

available at www.sciencedirect.comwww.elsevier.com/locate/brainres

**BRAIN
RESEARCH**

Research Report
Integrating automatic and controlled processes into neurocognitive models of social cognition
Ajay B. Satpute, Matthew D. Lieberman*
University of California, Los Angeles, CA 90024, USA

ARTICLE INFO
Article history:

Accepted 5 January 2006

Available online 20 February 2006

Keywords:

Social Cognitive neuroscience

Social cognition

Automaticity

Control

Dual process model

ABSTRACT

Interest in the neural systems underlying social perception has expanded tremendously over the past few decades. However, gaps between behavioral literatures in social perception and neuroscience are still abundant. In this article, we apply the concept of dual-process models to neural systems in an effort to bridge the gap between many of these behavioral studies and neural systems underlying social perception. We describe and provide support for a neural division between reflexive and reflective systems. Reflexive systems correspond to automatic processes and include the amygdala, basal ganglia, ventromedial prefrontal cortex, dorsal anterior cingulate cortex, and lateral temporal cortex. Reflective systems correspond to controlled processes and include lateral prefrontal cortex, posterior parietal cortex, medial prefrontal cortex, rostral anterior cingulate cortex, and the hippocampus and surrounding medial temporal lobe region. This framework is considered to be a working model rather than a finished product. Finally, the utility of this model and its application to other social cognitive domains such as Theory of Mind are discussed.

© 2006 Elsevier B.V. All rights reserved.

1. Introduction

Recently, the United States Senate held confirmation hearings to determine whether John Roberts ought to be the next Chief Justice of the Supreme Court. Several democratic senators, looking for black marks on Roberts' record to allow them to cast a no vote against the nominee of a conservative White House, expressed frustration over their inability to determine Roberts' beliefs and judicial philosophy because of the limited number of documents authored by Roberts' during his short tenure as a judge. Indeed, senators seized upon briefs written by Roberts when he served as a White House council in the Reagan administration 20 years earlier as a young attorney to demonstrate that he held unacceptably conservative views. Roberts deftly sidestepped these allegations by reminding the

senators that, "I was a staff lawyer. I didn't have a position. The administration had a position".

This social discourse between the senators and Roberts highlights the three main aspects of attribution theory as studied by dual-process researchers of social cognition over the past 20 years: task goals (determining Roberts' judicial positions), hypothesis-confirming automatic attributions (using briefs Roberts wrote for the Regan administration as indicative of his position), and controlled corrections of the automatic attributions (the diagnosticity of these briefs as a basis of their assessments) are each mainstays of the modern theories of attribution (Gilbert, 1989; Trope and Gaunt, 1999). This research has been propelled forward for decades from the earlier finding (Ichheiser, 1943; Jones and Harris, 1967; Ross, 1977) that individuals consistently overestimate the degree to

* Corresponding author.

E-mail address: lieber@ucla.edu (M.D. Lieberman).

which an observed behavior is indicative of the corresponding personality trait or attitude. Thus, in the Roberts case, the Senators may have ignored Roberts' constrained situation when writing those briefs as White House council, as his social role required him to advocate the position held by the White House regardless of his own personal views on the matter. The crux of dual-process models of attribution has been to argue that in terms of mental effort, inferring the corresponding trait from behavior is easy (i.e., writing sexist briefs indicates sexist beliefs), whereas recognizing the exculpatory role of situational influences is hard (i.e., any position one is obligated to because of professional responsibilities does not reflect on one's own dispositions).

Aside from social attribution, dual-process models have been a guiding theoretical tool in a variety of other domains in social cognition. Indeed, there seem to be few domains in social cognition that have not benefited by its application (see Chaiken and Trope, 1999). Moreover, they form a substantial theoretical basis for cognitive psychology as well including categorization (Murphy, 2002), memory (Squire and Zola, 1996), reasoning (Sloman, 1996), and decision-making (Kahneman, 2003). Though dual-process models are acknowledged to be a rough cut at the complexity of social behavior, they have provided behavioral researchers with an investigative tool that is generative, explanatory, and simple. In common, these models incorporate an age-old distinction between automaticity and control: that some processes occur quickly without requiring explicit thought or effort while others occur slowly and do require explicit thought and effort.

Although dual-process models of automaticity and control have made important contributions to behavioral research on social cognition, the distinction has been surprisingly absent from theorizing about the neural bases of social perception. Dozens of neuroimaging studies have identified brain regions involved in theory of mind (ToM), perspective taking, trait

identification, imitation, and other facets of social cognition and perception. Thus far, these studies have implicated medial prefrontal cortex (MPFC; BA 8/9/10), lateral prefrontal cortex (LPFC), lateral temporal cortex (LTC) including the superior temporal sulcus (STS) and tempo-parietal junction (TPJ), precuneus in medial posterior parietal cortex (MPPC), lateral posterior parietal cortex (LPPC), amygdala, ventral striatum (VS), and ventromedial prefrontal cortex (VMPFC; BA 11) as important regions of interest in social cognition. However, none of these studies have examined the neural correlates of automatic versus controlled social perception, nor do they invoke the distinction to explain differences from one study to the next despite its continued relevance to behavioral work. This might be because much of the cognitive neuroscience work in this area comes from perception or developmental traditions that do not rely as much on the dual-processing distinction. Growing interest in the neural bases of social cognition has pulled researchers from these differing traditions together, which has prompted a need for bridging gaps between their various theoretical takes.

We have previously proposed a division of neural processes (Lieberman et al., 2002; Lieberman et al., 2004b), which roughly corresponds to automatic and controlled social perception (Gilbert, 1989; Smith and DeCoster, 1999; Strack and Deutsch, 2004). In this model, we have proposed that several brain regions, together called the X-system (for the 'x' in reflexive) are involved in automatically coding the trait and evaluative implications of observed behaviors, whereas another set of brain regions called the C-system (for the 'c' in reflection) are responsible for holding inferential goals in mind and for taking situational constraint information and other prior knowledge into account to alter the dispositional inferences drawn from observed behaviors (Fig. 1). In this article, we review the neural bases of these two systems, how they may differentially contribute to social perception, and how

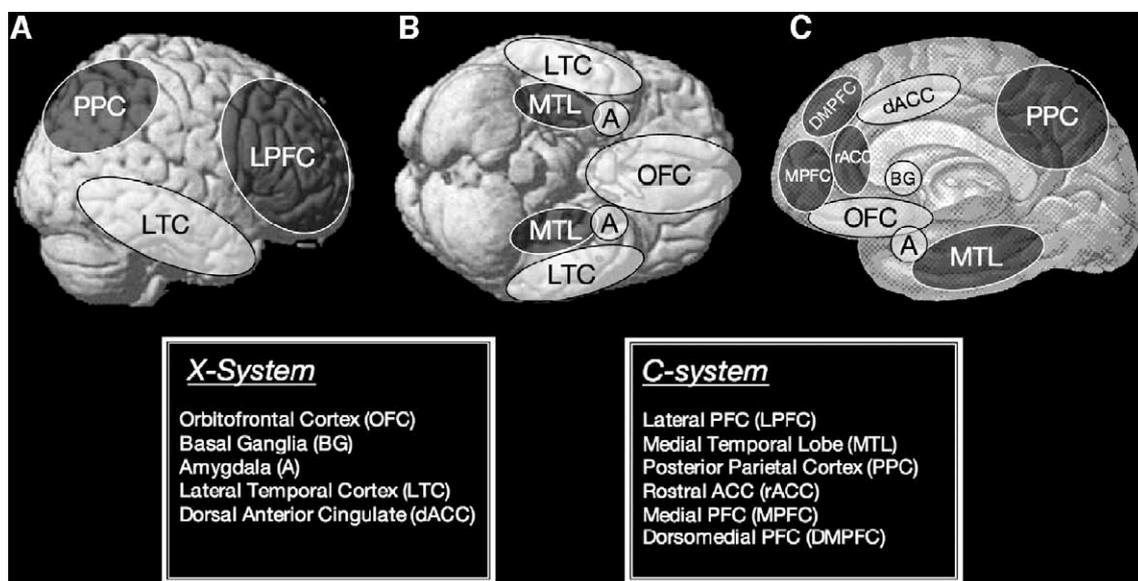


Fig. 1 – Neural correlates of the C-system and X-system displayed on a canonical brain rendering from (A) lateral, (B) ventral, and (C) medial views. Note: the hippocampus, nucleus accumbens, and amygdala are actually subcortical structures that are here on the cortical surface for ease of presentation.

this distinction may be useful for future research on social perception.

