple attitude stimuli (i.e., an aversive task, shock, or facial expression). Although attitudes can be formed under several of these types of situations, we have suggested that they are much more likely to be formed by a complex interaction of social and contextual factors. Using neuroimaging techniques, we are just beginning to uncover how different brain regions contribute to the formation and change of attitudes, and particularly how these can be applied to important social issues such as prejudice and political behavior. A more thorough investigation of the interplay between attitudes, the situation, and the mind is crucial for a deeper understanding of attitudes, both whence they came from and where they are going.

Note

1. Although the NAcc is correlated with buying behavior, it is unclear from these fMRI studies whether this is the result of the motivational dopaminergic "wanting" or the more hedonic opioid "liking" subdivisions of the NAcc (Berridge, Robinson, & Aldridge, 2009).

References

Adolphs, R., Russell, J. A., & Tranel, D. (1999). A role for the human amygdala in recognizing emotional arousal from unpleasant stimuli. *Psychological Science, 10,* 167–171.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1995). Fear and the human amygdala. *Journal of Neuroscience, 15,* 5879–5891.

Adolphs, R., Tranel, D., Hamann, S., Young, A. W., Calder, A. J., Phelps, E. A., et al. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia, 37,* 1111–1117.

Aggleton, J. P., Burton, M. J., & Passingham, R. E. (1980). Cortical and subcortical afferents to the amygdala of the rhesus monkey (macaca mulatta). *Brain Research, 190,* 347–368.

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

Amaral, D. G. & Price, J. L. (2004). Amygdalo-cortical projections in the monkey (macaca fascicularis). *The Journal of Comparative Neurology, 230,* 465–496.

Anderson, A. K., Christoff, K., Steppen, I., Panltz, D., Ghahremani, D. G., Glover, G., et al. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience, 6,* 196–202.

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature, 411,* 305–309.

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ : Prentice Hall.

Beer, J. S., Heery, E. A., Keltner, D., Scabini, D., & Knight, R. T. (2003). The regulatory function of self–conscious emotion: Insights from patients with orbitofrontal damage. *Journal of Personality and Social Psychology, 85,* 589–593.

Berntson, G. G., Bechara, A., Damasio, H., Tranel, D., & Cacioppo, J. T. (2007). Amygdala contribution to selective dimensions of emotion. *Social Cognitive and Affective Neuroscience, 2,* 123–129.

Berridge, K. C., Robinson, T. E. & Aldridge, J. W. (2009). Dissecting components of reward: "Liking," "wanting," and learning. *Current Opinion in Pharmacology, 9,* 65–73.

Blair, R. J. R. (2004). The roles of orbital frontal cortex in the modulation of antisocial behavior. *Brain and Cognition, 55,* 198–208.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A. & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron, 30,* 619–639.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55,* 429–444.

Bunge, S. A. & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science, 15,* 118–121.

Bush, G., Luu, P., Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulated cortex. *Trends in Cognitive Science, 4,* 215–222.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115,* 401–423.

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.

Canli, T., Congdon, E., Gutknecht, L., Constable, R. T., & Lesch, K. P. (2005). Amygdala responsiveness is modulated by tryptophan hydroxylase-2 gene variation. *Journal of Neural Transmission, 112,* 1479–1485.

Cardinal, R. N., Parkinson, J. A., Lachenal, G., Halkerston, K. M., Rudarakanchana, N., Hall, J., et al. (2002). Effects of selective excitotoxic lesions of the nucleus accumbens core, anterior cingulated cortex, and central nucleus of the amygdala on autoshaping performance in rats. *Behavioral Neuroscience, 116,* 553–567.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 1,* 747–749.

Carver, C. S. & Scheier, M. F. (2001). *On the self-regulation of behavior.* Cambridge, MA : Cambridge University Press.

Chudasama, Y. & Robbins, T. W. (2003). Dissociable contributions of the orbitofrontal and infralimbic cortex to Pavlovian autoshaping and discrimination reversal learning: Further evidence for the functional heterogeneity of the rodent frontal cortex. *The Journal of Neuroscience, 23,* 8771–8780.

Cools, R. (2005). Dopaminergic modulation of cognitive function — Implications for l-DOPA treatment in Parkinson's disease. *Neuroscience and Biobehavioral Reviews, 30,* 1–23.

Cunningham, W. A., Kesek, A., Mowrer, S.M. (2009). Distinct orbitofrontal regions encode stimulus and choice valuation. *Journal of Cognitive Neuroscience.* 21, 1956–1966.

Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology, 85,* 639–649.

Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes . *Journal of Cognitive Neuroscience, 16,* 1717–1729.

Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2005). Neural correlates of evaluation associated with promotion and prevention regulatory focus. *Cognitive, Affective, & Behavioral Neuroscience, 5,* 202–211.

Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective flexibility: Evaluative processing goals shape amygdala activity. *Psychological Science, 19,* 152–160.

Cunningham, W. A. & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences, 11,* 97–104.

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: A multilevel framework for attitudes and evaluation. *Social Cognition, 25,* 736–760.

Davis, M. (1992). The role of the amygdala in fear and anxiety. *Annual Review of Neuroscience, 15,* 353–375.

Davis, M. & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry, 6,* 13–34.

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience, 8,* 1611–1618.

Elliott, R., Newman, J. L., Longe, O. A., & William Deakin, J. F. (2003). Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: A parametric functional magnetic resonance imaging study. *The Journal of Neuroscience, 23,* 303–307.

Everitt, B. J., Cardinal, R. N., Hall, J., Parkinson, J. A., & Robbins, T. R. (2000). Differential involvement of amygdala subsystems in appetitive conditioning and drug addiction. In J. P. Aggleton (Ed.), *The amygdala: A functional analysis* (pp. 353–390). Oxford : Oxford UP.

Everitt, B. J., Parkinson, J. A., Olmstead, M. C., Arroyo, M. Robledo, P. & Robbins, T. W. (1999). Associative processes in addiction and reward: The role of amygdala-ventral striatal subsystems. *Annals of the New York Academy of Sciences, 877,* 412–438.

Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 14, pp. 161–202). New York : Academic Press.

Fellows, L. K. & Farah, M. J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: Evidence from a reversal learning paradigm. *Brain, 126,* 1830–1837.

Frank, M. J. & Claus, E. D. (2006). Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making and reversal. *Psychological Review, 113,* 300–326.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences, 41,* 16311–16316.

Freese, J. & Amaral, D. (2009)Neuroanatomy of the primate amygdala. In P. Whalen & E. Phelps, (Eds.). *The human amygdale* (pp. 3–42). New York : Guilford.

Garavan, H., Pendergrass, J. C., Ross, T. J., Stein, E. A., & Risinger, R. C. (2001). Amygdala response to both positive and negatively valenced stimuli. *NeuroReport, 12,* 2779–2783.

Gawronski, B. & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132,* 692–731.

Gerfen, C. R. (2000). Molecular effects of dopamine on striatal projection pathways. *Trends in Neurosciences, 23,* 64–70.

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science, 22,* 1104–1107.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480.

Hamann, S. B., Ely, T., D., Hoffman, J. M., & Kilts, C. D. (2002). Activation of the human amygdala in positive and negative emotion. *Psychological Science, 13,* 135–141.

Hamann, S. B. & Mao, H. (2002). Positive and negative emotional verbal stimuli elicit activity in the left amygdala. *Brain Imaging, 13,* 15–19.

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,* 2351–2354.

Hazy, T. E., Frank, M. J. & O'Reilly, R. C. (2007). Towards an executive without a homunculus: Computational models of the prefrontal cortex/basal ganglia system. *Philosophical Transactions of the Royal Society B, 362,* 1601–1613.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52,* 1280–1300.

Holroyd, C. B. & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and error-related negativity. *Psychological Review, 109,* 679–709.

Houk, J. C., Adams, J. L., & Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 249–270). Cambridge, MA : MIT Press.

Hutchison, W. D., Davis, K. D., Lozano, A. M., Tasker, R. R., & Dostrovsky, J. O. (1999). Pain-related neurons in the human cingulate cortex. *Nature Neuroscience, 2,* 403–405.

Isenberg, N., Silbersweig, D., Engelien, A., Emmerich, S., Malavade, K., Beattie, B., et al. (1999). Linguistic threat activates the human amygdala. *Proceedings of the National Academy of Sciences, 96,* 10456–10459.

Jog, M. S., Kubota, Y., Connolly, C. I., Hillegaart, V., & Graybiel, A. M. (1999). Building neural representations of habits. *Science, 286,* 1745–1749.

Kaplan, J. T., Freedman, J., & Iacoboni, M. (2007). Us versus them: Political attitudes and party affiliation influence neural responses to faces of presidential candidates. *Neuropsychologia, 45,* 55–64.

