3

Reflexive and Reflective Judgment Processes

A Social Cognitive Neuroscience Approach

Matthew Lieberman

CONTENTS

Introduction 44
Reflection-Reflexion Model of Judgment 46
How the X- and C-Systems Make Decisions 51
Reflective Processes Disrupt Reflexive Processes 55
Individual Differences in Reflective Judgment Processes 58
Conclusion 60

INTRODUCTION

One of the driving forces in social cognition has been the goal of understanding the mental mechanisms that can produce the large array of paradoxical findings that have excited social psychology students for half a century. From persuasion to person perception, decision making to dissonance, and judgment to job descrimination, the distinction between automatic and controlled processes has provided tremendous empirical leverage in the crusade to divide and conquer mental phenomena (in this volume, see Brewer; Chartrand & Jefferis; Galinsky, Martorana, & Ku; Johnston & Miles; cf. Kruglanski et al.). Controlled processes (sometimes referred to as *explicit*, *conscious*, or *rational processes*) typically involve some combination of effort, intention, and awareness, tend to interfere with one another, and are usually experienced as self-generated thoughts. Automatic processes (sometimes referred to as *implicit*, *nonconscious*, or *intuitive processes*) typically lack effort, intention, or awareness, tend not to interfere with one another, and are usually experienced as perceptions or feelings. Deciding how to budget one's finances effectively to cover

Please address correspondence to Matthew Lieberman, Department of Psychology, Franz Hall, University of California, Los Angeles, Los Angeles, CA 90095-1563; e-mail: lieber@ucla.edu

necessary expenses in both the short and the long term is a controlled process that will likely require effort, intention, and awareness. Deciding to see the new movie starring one's favorite actor is more automatic, with the decision simply appearing in the form of an impulse.

Although progress has been made in terms of determining many of the behavioral and cognitive consequences of automatic and controlled processing, the theoretical utility of this progress has been hampered because automaticity and control refer to the processing phenomenology rather than to the qualitatively distinct representations that are utilized in the two kinds of processing. There is a tendency for us to believe that the same kinds of underlying symbolic representations are used regardless of the automaticity of their use in terms of effort, intention, and awareness (Smith, 1996). This view would suggest that as the processing of certain representations becomes increasingly automatic, the internal structure of the representations remains qualitatively the same but the individual's ability to initiate, prevent, and consciously guide the process changes. There is a increasing volume of work in the cognitive neurosciences that conflicts with this view, instead offering up distinct neural bases for automatic and controlled processes.

A second limitation of current models is that automaticity and control are often viewed as the anchors for two ends of a continuum. Controlled processes are those that involve effort, awareness, and intention. But what are automatic processes? It seems somewhat incomplete to answer that automatic processes are the ones that lack the qualities that controlled processes possess. Rather than defining automatic processes negatively, in terms of which components of controlled processes are lacking, a cognitive neuroscience approach allows us to examine automatic and controlled processes positively in terms of the qualitatively distinct computational properties that emerge from the neurophysiology and connectivity to other neural systems.

Additionally, traditional measures of automaticity and control cannot assess interactions between automatic and controlled processes. No matter how automatic and controlled processes are interacting, the resulting judgment or behavioral response will reflect a unified product that will look more or less automatic overall (Cacioppo & Gardner, 1999). Imagine a process that is two parts controlled and two parts automatic. A shift in processing that adds one part controlled process will be difficult to discriminate from a shift that instead subtracts one part automatic process if the test of automaticity measures the linear summation or relative contributions of the automatic and controlled components. Using neuroimaging techniques, the independent changes in automatic and controlled processes can be assessed prior to the output stage of behavior.

The goal of this chapter is to provide a simple model of the distinct neural correlates of automatic and controlled processes as they relate to social and affective judgment. Once presented, a new way to understand the negative consequences of introspection in judgment will be presented, complementing existing models (Wilson & Schooler, 1991). Finally, a connection will be made between the neurocognitive systems involved in automatic and controlled judgment processes and the personality-related reactivity of these systems. Because the efficiency and sensitivity of these brain structures vary, these individual differences should be associated with different judgment and decision-making styles. The field of social cognitive neuroscience is in its infancy (Lieberman, 2000a; Ochsner & Lieberman, 2001), and thus a disclaimer is necessary: Some of the arguments made in this chapter are bound to appear simplistic or flat-out mistaken in the clear light of history. Nevertheless, some theorizing and speculating are in order if we are ever to pull the edges of the separate disciplines close enough to overlap with one another. This chapter will, it is hoped, serve as an invitation to social psychologists to take a cognitive neuroscientist to lunch and start exploring the ways that collaborations bringing the disciplines closer can lead to progress on both new and old problems in judgment and decision making.