2. The reflective and reflexive systems

Since the original formulation of the X- and C-systems and in the development of this model since (Lieberman, *in press*), several criteria have been used to identify candidate regions and assign them to one system or the other. Neural components of the C-system are proposed to be fast learning, slow operating, symbolic, or propositional structures (Table 1). The operations of the C-system are typically experienced as an internal linguistic monologue emerging in a freely chosen way from oneself and are associated with the experience of agency or will. Because of the symbolic computational ability of the C-system, it is best suited to represent asymmetric relationships (e.g., 'if x then y' which does not imply 'if y then x'), exceptions, special cases, negation, and counterfactuals. Considering that symbolic capabilities are recent developments in evolution (Deacon, 1998), C-system structures are also likely to be phylogenetically younger than X-system structures.

Neural components of the X-system are proposed to be slow learning, fast operating, bidirectional, parallel-processing structures (Table 1). The operations of the X-system may or may not be conscious, but are not reflectively conscious. Because of the bidirectional nature of X-system representations and processes, it is best suited for processing implicit semantic and evaluative associations (e.g., dog ↔ cat, terrorists ↔ bad), rather than asymmetric relationships. These associations are not fixed or rigid; rather, the semantic and affective evaluation of a given stimulus by the X-system is in flux with the environmental context and constraint satisfaction processes within the X-system neural architecture (Mitchell et al., 2003).

The C-system is posited to include lateral prefrontal cortex, posterior parietal cortex, medial prefrontal cortex, rostral anterior cingulate cortex (rACC), and the hippocampus and surrounding medial temporal lobe region (MTL). The X-system is posited to include the amygdala, basal ganglia, ventromedial prefrontal cortex, dorsal anterior cingulate cortex (dACC), and lateral temporal cortex. Many of these claims are based on work that does not seem to be directly related to social cognition. This is because our model of symbolic and non-

symbolic processes cuts across the distinction between cognitive and social cognitive processes. Importantly, the model reveals points of contact between what may be considered disparate cognitive and social cognitive neural circuitry.

The collection of brain regions assigned to one system or the other is an evolving process as relevant data are collected on previously unassigned regions or new data suggest that a previous assignment should be updated. Thus, despite substantial cognitive, morphological, and phylogenetic reasons for the current assignments, it is unlikely that this model has seen the end of its evolution. It should be taken as a working model rather than a finished product.

3. The X-system for reflexive social cognition

3.1. Amygdala

The amygdala has been proposed to have numerous functions associated with both fearful (LeDoux, 2003) and rewarding (Baxter and Murray, 2002) stimulus properties, though it may be more sensitive to negative than positive valence (Wager et al., 2003). Projections from the amygdala produce a cascade of behavioral responses geared towards fight or flight style responses (LeDoux, 1996, 2003). Placing the amygdala in the X-system is supported by studies showing that the amygdala responds to subliminal presentations of fearful stimuli (Cunningham et al., 2004; Liddell et al., 2005; Morris et al., 1999; Pasley et al., 2004; Whalen et al., 1998), fear conditioning in the absence of awareness of the unconditioned stimulus (Morris et al., 1998), and by relative robustness of response despite attentional modulation (Öhman, 2005). The response properties of the amygdala towards fearful stimuli may be modulated by anxiety (Pessoa et al., 2002), possibly indicating that the automatic evaluation of stimulus relevance may depend on a person's personality traits as determined by genetics or personal history (Hariri et al., 2002).

In contrast to the general guidelines for assigning a region to the X-system, the amygdala is capable of fast learning under certain circumstances, although it does typically learn slowly over time in conditioning paradigms. This anomaly of amygdala function may be the result of further specialization of the amygdala than is generally considered in human neuroimaging research. Swanson has argued on the basis of morphological characteristics and neurotransmitter receptor distributions that there are components of the amygdala, the lateral and basolateral nuclei, that are actually part of the cortex in the medial temporal lobe, and that these are the regions of the amygdala that are critical for single-trial learning in the amygdala (Swanson and Petrovich, 1998).

3.2. Basal ganglia

The set of structures comprising the basal ganglia have been implicated in automatic components of affect, cognition, and behavior. The putamen and caudate nuclei seem to underlie implicit learning, probabilistic learning, and motor skill acquisition, which is slow to acquire but fast to execute once learned (Knowlton et al., 1996; Poldrack et al., 1999; Lieberman

Table 1 – Features associated with X- and C-systems

X-system	C-system
Parallel processing	Serial processing
Fast operating	Slow operating
Slow learning	Fast learning
Non-reflective consciousness	Reflective consciousness
Phylogenetically older	Phylogenetically newer
Representation of symmetric relations	Representation of asymmetric relations
Representation of common cases	Representation of special cases
	Representation of abstract concepts (e.g., negation, time)

et al., 2004a). The ventral striatum, an interconnected circuit consisting of the nucleus accumbens and ventromedial aspects of the putamen, globus pallidus, and caudate nuclei, seems to be involved in the evaluation and prediction of both pleasant and unpleasant stimuli (Knutson et al., 2002, 2004), though it has most often been associated with a variety of reward-related stimuli including administration of drugs, winning money, and viewing pictures of a romantic partner (Bartels and Zeki, 2000; Breiter et al., 1997; Zalla et al., 2000).

Recent work in our laboratory has indicated that the ventral striatum is particularly responsive to processing subjective value, or providing an assessment of value that is contingent upon the current social context of the perceiver (Tabibnia et al., submitted for publication). That is, ventral striatum seems to be more responsive to the perceived fairness of a monetary offer than to the objective size of the monetary offer. For example, \$4 offered out of \$10 shows greater reward-related striatal activity than \$4 offered out of \$20.

Taken together, the basal ganglia seems to play a role in learning statistical regularities of the world and attaching emotional and motivational significance to them. Lieberman (2000) juxtaposed this learning environment to that underlying social intuition. Phenomenally, social intuition can be defined as the feelings, judgments, or hunches people have towards other social targets or situations that are often experienced in the absence of a well-articulated reason. In accordance with implicit learning paradigms, the basal ganglia may pick up on well-learned probabilistic regularities that exist in social interactions and situations, which may form the basis of social intuitions. Speculatively, this may include the implicit learning of social rules, expectations, and norms, and linking them with corresponding emotional and motivational states such as perceived fairness or unfairness when these expectations are violated. Critically, social psychologists have shown that these intuitive social processes often occur automatically without explicit intentions (Ambady, 1999; Cheng and Chartrand, 2003; Lowery et al., 2001).

3.3. Ventromedial prefrontal cortex (VMPFC)

The ventromedial prefrontal cortex lies at the intersection of the medial region of prefrontal cortex and the orbital or bottom region of prefrontal cortex (BA11). The damage to the famous patient Phineas Gage was primarily in VMPFC and Gage's deficits in social cognition are well known (Damasio, 1994). This area has strong parallel connections with the basal ganglia, amygdala and other limbic structures (Gaffan and Murray, 1990; Öngür and Price, 2001), and also seems to be required for the long-term formation of intuitions regarding stimulus-outcome probabilities prior to, or perhaps even irrespective of, explicit awareness of them (Bechara et al., 1997). Milne and Grafman (2001) have also observed that patients with VMPFC damage do not show evidence of implicit gender stereotyping on the implicit association test (Greenwald et al., 1998). Most recently, Deppe et al. (2005) demonstrated that susceptibility to framing effects is positively related to activity in VMPFC. Framing effects have been considered a consequence of capitalizing on operational properties of intuitive over deductive reasoning (Kahneman,

2003), further supporting the role of VMPFC in conjunction with the basal ganglia in the formation of automatic intuitions. Finally, a recent ERP study (Carretie et al., 2005) localized early responses (~150ms) to subliminally presented threat stimuli to VMPFC meeting two criteria of the X-system in sensitivity of subliminal presentations and speed of response.