Kato, J., Ide, H., Kabashima, I., Kadota, H., Takano, K., & Kansaku, K. (2009). Neural correlates of attitude change following positive and negative advertisements. *Frontiers in Behavioral Neuroscience, 3,* 1–13.

Kawasaki, H., Kaurfman, O., Damasio. H., Damasio, A. R., Granner, M., Bakken, H., et al. (2001). Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. *Nature Neuroscience, 4,* 15–16.

Kelley, A. E. & Berridge, K. C. (2002). The neuroscience of natural rewards: Relevance to addictive drugs. *The Journal of Neuroscience, 22,* 3306–3311.

Khamassi, M., Lacheze, L., Girard, B., Berthoz, A., & Guillot, A. (2005). Actor-critic models of reinforcement learning in the basal ganglia. *Adaptive Behavior, 13,* 131–148.

Kim, H., Shimojo, S., & O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biology, 4,* 1453–1461.

Kluver, H. & Bucy, P.C. (1937). Psychic blindness and other symptoms following bilateral temporal lobectomy in rhesus monkeys. *American Journal of Psychology, 119,* 352–353.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience, 21,* 159–164.

Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fMRI. *Neuroimage, 18,* 263–272.

Knutson, B., Scott, R., Wimmer, G. E., Prelec, D., & Loewenstein, G. (2007). Neural predictors of purchases. *Neuron, 53,* 147–156.

Knutson, K. M., Wood, J. N., Spampinato, M. V., & Grafman, J. (2006). Politics on the brain: An fMRI investigation. *Social Neuroscience, 1,* 25–40.

Kringelbach, M. L. (2005). The human orbitofrontal cortex: Linking reward to hedonic experience. *Nature Reviews Neuroscience, 6,* 691–702.

Kringelbach, M. L., O'Doherty, J., Rolls, E. T., & Andrews, C. (2003). Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness . *Cerebral Cortex, 13,* 1064–1071.

Kringelbach, M. L. & Rolls, E. T. (2003). Neural correlates of rapid reversal learning in a simple model of human social interaction. *NeuroImage, 20,* 1371–1383.

Kringelbach, M. L. & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. *Progress in Neurobiology, 72,* 341–372.

LaBar, K. S., Gatenby, C., Gore, J. C., LeDoux, J. E., Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: A mixed trial fMRI study. *Neuron, 20,* 937–945.

LaBar, K. S., LeDoux, J. E., Spencer, D. D., & Phelps, E. A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *Journal of Neuroscience, 15,* 6846–6855.

Lavine, H. (2001). The electoral consequences of ambivalence toward presidential candidates. *American Journal of Political Science, 45,* 915–929.

LeDoux, J. E., Cicchetti, P., Xagoraris, A., & Romanski, L. M. (1990). The lateral amygdaloid nucleus: Sensory interface of the amygdala in fear conditioning. *Journal of Neuroscience, 10,* 1062–1069.

Liberzon, I., Phan, K. L., Decker, L. R., & Taylor, S. F. (2003). Extended amygdala and emotional salience: A PET activation study of positive and negative affect. *Neuropsychopharmacology , 28,* 726–733.

Lodge, M. & Taber, C. S. (2005). The automaticity of affect for political leaders, groups, and issues: An experimental test of the hot cognition hypothesis. *Political Psychology, 26,* 455–482.

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288,* 1835–1838.

Mather, M., Canli, T., English, T., Whitfield, S., Wais, P., Ochsner, K., Gabrieli, J. D. E., & Carstensen, L. L. (2004). Amygdala responses to emotionally valenced stimuli in older and younger adults. *Psychological Science, 15,* 259–263.

McGraw, K. M., Hasecke, E., & Conger, K. (2003). Ambivalence, uncertainty, and processes of candidate evaluation. *Political Psychology, 24,* 421–448.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron, 50,* 655–663.

Montague, P. & Berns, G. (2002). Neural economics and the biological substrates of valuation. *Neuron, 36,* 265–284.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., et al. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature, 383,* 812–815.

Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the amygdala. *Nature, 393,* 467–470.

Muraven, M., Tice, D. M., Baumeister, R. F. (1998). Self-control as a limited resource. *Journal of Personality and Social Psychology, 74,* 774–789.

Murray, E. A. (2007). The amygdala, reward and emotion. *TRENDS in Cognitive Sciences, 11,* 489–497.

Murray, E. A., O'Doherty, J. P., & Schoenbaum, G. (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *The Journal of Neuroscience, 271,* 8166–8169.

Nobre, A., Coull, J., Frith, C., Mesulam, M. (1999). Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nature Neuroscience,* 2, 11–12.