REFLECTION-REFLEXION MODEL OF JUDGMENT

At every turn and at each moment in our daily lives, we are making countless implicit judgments and decisions that allow us seamlessly to make sense of and navigate our social world. We intuitively make sense of the nonverbal messages in the environment and often reciprocate appropriately without any effort (Ambady & Rosenthal, 1993; Chartrand & Jefferis, this volume; Chen & Bargh, 1997; Word, Zanna & Cooper, 1974); automatically judge objects as more likeable based on previous exposure or their position in a display (Nisbett & Wilson, 1977; Zajonc, 1968); spontaneously make sense of behavior in terms of intentions and traits (Gilbert, 1989; Heider & Simmel, 1944; Winter & Uleman, 1984); and decide whether to help strangers based merely on the syntax of the request, without careful consideration (Langer, Blank, & Chanowitz, 1978).

As long as the judgments to be made address familiar stimuli that are functioning in the way we are accustomed, our judgments can usually proceed automatically without ever becoming a focus of attention. However, when our expectations are violated, doubt and ambiguity ensue, followed quickly by more explicit decision-making processes (Dewey, 1910; Whitehead, 1911). For instance, when we turn a doorknob to open a door, we are making a number of assumptions about the nature of the doorknob in terms of structure and function. As long as the doorknob works as expected, these assumptions remain tacit, allowing us to focus on other thoughts, and most of the time we are not even reflectively aware of doorknobs at all (i.e., aware of the doorknob as "a doorknob"). Doorknobs recruit controlled processing only when they cease to function as doorknobs because broken doorknobs are anomalous and cannot be assimilated by our more automatic processes (Dreyfus, 1991; Heidegger, 1927). A broken doorknob that stands between us and where we want to be creates an expectancy violation that requires effort, attention, and reasoning to resolve. In these cases, controlled decision-making processes occur primarily when automatic processes have failed to achieve our goals.

Previously, my collaborators and I have termed those processes that spontaneously link our goals to behavior *reflexive* (Lieberman, Gaunt, Gilbert, & Trope, 2002b). When reflexive processes fail, *reflective* processes are recruited to deal with circumstances that are exceptions to our implicit expectations. Although these terms may seem like yet another name for the same old dual-process dichotomy, reflexive and reflective processes are defined functionally and neurally, not just in terms of resistance to cognitive load and other standard measures of automaticity. Dual-process theories typically propose that the occurrence of controlled processing depends on the availability of cognitive resources and the motivation to be accurate. The reflection-reflexion model, taking its cue from cybernetic control models (Carver & Scheier, 1981; Miller, Galanter, & Pribram, 1960; Wiener, 1948), suggests that in addition to motivation and resources, the occurrence of reflective processes is determined by the success or failure of reflexive processes. In other words, reflective processes can be functionally defined as those designed for and recruited to handle situations that prove too difficult for reflexive processes. Additionally, the word *control* has come under increasing scrutiny for its association with free will and the implied homunculus that creates a philosophical regress (Dennett, 1984; Wegner & Bargh, 1998). *Reflexive* and *reflective* indicate the phenomenological experience associated with these processes, without sneaking in a reified self through the backdoor (James, 1913).