3.4. Dorsal anterior cingulate cortex (dACC)

Previously, the anterior cingulate cortex as a whole was included as a C-system structure, however, given that dACC is involved in the emotional distress of physical and social pain (Eisenberger et al., 2003; Lieberman and Eisenberger, 2005), it may be more appropriate to characterize the dACC as an automatic alarm system that may reside in the X-system but be critical to 'notifying' the C-system that its abilities to respond flexibly and deal with novelty are needed. More on the relationship of dACC and rACC to the X- and C-systems can be found below in the section on rACC.

3.5. Lateral temporal cortex

The lateral temporal cortex (LTC) consists of lateral and inferior portions of the temporal lobes, the temporal poles, and the superior temporal sulcus. It has been implicated in semantic and conceptual processing in both imaging studies (Crinion et al., 2003; Mummery et al., 1999; Rissman et al., 2003; Rossell et al., 2001) and neuropsychological investigations of patients with semantic dementia (Garrard and Hodges, 2000; Mummery et al., 2000). LTC is particularly relevant to social cognition for semantic and perceptual processes rather than affective processes. These regions may contain the semantic information responsible for constructing stereotypes (Hart et al., 2000), individual impressions (Mitchell et al., 2002), and dispositional attributions (Lieberman et al., 2002).

Perhaps the most studied region of LTC from a social cognitive perspective is the superior temporal sulcus (STS). This area seems to be particularly relevant for recognition of other people, the actions they perform, and for positing intentions. In primates, single-cell recording studies in the STS have found a significant proportion of cells to be face-selective, some of which respond to specific individuals (Bruce et al., 1981; Perrett et al., 1982; Oram and Perrett, 1996). Some cells also appear selective for which direction a target is facing (e.g., towards the subject, or left or right profile views), translational body motion, and many respond to interactive combinations thereof (Oram and Perrett, 1996; Jellema et al., 2004). Cells in STS are also sensitive to hand-object interactions (Perrett et al., 1989) and eye-gaze of social targets (Jellema et al., 2000), indicating that STS may be involved with imputing intentions or goal states in others.

An analogous region in posterior STS has been found in humans. Imaging studies have shown STS activity in response to action observation (Rizzolatti et al., 1996; Iacoboni et al., 2001; Puce et al., 1998), biological motion (Grossman et al., 2000; Bonda et al., 1996; Pelphrey et al., 2003), and eye-gaze tracking (Puce et al., 1998; Pelphrey et al., 2004). Posterior STS has also been activated in studies exploring theory of mind,

perspective taking, and attribution of intentionality to perceptual Heider-Simmel type displays (Goel et al., 1995; Brunet et al., 2000; Gallagher et al., 2000; Saxe et al., 2004; Saxe and Kanwisher, 2003; Castelli et al., 2000). More recently, it has been found that activity in pSTS linearly increased as the perception of animate motion increased (Shultz et al., 2005), supporting the involvement of STS in forming intentional attributions (Blakemore and Decety, 2001).

Though STS is shown to be functionally and anatomically heterogeneous, collectively this region seems to be involved in automatic visual perception of social targets, identification of their actions, and attribution of their mental goal states to their behaviors (Allison et al., 2000). Further supporting the placement of this region in the X-system is the neural evidence that neuronal discharge in this region occurs relatively fast, within 200 ms of stimulus onset (Perrett et al., 1985; Allison et al., 2000), and the behavioral evidence that perception of biological motion and attribution of intentional states to others in Heider-Simmel type displays as well as to presented facial expressions is known to occur rapidly and effortlessly (Johansson, 1973; Scholl and Tremoulet, 2000; Ambady and Rosenthal, 1992).

4. The C-system for reflective social cognition

4.1. Lateral prefrontal cortex

Lateral prefrontal cortex is the primary structure involved in C-system processes and has been consistently activated in many effortful and resource-demanding cognitive tasks that require symbol manipulation including working memory (Cabeza and Nyberg, 2000; Baddeley, 1986), reasoning and logic (Noveck et al., 2004), fluid intelligence (Gray et al., 2003; Prabhakaran et al., 1997), analogy (Hummel and Holyoak, 2003), and mathematical problem-solving (Prabhakaran et al., 2001). It has been implicated in asymmetric relational reasoning such as causal reasoning in which relationships between events are unidirectional rather than bidirectional (e.g., X causes Y does not entail that Y causes X; Satpute et al., 2005); an ability which seems require symbolic computations (Hummel, 1999; Waldmann and Walker, 2005; Lieberman et al., 2002). Moreover, activity in LPFC is related to the experience of volitional or 'willful' actions (Lau et al., 2004), which is a critical component of C-system processes. In regard to social cognition, LPFC may also be involved in emotion-regulation (Ochsner and Gross, 2005) and behavioral inhibition (Aron et al., 2004).

4.2. Posterior parietal cortex

In combination with LPFC, lateral regions of posterior parietal cortex are commonly activated in tasks requiring working memory (Cabeza and Nyberg, 2000), and may also be involved in reasoning (Kroger et al., 2002; Satpute et al., 2005). Medial regions of PPC are activated under conditions of self-focused attention (Gusnard et al., 2001; Lieberman and Pfeifer, 2005) and perspective taking (Ruby and Decety, 2001), which may require symbolic representations in order to decouple other from self.

4.3. Rostral anterior cingulate cortex (rACC)

The ACC is often associated with conflict detection or error processing (Botvinick et al., 2001) and divided into regions dedicated to the processing of emotional (rACC) and cognitive (dACC) conflict (Bush et al., 2000). Although the rostral/dorsal distinction has been helpful in organizing a diverse set of findings associated with ACC activity in the 1990s, this distinction may not be able to account for a number of findings that have emerged since the time of that proposal.

Eisenberger and Lieberman (2004) suggested that the rACC and dACC could be better characterized in terms of non-symbolic and symbolic representations of conflict, respectively. They review studies of cognitive conflict, emotional processing, and pain and within each domain, some studies activated dACC and some activated rACC. However, if the object of conflict was explicit or reflected upon there was a greater tendency for rACC activity over dACC activity. For instance, negative emotions which by definition have a particular object, person, or event that they are focused on (i.e., being mad at a person) tended to activate rACC, whereas anxiety, which has been defined as fear in the absence of a particular object, is more associated with dACC activity. Similarly, unanticipated pain that is detected bottom-up from sensory processes activated dACC (Rainville et al., 1997; Eisenberger et al., 2003); however, anticipated pain recruits rACC but not dACC (Ploghaus et al., 2003).

Additionally, Smith (1945) observed that, in macaques, the entire ACC was morphologically analogous to human dACC. In other words, in primates that lack the capacity for symbolic processing (Thompson and Oden, 2000), there is no rACC. More recent evidence points to "spindle cells" (Nimchinsky et al., 1999) that appear in rACC, but not dACC, and are present in diminishing densities in the ACC of children and other primates paralleling the drop-off in symbolic capacities (Craig, 2004).

4.4. Medial temporal lobe

The medial temporal lobe (MTL) has been placed in the C-system because of its integral role in recollection, or memory retrieval accompanied with the conscious experience of a particular learning episode (Eldridge et al., 2000). On a phenomenal level, directed memory retrieval involving a conscious experience of the prior learning episode requires a temporary override of current perceptual inputs presented to the X-system so as to partake in the recollective experience. One use of this kind of recollective memory may be to serve as a repository for situations in which control could be useful (Lieberman, *in press*). That is, MTL may be useful in retrieving memories that either aid in identifying when control is needed and provide alternative behavior possibilities that run counter to habitual response patterns.