Northoff, G., Richter, A., Gessner, M., Schlagenhauf, F., Fell, J., Baumgart, F., et al. (2000). Functional dissociation between medial and lateral prefrontal cortical spatiotemporal activation in negative and positive emotions: A combined fMRI/MEG study. *Cerebral Cortex, 10,* 93–107.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience, 14,* 1215–1229.

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.

O'Doherty, J., Critchley, H., Deichmann, R., & Dolan, R. J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *The Journal of Neuroscience, 23,* 7931–7939.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 16,* 452–454.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4,* 95–102.

Ohman, A. (2005). The role of the amygdala in human fear: Automatic detection of threat. *Psychoneuroendocrinology, 30,* 953–958.

Olsson, A., Nearing, K. I., & Phelps, E. A. (2007). Learning fears by observing other: The neural systems of social fear transmission. *Social Cognitive and Affective Neuroscience, 2,* 3–11.

O'Reilly, R. C., Noelle, D. C., Braver, T. S., & Cohen, J. D. (2002). Prefrontal cortex and dynamic categorization tasks: Representational organization and neuromodulatory control. *Cerebral Cortex, 12,* 246–257.

Packard, M. G. & McGaugh, J. L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of Learning and Memory, 65,* 65–72.

Padoa-Schioppa, C. & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature, 441,* 223–226.

Petty, R. E. & Cacioppo, J. T. (1984). The effects of involvement on response to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology, 46,* 69–81.

Phelps, E. A., Cannistraci, C. J., & Cunningham, W. A. (2003). Intact performance on an indirect measure of race bias following amygdala damage. *Neuropsychologia, 41,* 203–208.

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.

Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Gore, J. C., Grillon, C., & Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nature Neuroscience, 4,* 437–441.

Rahn, W. M. (1993). The role of partisan stereotypes in information processing about political candidates. *American Journal of Political Science, 37,* 472–496.

Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S., et al. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience, 6,* 1323–1328.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 206,* 443–447.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10,* 284–294.

Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain and Cognition, 55,* 11–29.

Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery, and Psychiatry, 57,* 1518–1524.

Rolls, E. T., McCabe, C., & Redoute, J. (2008). Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebral Cortex, 18,* 652–663.

Rougier, N.P., Noelle, D., Braver, T.S., Cohen, J.D. & O'Reilly, R.C. (2005). Prefrontal cortex and the flexibility of cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences, 102,* 7338–7343.

Roskos-Ewoldsen, D. R. & Fazio, R. H. (1992). On the orienting value of attitudes. *Journal of Personality and Social Psychology, 63,* 198–211.

Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: An evolved system for relevance detection. *Reviews in the Neurosciences, 14,* 303–316.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275,* 1593–1599.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London, 298,* 199–209.

Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., & Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron, 39,* 701–711.

Stefanacci, L. & Amaral, D. G. (2000). Topographic organization of cortical inputs to the lateral nucleus of the macaque monkey amygdala: A retrograde tracing study. *The Journal of Comparative Neurology, 421,* 52–79.

Stuss, D. T. & Benson, D. F. (1986). The frontal lobes and control of cognition and memory. In E. Perecman (Ed.), *The frontal lobes revisited* (pp. 144–158). New York : Raven.

Sutton, R. S. & Barto, A. G. *Reinforcement learning.* MIT Press, Cambridge, MA, 1998.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science, 26,* 515–518.

Thorndike, E. (1911). *Animal intelligence: Experimental studies.* New York : Macmillan.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: A functional magnetic resonance imaging investigation. *Psychological Science, 19,* 1131–1139.

Wallis, J. D. & Miller, E. K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *European Journal of Neuroscience, 18,* 2069–2081.

Watson, J.B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–14.

Weiskrantz, L. (1956). Behavioral changes associated with ablation of the amygdaloid complex in monkeys. *Journal of Comparative Physiological Psychology, 49,* 381–391.

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. *Journal of Cognitive Neuroscience, 18,* 1947–1958.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of Neuroscience, 18,* 411–418.

Yin, H. H., Ostlund, S. B., Knowlton, B. J., & Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience, 22,* 513–523.

Zald, D. H. & Pardo, J. V. (1997). Emotion, olfaction, and the human amygdala: Amygdala activation during aversive olfactory stimulation. *Proceedings of the National Academy of Sciences, 94,* 4119–4124.

Zelazo, P. D. (2003). The development of conscious control in childhood. *Trends in Cognitive Sciences, 8,* 12–17.