Properties of MTL and its interaction with striatal structures support this notion. The hippocampal structure in MTL has been implicated as a fast and flexible learning structure via a sparse representational coding scheme (McClelland et al., 1995) and may also be involved in overriding the slow and rote learning of habitual responses produced by striatal

systems (Miller and Cohen, 2001; Sloman, 1996; Poldrack and Packard, 2003). If a situation is encountered frequently enough, the X-system will change its habits to accommodate the situation. But since the X-system is generally slow to change, these C-system memories could be useful in the interim.

4.5. Medial prefrontal cortex (MPFC)

In previous iterations of the X- and C-systems, MPFC (BA 9/10; dorsal to VMPFC) has not been assigned to either system despite its obvious relevance to self and social cognition (Frith and Frith, 2003; Lieberman and Pfeifer, 2005). One consideration for assigning this region to the C-system is the fact that it is relatively larger in humans than other primates (Semendeferi et al., 2001). Additionally, cognitive load parametrically decreases the activity observed in MPFC (McKiernan et al., 2003) suggesting that it may not be able to operate in parallel with other controlled processes. At this point, the designation to the C-system is tentative and discussed further in the section on the automatization of social attribution processes.

5. Application to social perception

In our previous work (Lieberman et al., 2002), we described how the X- and C-systems could instantiate existing dual-process models of dispositional attribution. Specifically, this account focused on how LPFC would be involved in holding an attributional goal online, which would bias constraint satisfaction processes in LTC supporting automatic attributions to produce hypothesis-consistent attributions. These initial attributions could be corrected on the basis of situational information and the situation's role in producing the observed behavior in the form of propositional rules (e.g., if Roberts was only following orders as an advocate, then we should not assume that his behavior reflects his personal beliefs). The corrective process was presumed to rely on LPFC, although it now seems possible that MPFC regions may also be involved in this explicit consideration of social pressures. Here, rather than extending or refining this model, we take a step back to examine the forest rather than the trees, and in doing so make a much more modest proposal: that our understanding of the neural bases of social perception, ToM, and other related processes would profit from assessing whether some of the regions supporting these processes operate in more automatic or controlled ways. Knowledge of whether a particular ToM task relies on symbolic processes or not would indicate what kinds of manipulations would influence performance on this task.

The use of this approach relies on some important assumptions. First, it is assumed that some neural regions can be exclusively assigned to the X- or C-system. Second, it is assumed that a task that activates one of these neural regions implements the corresponding computational process (in this case, symbolic or non-symbolic computations). Hence, the use of this approach relies on 'reverse inference' logic (Poldrack, *in press*): the inference that a particular cognitive process is engaged by performance of a given task as indicated by a neural region activated by the task. At first glance, this may seem circular since it may be argued that these systems are in

part defined using the very data they purport to explain. However, this circularity is lessened since the tasks we used to define the X- and C-systems are separable from those which we hope to apply this paradigm to, and since converging methods including morphological, neuropsychological, and phylogenetic data are all used in assigning neural regions to systems.

5.1. Tests of automaticity in social attribution

5.1.1. Subliminal presentation

Behavioral research on social cognition typically uses one of two techniques to assess whether a process of interest is automatic or controlled. One technique is subliminal presentation of stimuli outside of awareness. If some process occurs even when there is no conscious awareness that the process has been invoked, this is strong evidence that the process is automatic (Bargh and Chartrand, 1999). A neural region of interest (ROI) that responds to subliminal presentation sufficiently demonstrates its involvement in the automatic processing of the stimulus, though it does not exclude the ROI's participation in controlled processes. However, our framework allows for subliminal presentation to be sufficient for an assignment of the region to the X-system. Subliminal presentation as a technique though is limited in its application because we live in a supraliminal world, albeit one that is not always reflected upon (Langer, 1989). Thus, while it has been applied to the processing of static images of emotional faces (Morris et al., 1999; Whalen et al., 1998), it would not be possible to present stories or picture sequences that depend on theory of mind, subliminally.

5.1.2. Cognitive load

The second commonly used technique, cognitive load, has been used repeatedly to disentangle the automatic and controlled components of attribution (Gilbert et al., 1988; Trope and Gaunt, 2000), stereotyping (Fein and Spencer, 1997; Gilbert and Hixon, 1991; Payne, 2001), and person knowledge (Macrae and Bodenhausen, 2000). Cognitive load consists of performing a resource-demanding task concurrently with a task of interest. To the extent that the task of interest is automatic, it will not be influenced by the load manipulation. One reason for this is that controlled processes operate serially rather than in parallel, resulting in task-switching costs. A second reason is that controlled processes may rely on a shared pool of resources (Baumeister et al., 1998; Kahneman, 1973); performing highly divergent but resource-demanding tasks one after the other results in worse performance on the secondary task. For instance, emotional self-regulation (e.g., suppressing emotions in response to a provocative stimulus) results in poorer subsequent performance in solving anagrams (Baumeister et al., 1998). In other words, the neural systems that support controlled processes may be similar to muscles in that they may tire in the short term with intense use and may strengthen over the long term if 'exercised' regularly. This indicates that though C-system structures may have their own distinct functional specialties, they may nonetheless rely on a shared resource pool.

Adapting the cognitive load paradigm to a functional neuroimaging environment is not straightforward, for how

are the hemodynamic consequences of cognitive load to be interpreted? Supposing dual-task performance causes an increase in activity in a given ROI relative to the task of interest alone. This increase could be attributed to (i) greater effort to maintain performance in the task of interest, (ii) operations required by the load task, or (iii) an interactive property between the two tasks (e.g., a third process is invoked to regulate task switching). To narrow down these possibilities, it is useful to know the neural properties of both tasks being performed individually. If the ROI is activated by cognitive load on its own and not by the task of interest, greater activity in the ROI during the dual-task situation could be interpreted as drawing upon more resources to maintain adequate performance for the load task. The same applies to the alternative situation involving the task of interest. In both cases, however, interactions between processes involved in both tasks may also be occurring. Some possibilities might be: (i) integration—the involved processes may integrate with each other making two independent tasks merge into a single combined task, (ii) process recruitment—the dual-task situation may recruit additional processes involved in monitoring and switching between the separate tasks, or (iii) strategic changes—the dual-task situation may cause participants to adopt new strategies in task performance. Interactive properties are also the case when an ROI is active under dual-task conditions, but not in either task presented alone.

Though the inferential consequences of load are numerous, in conjunction with theoretical expectations and corroboration with behavioral results, cognitive load may prove to be a useful tool to ascertain whether regions should be assigned to X- or C-systems. For example, in perhaps the only neuroimaging study to use cognitive load thus far, Foerde et al. (in preparation) found that performance on an implicit learning task correlated with hippocampal activity under single-task conditions but correlated with striatum under dual-task conditions, supporting their model of competition between these two memory systems. From our theoretical perspective, these results demonstrate that the hippocampus is rightly characterized as a C-system structure since cognitive load reduces its efficacy. It also supports the role of striatum in the X-system since cognitive load increased its efficacy (note that though basal ganglia may not be correlated with performance in the single-task condition, this is because it is typically overruled by hippocampal output to behavioral responses; Poldrack and Packard, 2003).

For initial studies using cognitive load, one may consider the following suggestions:

1. Choose a cognitive load manipulation whose neural properties are well-studied and are not expected to overlap with the neural properties of the task of interest.
2. Choose a load manipulation that can be parametrically modulated. This will allow for tracking of the influence of load on ROIs that respond to the task of interest as well as decoupling ROIs activated by increasing load from those activated by the task of interest.
3. Include a functional run to isolate those regions involved in the load manipulation alone.

4. Incorporate a behavioral task in which there is a fair amount of variability in task performance, which can be then correlated with neural data (e.g., Foerde et al., in preparation).

For social perception, some good candidate load manipulations include working memory tasks whose neural properties have been studied and seem to rely mostly on DLPFC, which seems to be involved in working memory and reasoning tasks that are not socially specific (Cabeza and Nyberg, 2000). Moreover, the degree of interference produced by cognitive load can be easily manipulated by parametric variation of working memory load, allowing for more specific behavioral and neural predictions to be drawn. According to our model, though subregions of the C-system serve distinct functions, because of the resource-limited nature of the C-system, the activation of one subregion reduces the resources available to other subregions. This can happen during both simultaneous task performance and sequential task performance. Consequently, performance on a C-system based task drops significantly with the addition of cognitive load even though both tasks may recruit separate neural regions. Hemodynamically, this may manifest itself as a reduction in activity in the dual-task condition, which should correlate with performance decrements.

5.2. Automization of social attribution processes

There is a general tendency to think of tasks involving attribution of mental states such as ToM as involving high level processes because only humans and perhaps a few other animals (Emery and Clayton, 2005; Heyes, 1998) share this capacity. However, there are other unique capacities of human higher cognition that initially depend on the C-system but with practice come to be largely automated and supported by the X-system. For instance, despite the probable dependence of initial self-awareness and the early development of self-knowledge on MPFC in the C-system, there is reason to believe that with repeated self-processing in a particular domain, the X-system can produce self-related outputs. In one study (Lieberman et al., 2004b), we observed that when actors and athletes were asked to judge the self-descriptiveness of traits related to acting and athletics, these individuals activated most of the neural regions in the X-system for the trait words that were in their high experience domain, but not for words in the other domain. Additionally, 10 year olds who are just beginning to establish integrative self-knowledge, show greater MPFC activity than adults when judging the self-descriptiveness of trait words (Pfeifer et al., submitted for publication).

Just as inception of self-awareness and self-knowledge may initially require C-system regions but may become automated over time, ToM processes may follow a similar trajectory. That is, MPFC may be necessary for a person to ever have the ability to think about the psychological states and traits of other people; however, over time, some of the mental operations may become automatized and supported by regions in the X-system instead. Although no research has explicitly examined the automatization of ToM processes and the corresponding shift from C- to X-system structures, some

existing research is suggestive. Wang et al. (submitted for publication) examined the neural correlates of irony comprehension in children and adults. Comprehending irony requires thinking about the communicative intent of the speaker and thus has been used to assess ToM in several studies (Shamay-Tsoory et al., 2005). Wang et al. found that although children produced significantly more MPFC activity during the ironic trials compared to the non-ironic trials, adults produced no MPFC activity during ironic trials compared to non-ironic trials despite the fact that there was no difference in accuracy or reaction times between the children and the adults. The fact that there is a strong negative correlation ($r = -0.71$) between a child's age (range: 9–14) and the magnitude of activity in MPFC also suggests that, with greater experience and development, some of the operations of the MPFC can be automated and performed by other regions of the brain. Indeed, a recent paper by Bird et al. (2004) observed that an adult with extensive damage to MPFC showed intact ToM performance on five different tasks. Thus, MPFC may be activated in adults during some mental acts involving ToM, but it may no longer be necessary for successful performance.

These studies of the self and ToM are suggestive in that MPFC may perform symbolic functions unique to human social cognition that are to some extent slowly learned by X-system structures over time. If this account is true, our model would predict that ToM processes which have been encoded in the X-system will not be impacted by cognitive load. Conversely, those that still rely on MPFC will be affected by cognitive load. Given that over time both MPFC and X-system structures contain the information to solve the problem of mental inference, MPFC may correlate with performance during single-task conditions whereas X-system structures may correlate with performance during dual-task conditions, analogous to the Foerde et al. (in preparation) study. No study to date has explicitly examined which brain regions involved in ToM processes continue to produce the same outputs whether the individual is under cognitive load or not. Studies along these lines should help to better constrain our understanding of the qualitatively different kinds of processes involved in social cognition and the computations performed by the different brain regions involved.

5.3. New directions in fMRI studies of social cognition

To the extent that the descriptions of the X- and C-systems are valid, there are a number of computational and experiential correlates that would be hypothesized for activity in various brain regions as it relates to social cognition. Some of these hypotheses may help to organize existing findings, whereas others may suggest new research that would refine our understanding of social cognition and its neural bases.

Accordingly, we suggest that MPFC should be considered a C-system structure, and LTC a repository of mental state attributions and trait inferences derived from MPFC (Lieberman et al., 2002). Though MPFC is activated without an explicit instruction to do so in situations that allow for mentalizing (German et al., 2004), such as in watching Heider-Simmel displays (Castelli et al., 2000) or social interactions (Iacoboni et al., 2005), the spontaneous engagement of a process does not guarantee that its operations are automatic or possess the

qualities associated with the X-system. As suggested by the developmental neuroimaging data above, MPFC may be involved in more explicit representations of social inference that become automated over time in LTC. Positive relationships between executive function and ToM abilities further support the notion that some aspects of mentalizing involve symbolic computation, and though ToM has been dissociated from more basic working memory tasks (Stone et al., 1998), taken together, this fits the idea that MPFC and DLPFC serve functionally separable but computationally related aspects of the C-system. This hypothesis further predicts that resource-limiting experiments in the form of either cognitive load or self-regulation should particularly impair performance on higher-order mentalizing tasks that rely more heavily on MPFC (e.g., recognition of faux pas).

Activity in MPFC should be predicted by situations in which the computational properties of LTC are insufficient for performance. As a component of the X-system, LTC is expected to behave on the basis of habitually drawn inferences. LTC is not equipped to handle attributions that are novel or at odds with automatic inferences. This includes negation of an attribution and the ability to correct a misguided attribution. In such cases where the habitual or automatic mental attribution drawn is inaccurate or at odds with the behavior of the social target or goals of the perceiver, MPFC may be recruited to provide a more deliberated response. For instance, suspicion, defined as when a perceiver may consider alternative motivations than a surface reason for a target's behavior (Fein, 1996), may preferentially engage MPFC over LTC in social inference. Similarly, interactions between MPFC and LTC may be found as a function of the perceiver's automatic inference of personality and dispositional states of the social target, the perceiver's representations of the target's contextual goals, and finally, the perceiver's own motivational and goal states.

5.4. Relevance to the uniqueness of social cognition as an irreducible domain

One issue of budding interest among neuroscientists interested in social cognition is whether the processes we are studying are fundamentally different from “non-social” cognitive processes. From the perspective of our model, social cognitive systems and cognitive systems are equally divisible along the lines of computational properties of the X- and C-system. Social processes may involve symbolic operations, as required by self-regulation, social reasoning, and early stages or non-habitual aspects of Theory of Mind. Social processes may also be subsymbolic or associative, as with stereotypes, dispositional attributions, and perhaps well-learned Theory of Mind processes. Though there is a fair amount of evidence to entertain the possibility that social and non-social cognitive processes have non-overlapping neural regions, from the perspective of this model, they nonetheless share fundamental computational similarities corresponding to symbolic and associative operations. With respect to C-system processes, social and non-social cognitive processes are hypothesized to be competing over the same resources. As such, the use of, say, dorsolateral prefrontal cortex, considered to be a cognitive area involved in working memory, will reduce the efficacy

of ventrolateral prefrontal cortex, an area involved in self-regulatory behavior. With respect to X-system processes, there should be little to no interference in the operation of one X-system process towards another so long as they are not producing competing interpretations of perceptual information. Hence, the value of this approach is that it provides a computational bridge between seemingly disparate social and non-social cognitive operations that may reveal dependencies that are otherwise unintuitive.

6. Conclusion

Though dual-process models are acknowledged to be a rough cut at the complexity of social behavior, they have provided behavioral researchers with an investigative tool that is generative, explanatory, and simple. In this article, we extended the use of dual-process models to neural systems and in prior work, to phenomenal aspects, too (Lieberman et al., 2002). In total, the model layers neural, computational, and phenomenal aspects of social perception together. This application is nascent in that the kind of evidence that is required to make definitive claims about the automatic or controlled nature of a neural region is largely unavailable. Nonetheless, the available behavioral and neural studies allow for an initial framework to be developed. Ultimately, our division of neural regions into reflective and reflexive systems seeks to provide a useful bridge between a great deal of behavioral work in social cognition that has benefited from a dual-processing approach and the rising interplay between social cognition and neuroscience.

REFERENCES

- Allison, T., Puce, A., McCarthy, G., 2000. Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* 4, 267–278.
- Ambady, N., 1999. The perils of pondering: distraction, deliberation, and thin slice judgments. Unpublished manuscript, Harvard University.
- Ambady, N., Rosenthal, R., 1992. Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis. *Psychol. Bull.* 111, 256–274.
- Aron, A.R., Robbins, T.W., Poldrack, R.A., 2004. Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* 8, 170–177.
- Baddeley, A., 1986. *Working Memory*. Clarendon Press, Oxford, England.
- Bargh, J.A., Chartrand, T.L., 1999. The unbearable automaticity of being. *Am. Psychol.* 54, 462–479.
- Bartels, A., Zeki, S., 2000. The neural basis of romantic love. *NeuroReport* 11, 3829–3834.
- Baumeister, R.F., Bratslavsky, E., Muraven, M., Tice, D.M., 1998. Ego depletion: is the self a limited resource? *J. Pers. Soc. Psychol.* 65, 317–338.
- Baxter, M.G., Murray, E.A., 2002. The amygdala and reward. *Nat. Rev., Neurosci.* 3, 563–573.
- Bechara, A., Damasio, H., Tranel, D., Damasio, A.R., 1997. Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293–1294.
- Bird, C.M., Castelli, F., Malik, O., Frith, U., Husain, M., 2004. The impact of extensive medial frontal lobe damage on 'Theory of Mind' and cognition. *Brain* 127, 914–928.
- Blakemore, S.J., Decety, J., 2001. From perception of action to the understanding of intention. *Nat. Rev., Neurosci.* 2, 561–567.
- Bonda, E., Petrides, M., Ostry, D., Evans, A., 1996. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J. Neurosci.* 16, 3737–3744.
- Botvinick, M.M., Braver, T.D., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.
- Breiter, H.C., Gollub, R.L., Weisskoff, R.M., Kennedy, D.N., Makris, N., Berke, J.D., et al., 1997. Acute effects of cocaine on human brain activity and emotion. *Neuron* 19, 591–611.
- Bruce, C., Desimone, R., Gross, C.G., 1981. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* 46, 369–384.
- Brunet, E., Sarfati, Y., Hardy-Bayle, M., Decety, J., 2000. A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage* 11, 157–166.
- Bush, G., Luu, P., Posner, M.I., 2000. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* 4, 215–222.
- Cabeza, R., Nyberg, L., 2000. Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J. Cogn. Neurosci.* 12, 1–47.
- Carretie, L., Hinojosa, J.A., Mercado, F., Tapia, M., 2005. Cortical response to subjective unconscious danger. *NeuroImage* 24, 615–623.
- Castelli, F., Happe, F., Frith, U., Frith, C., 2000. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage* 12, 314–315.
- Chaiken, S., Trope, Y. (Eds.), 1999. *Dual-Process Theories in Social Psychology*. Guilford Press, New York, NY.
- Cheng, C.M., Chartrand, T.L., 2003. Self-monitoring without awareness: Using mimicry as a nonconscious affiliation strategy. *J. Pers. Soc. Psychol.* 85, 1170–1179.
- Craig, A.D., 2004. Human feelings: why are some more aware than others. *Trends Cogn. Sci.* 8, 239–241.
- Crinion, J.T., Lambon-Ralph, M.A., Warburton, E.A., Howard, D., Wise, R.J., 2003. Temporal lobe regions engaged during normal speech comprehension. *Brain* 126, 1193–1201.
- Cunningham, W.A., Johnson, M.K., Raye, C.L., Chris Gatenby, J., Gore, J.C., Banaji, M.R., 2004. Separable neural components in the processing of black and white faces. *Psychol. Sci.* 15, 806–813.
- Damasio, A.R., 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam, New York.
- Deacon, T.W., 1998. *The Symbolic Species: The Co-Evolution of Language and Brain*. W.W. Norton.
- Deppe, M., Schwindt, W., Kramer, J., Kugel, H., Plassmann, H., Kenning, P., Ringelstein, E.B., 2005. Evidence for a neural correlate of a framing effect: bias-specific activity in the ventromedial prefrontal cortex during credibility judgments. *Brain Res. Bull.* 67, 413–421.
- Eisenberger, N.I., Lieberman, M.D., 2004. "Why it hurts to be left out": the neurocognitive overlap between physical and social pain. *Trends Cogn. Sci.* 8, 294–300.
- Eisenberger, N.A., Lieberman, M.D., Williams, K.D., 2003. Does rejection hurt? An fMRI study of social exclusion. *Science* 302, 290–292.
- Eldridge, L.L., Knowlton, B.J., Furmanski, C.S., Bookheimer, S.Y., Engel, S.A., 2000. Remembering episodes: a selective role for the hippocampus during retrieval. *Nat. Neurosci.* 3, 1149–1152.
- Emery, N.J., Clayton, N.S., 2005. Animal cognition. In: Bolhuis, J., Giraldeau, L.-A. (Eds.), *The Behavior of Animals: Mechanisms, Function, and Evolution*. Blackwell Publishing, Malden, MA, pp. 170–196.
- Fein, S., 1996. Effects of suspicion on attributional thinking and the correspondence bias. *J. Pers. Soc. Psychol.* 70, 1164–1184.
- Fein, S., Spencer, S.J., 1997. Prejudice as self-image maintenance:

- affirming the self through derogating others. *J. Pers. Soc. Psychol.* 73, 31–44.
- Foerde, K.E., Knowlton, B.J., and Poldrack, R.A., in preparation. Secondary task effects on classification learning examined using fMRI.
- Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalising. *Phil. Trans. R. Soc. Lond. B* 358, 459–473.
- Gaffan, D., Murray, E.A., 1990. Amygdalar interaction with the mediodorsal nucleus of the thalamus and the ventromedial prefrontal cortex in stimulus-reward associative learning in the monkey. *J. Neurosci.* 10, 3479–3493.
- Gallagher, H.L., Happe, F., Brunswick, N., Fletcher, P.C., Frith, U., Frith, C.D., 2000. Reading the mind in cartoons and stories: an fMRI study of ‘theory of mind’ in verbal and nonverbal tasks. *Neuropsychologia* 38, 11–21.
- Garrard, P., Hodges, J.R., 2000. Semantic dementia: clinical, radiological and pathological perspectives. *J. Neurol.* 247, 409–422.
- German, T.P., Niehaus, J.L., Roarty, M.P., Giesbrecht, B., Miller, M.B., 2004. Neural correlates of detecting pretense: Automatic engagement of the intentional stance under covert conditions. *J. Cogn. Neurosci.* 16, 1805–1817.
- Gilbert, D.T., 1989. Thinking lightly about others. Automatic components of the social inference process. In: Uleman, J.S., Bargh, J.A. (Eds.), *Unintended Thought*. Guilford, New York, pp. 189–211.
- Gilbert, D.T., Hixon, J.G., 1991. The trouble of thinking: activation and application of stereotypic beliefs. *J. Pers. Soc. Psychol.* 60, 509–517.
- Gilbert, D.T., Pelham, B.W., Krull, D.S., 1988. On cognitive business: when person perceivers meet persons perceived. *J. Pers. Soc. Psychol.* 54, 733–740.
- Goel, V., Grafman, J., Sadato, N., Hallett, M., 1995. Modeling other minds. *NeuroReport* 6, 1741–1746.
- Gray, J.R., Chabris, C.F., Braver, T.S., 2003. Neural mechanisms of general fluid intelligence. *Nat. Neurosci.* 6, 316–322.
- Greenwald, A.G., McGhee, D.E., Schwartz, J.L.K., 1998. Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480.
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., Blake, R., 2000. Brain areas involved in the perception of biological motion. *J. Cogn. Neurosci.* 12, 711–720.
- Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E., 2001. Medial prefrontal cortex and self-referential mental activity: relation of a default mode of brain function. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4259–4264.
- Hariri, A.R., Mattay, V.S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., Egan, M.F., Weinberger, D.R., 2002. Serotonin transporter genetic variation and the response of the human amygdala. *Science* 297, 400–403.
- Hart, A.J., Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H., Rauch, S.L., 2000. Differential response in the human amygdala to racial outgroup vs. ingroup face stimuli. *Brain Imaging* 11, 2351–2355.
- Heyes, C.M., 1998. Theory of mind in nonhuman primates. *Behav. Brain Sci.* 21, 101–148.
- Hummel, J.E., 1999. Binding problem. In: Wilson, R.A., Keil, F.C. (Eds.), *The MIT Encyclopedia of the Cognitive Sciences*. MIT Press, Cambridge, MA, pp. 85–86.
- Hummel, J.E., Holyoak, K.J., 2003. A symbolic-connectionist theory of relational inference and generalization. *Psychol. Rev.* 110, 220–264.
- Iacoboni, M., Koski, L.M., Brass, M., Bekkering, H., Woods, R.P., Dubeau, M.-C., Mazziotta, J.C., Rizzolatti, G., 2001. Reafferent copies of imitated actions in the right superior temporal cortex. *Proc. Natl. Acad. Sci.* 98, 13995–13999.
- Iacoboni, M., Lieberman, M.D., Knowlton, B.J., Molnar-Szakacs, I., Moritz, M., Throop, C.J., Fiske, A.P., 2005. Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *NeuroImage* 21, 1167–1173.
- Ichheiser, G., 1943. Misinterpretations of personality in everyday life and psychologist’s frame of reference. *Character Pers.* 12, 145–160.
- Jellema, T., Baker, C.I., Wicker, B., Perrett, D.I., 2000. Neural representation for the perception of the intentionality of actions. *Brain Cogn.* 44, 280–302.
- Jellema, T., Maassen, G., Perrett, D.I., 2004. Single cell integration of animate form, motion, and location in the superior temporal cortex of the macaque monkey. *Cereb. Cortex* 14, 781–790.
- Johansson, G., 1973. Visual perception of biological motion and a mode for its analysis. *Percept. Psychophys.* 14, 201–211.
- Jones, E.E., Harris, V.A., 1967. The attribution of attitudes. *J. Exp. Soc. Psychol.* 3, 1–24.
- Kahneman, D., 1973. *Attention and Effort*. Prentice-Hall, Eaglewood Cliffs, NJ.
- Kahneman, D., 2003. A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol.* 58, 697–720.
- Knowlton, B.J., Mangels, J.A., Squire, L.R., 1996. A neostriatal habit learning system in humans. *Science* 273, 1399–1402.
- Knutson, B., Adams, C.M., Fong, G.W., Hommer, D., 2002. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.
- Knutson, B., Bjork, J.M., Fong, G.W., Hommer, D., Mattay, V.S., Weinberger, D.R., 2004. Amphetamine modulates human incentive processing. *Neuron* 43, 261–269.
- Kroger, J.K., Saab, F.W., Fales, C.L., Bookheimer, S.Y., Cohen, M.S., Holyoak, K.J., 2002. Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: a parametric study of relational complexity. *Cereb. Cortex* 12, 477–485.
- Langer, E.J., 1989. *Mindfulness*. Addison Wesley, Reading, MA.
- Lau, H.C., Rogers, R.D., Ramnani, N., Passingham, R.E., 2004. Willed action and attention to the selection of action. *NeuroImage* 21, 1407–1415.
- LeDoux, J.E., 1996. *The Emotional Brain*. Simon and Schuster, New York.
- LeDoux, J., 2003. The emotional brain, fear, and the amygdala. *Cell. Mol. Neurobiol.* 23, 727–738.
- Liddell, B.J., Brown, K.J., Kemp, A.H., Barton, M.J., Das, P., Peduto, A., Gordon, E., Williams, L.M., 2005. A direct brainstem-amygdala-cortical ‘alarm’ system for subliminal signals of fear. *NeuroImage* 24, 235–243.
- Lieberman, M.D., 2000. Intuition: a social cognitive neuroscience approach. *Psychol. Bull.* 126, 109–137.
- Lieberman, M.D., in press. The X- and C-systems: the neural basis of automatic controlled social cognition. In: Harmon-Jones, E., Winkelman P. (Eds.), *Fundamentals of Social Neuroscience*. New York: Guilford.
- Lieberman, M.D., Eisenberger, N.I., 2005. A pain by any other name (rejection, exclusion, ostracism), still hurts the same: the role of dorsal anterior cingulate in social and physical pain. In: Cacioppo, J.T., Visser, P., Pickett, C. (Eds.), *Social Neuroscience: People Thinking About People*. MIT Press, Cambridge, MA, pp. 167–187.
- Lieberman, M.D., Pfeifer, J.H., 2005. The self and social perception: three kinds of questions in social cognitive neuroscience. In: Easton, A., Emery, N. (Eds.), *Cognitive Neuroscience of Emotional and Social Behavior*. Psychology Press, Philadelphia.
- Lieberman, M.D., Gaunt, R., Gilbert, D.T., Trope, Y., 2002. Reflection and reflexion: a social cognitive neuroscience approach to attributional inference. *Adv. Exp. Soc. Psychol.* 34, 199–249.
- Lieberman, M.D., Chang, G.Y., Chiao, J., Bookheimer, S.Y., Knowlton, B.J., 2004a. An event-related fMRI study of artificial grammar learning in a balanced chunk strength design. *J. Cogn. Neurosci.* 16, 427–438.
- Lieberman, M.D., Jarcho, J.M., Satpute, A.B., 2004b. Evidence-based

- and intuition-based self-knowledge: an fMRI study. *J. Pers. Soc. Psychol.* 87, 421–435.
- Lowery, B.S., Hardin, C.D., Sinclair, S., 2001. Social influence effects on automatic racial prejudice. *J. Pers. Soc. Psychol.* 81, 842–855.
- Macrae, C.N., Bodenhausen, G.V., 2000. Social cognition: thinking categorically about others. *Annu. Rev. Psychol.* 51, 93–120.
- McClelland, J.L., McNaughton, B.L., O'Reilly, R.C., 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457.
- McKiernan, K.A., Kaufman, J.N., Kucera-Thompson, J., Binder, J.R., 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *J. Cogn. Neurosci.* 15, 394–408.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Milne, E., Grafman, J., 2001. Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *J. Neurosci.* 21 (RC150), 151–156.
- Mitchell, J.P., Heatherton, T.F., Macrae, C.N., 2002. Distinct neural systems subserve person and object knowledge. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15238–15243.
- Mitchell, J.P., Nosaek, B.A., Banaji, M.R., 2003. Contextual variations in implicit evaluation. *J. Exp. Psychol. Gen.* 132, 455–469.
- Morris, J.S., Friston, K.J., Büchel, C., Frith, C.D., Young, A.W., Calder, A.J., Dolan, R.J., 1998. A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain* 121, 47–57.
- Morris, J.S., Ohman, A., Dolan, R.J., 1999. A subcortical pathway to the right amygdala mediating “unseen” fear. *Proc. Natl. Acad. Sci. U. S. A.* 16, 1680–1685.
- Murphy, G.L., 2002. *The Big Book of Concepts*. MIT Press.
- Mummery, C.J., Shallice, T., Price, C.J., 1999. Dual-process model in semantic priming: a functional imaging perspective. *NeuroImage* 9, 516–525.
- Mummery, C.J., Patterson, K., Price, C.J., Ashburner, J., Frackowiak, R.S.J., Hodges, J.R., 2000. A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. *Ann. Neurol.* 47, 36–45.
- Nimchinsky, E.A., Gilissen, E., Allman, J.M., Perl, D.P., Erwin, J.M., Hof, P.R., 1999. A neuronal morphological type unique to humans and great apes. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5268–5273.
- Noveck, I.A., Goel, V., Smith, K.W., 2004. The neural basis of conditional reasoning with arbitrary content. *Cortex* 40, 613–622.
- Ochsner, K.N., Gross, J.J., 2005. The cognitive control of emotion. *Trends Cogn. Sci.* 9, 242–249.
- Öhman, A., 2005. The role of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology* 30, 953–958.
- Öngür, D., Price, J.L., 2001. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys, and humans. *Cereb. Cortex* 10, 206–219.
- Oram, M.W., Perrett, D.I., 1996. Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *J. Neurophysiol.* 76, 109–129.
- Pasley, B.N., Mayes, L.C., Schultz, R.T., 2004. Subcortical discrimination of unperceived objects during binocular rivalry. *Neuron* 42, 163–172.
- Payne, B.K., 2001. Prejudice and perception: the role of automatic and controlled processes in misperceiving a weapon. *J. Pers. Soc. Psychol.* 81, 181–192.
- Pelphrey, K.A., Mitchell, T.V., McKeown, M.J., Goldstein, J., Allison, T., McCarthy, G., 2003. Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. *J. Neurosci.* 23, 6819–6825.
- Pelphrey, K.A., Viola, R.J., McCarthy, G., 2004. When strangers pass: processing of mutual and averted social gaze in the superior temporal sulcus. *Psychol. Sci.* 15, 598–603.
- Perrett, D.I., Rolls, E.T., Caan, W., 1982. Visual neurons responsive to faces in the monkey temporal cortex. *Exp. Brain Res.* 47, 329–342.
- Perrett, D.I., Smith, P.A.J., Mistlin, A.J., Chitty, A.J., Head, A.S., Potter, D.D., Broennimann, R., Milner, A.D., Jeeves, M.A., 1985. Visual analysis of body movements by neurons in the temporal cortex of the macaque monkey: a preliminary report. *Behav. Brain Res.* 16, 153–170.
- Perrett, D., Harries, M., Bevan, R., Thomas, S., Benson, P., Mistlin, A., Chitty, A., Hietanen, J., Ortega, J., 1989. Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.* 146, 87–114.
- Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L.G., 2002. Neural processing of emotional faces requires attention. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11458–11463.
- Pfeifer, Dappretto, and Lieberman, submitted for publication.
- Ploghaus, A., Becerra, L., Borras, C., Borsook, D., 2003. Neural circuitry underlying pain modulation: expectation, hypnosis, placebo. *Trends Cogn. Sci.* 7, 197–200.
- Poldrack, R.A., in press. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.*
- Poldrack, R.A., Packard, M.G., 2003. Competition between memory systems: converging evidence from animal and human studies. *Neuropsychologia* 41, 245–251.
- Poldrack, R.A., Prabakaran, V., Seger, C., Gabrieli, J.D.E., 1999. Striatal activation during cognitive skill learning. *Neuropsychology* 13, 564–574.
- Prabhakaran, V., Smith, J.A., Desmond, J.E., Glover, G.H., Gabrieli, J.D., 1997. Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cogn. Psychol.* 33, 43–63.
- Prabhakaran, V., Rypma, B., Gabrieli, J.D.E., 2001. Neocortical substrates of mathematical reasoning: a functional magnetic resonance imaging of neocortical activation during performance of the necessary arithmetic operations test. *Neuropsychology* 15, 115–127.
- Puce, A., Allison, T., Bentin, S., Gore, J.C., McCarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. *J. Neurosci.* 18, 2188–2199.
- Rainville, P., Duncan, G.H., Price, D.D., Carrier, B., Bushnell, M.C., 1997. Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science* 15, 968–971.
- Rissman, J., Blumstein, S.E., 2003. An event-related fMRI investigation of implicit semantic priming. *J. Cogn. Neurosci.* 15, 1160–1175.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., Fazio, F., 1996. Localization of grasp representations in humans by PET: 1. Observation versus execution. *Exp. Brain Res.* 111, 246–252.
- Ross, L., 1977. The intuitive psychologist and his shortcomings. In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, vol. 10. Academic Press, New York, pp. 173–220.
- Rossell, S.L., Bullmore, E.T., Williams, S.C.R., David, A.S., 2001. Brain activation during automatic and controlled processing of semantic relations: a priming experiment using lexical decision. *Neuropsychologia* 39, 1167–1176.
- Ruby, P., Decety, J., 2001. Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat. Neurosci.* 4, 546–550.
- Satpute, A.B., Sellner, D., Waldman, M.D., Tabibnia, G., Holyoak, K.J., Lieberman, M.D., 2005. An fMRI study of causal judgments. *Eur. J. Neurosci.* 22, 1233–1238.
- Saxe, R., Kanwisher, N., 2003. People thinking about thinking

- people: the role of the temporo-parietal junction in “theory of mind”. *NeuroImage* 19, 1835–1842.
- Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D.I., Kanwisher, N., 2004. A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia* 42, 1435–1446.
- Scholl, B.J., Tremoulet, P.D., 2000. Perceptual causality and animacy. *Trends Cogn. Sci.* 4, 299–309.
- Semendeferi, K., Schleichner, A., Zilles, K., Armstrong, E., Van Hoesen, G.W., 2001. Prefrontal cortex in humans and apes: A comparative study of area 10. *American Journal of Physical Anthropology* 114, 224–241.
- Shamay-Tsoory, S.G., Tomer, R., Aharon-Peretz, J., 2005. The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology* 19, 288–300.
- Shultz, J., Friston, K.J., O’Doherty, J., Wolpert, D.M., Frith, C.D., 2005. Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron* 45, 625–635.
- Slooman, S.A., 1996. The empirical case for two systems of reasoning. *Psychol. Bull.* 119, 3–22.
- Smith, W., 1945. The functional significance of the rostral cingulate cortex as revealed by its responses to electrical excitation. *J. Neurophysiol.* 8, 241–255.
- Smith, W., DeCoster, J., 1999. Associative and rule-based processing: A connectionist interpretation of dual process models. In: Chaiken, S., Trope, Y. (Eds.), *Dual-Process Theories in Social Psychology*. Guilford Press, New York, pp. 323–338.
- Squire, L.R., Zola, S.M., 1996. Structure and function of declarative and nondeclarative memory systems. *Proc. Natl. Acad. Sci.* 93, 13515–13522.
- Stone, V.E., Baron-Cohen, S., Knight, R.T., 1998. Frontal lobe contributions to theory of mind. *J. Cogn. Neurosci.* 10, 640–656.
- Strack, F., Deutsch, R., 2004. Reflective and impulsive determinants of social behavior. *Pers. Soc. Psychol. Rev.* 8, 220–247.
- Swanson, L.W., Petrovich, G.D., 1998. What is the amygdala? *Trends Neurosci.* 21, 323–331.
- Tabibnia, G., Satpute, A.B., and Lieberman, M.D., submitted for publication. The neural basis of fairness preference.
- Thompson, R.K.R., Oden, D.L., 2000. Categorical perception and conceptual judgments by nonhuman primates: the paleological monkey and the analogical ape. *Cogn. Sci.* 24, 363–396.
- Trope, Y., Gaunt, R., 1999. A dual-process model of overconfident attributional inferences. In: Chaiken, S., Trope, Y. (Eds.), *Dual-Process Theories in Social Psychology*. Guilford Press, New York, pp. 161–178.
- Trope, Y., Gaunt, R., 2000. Processing alternative explanations of behavior: correction or integration? *J. Pers. Soc. Psychol.* 79, 344–354.
- Wager, T.D., Phan, K.L., Liberzon, I., Taylor, S.F., 2003. Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *NeuroImage* 19, 513–531.
- Waldmann, M.R., Walker, J.M., 2005. Competence and performance in causal learning. *Learn. Behav.* 33, 211–219.
- Wang, Lee, Sigman, Dappretto, submitted for publication.
- 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. *J. Neurosci.* 18, 411–418.
- Zalla, T., Koechlin, E., Pietrini, P., Basso, G., Aquino, P., Sirigu, A., Grafman, J., 2000. Differential amygdala responses to winning and losing: a functional magnetic resonance imaging study in humans. *Eur. J. Neurosci.* 12, 1764–1770.