The X-System

The neural correlates associated with these two types of processes are the X-system (for the *X* in reflexive) and the C-system (for the *C* in reflective). Functionally, the X-system is responsible for linking affect and social meaning to currently represented stimuli, regardless of whether those stimuli are activated bottom-up as a result of ongoing perception or top-down as the contents of working memory in the form of goals, explicit thoughts, or retrieved memories. These links usually reflect conditioning between the various features of a stimulus or between the stimulus and the outcomes for which the stimulus is a cue. The former (*stimulus–stimulus* associations) might include implicit personality theories, stereotypes, and other forms of categorical cognition in which various characteristics, traits, or attributes are believed to co-occur. The latter (*stimulus–outcome* associations) generally refer to affective processes, in which one cue (e.g., an angry expression) indicates that ones' goals are about to be advanced or thwarted. These affective processes have long been thought to prepare the individual to act on the basis of affective judgment, and thus these processes are assumed to link directly to motor systems in the brain (Frijda, 1986; LeDoux, 1996; Rolls, 1999).

The three neural structures associated with the X-system are the amygdala, basal ganglia, and lateral temporal cortex. There is strong evidence that the amygdala learns about and spontaneously reacts to threat cues as varied as snakes, fear expressions, and out-group race faces (LeDoux, 1996; Phelps et al., 2000); responds to even subliminal presentations of these threat cues (Morris, Ohman, & Dolan, 1999; Whalen et al., 1998); and has direct projections to various motor systems that mediate the *fight-or-flight* response, implicating the amygdala as part of an automatic avoidance system. Alternatively, the basal ganglia seem to serve as part of an automatic approach system, responding to various predictors of reward (Depue & Collins, 1999; Lieberman, 2000a; Schultz, 1998). Neuroimaging studies of implicit learning have shown that the basal ganglia is activated by sequences of cues that predict a desired outcome (Berns, Cohen, & Mintun, 1997; Lieberman, Chang, Chiao, Bookheimer, & Knowlton, 2002a) and in response to cues that elicit positive affect ranging from pictures of loved ones (Bartels & Zeki, 2000) to cocaine administration (Breiter, Gollub, & Weisskoff, 1997). Like the amgydala, the basal ganglia also have numerous projections that can initiate behavior; indeed, the first symptoms of basal ganglia–related neurodegenerative diseases like Parkinson's disease and Huntington's disease are impairments in habitual action patterns (Grant & Adams, 1996). There is evidence to suggest that the basal ganglia and amygdala each participate to some degree in reflexive responses to positive and negative cues; however, they do appear to be selectively more involved in one valenced affect than the other (Tabibnia, 2002).

Lateral temporal cortex is involved in recognizing the identity, attributes, and behavior of social (and nonsocial) objects. The lower (or *ventral*) part of temporal cortex is part of the "what" system in the visual processing stream (Mishkin, Ungeleider, & Macko, 1983). The contiguous regions that make up the ventral visual stream make progressively more abstract inferences about the identity and category membership of the underlying stimulus. In early stages (toward the back of the brain), visual attributes are constructed from lines, colors, and textures. In later stages (toward the front of the temporal lobes), the overall identity and category memberships are activated such that attributes that are missing or occluded are activated in addition to those that are strictly visible. Thus, just seeing part of a familiar face or the color of a stranger's skin will lead people to fill in all sorts of other characteristics associated with individual or group identity. This ability to "go beyond the information given" is a major component of most theories hoping to explain the strengths and weaknesses of human social cognition (Bruner, 1957; Fiske & Taylor, 1991; Griffin & Ross, 1991).

The upper part of lateral temporal cortex, referred to as *superior temporal sulcus*, is specialized for behavior identification (Allison, Puce, & McCarthy, 2000; Decety & Grezes, 1999; Haxby, Hoffman, & Gobbin, 2000; Lieberman et al., 2002b). This region is largely involved in extracting the intentions behind particular behaviors and has been theorized to be critical to the *Theory of Mind* lacking in children with autism (Baron-Cohen, 1995). Neurons in this region fire in conjunction with very different behaviors as long as the underlying intention associated with each behavior is the same (Jellema & Perrett, 2001). Moreover, accidental behaviors such as tripping and dropping an object do not activate these neurons (Perrett, Jellema, Frigerio, & Burt, 2001). Ventral temporal cortex and superior temporal cortex "talk" with one another to allow inferences about individual and group traits based on observed behaviors and inferences about the implications of behaviors based on who performed them.

To summarize, the X-system is composed of the amygdala, basal ganglia, and lateral temporal cortex to form a very efficient knowledge base about the social and affective characteristics of social phenomena that often sets into motion behaviors based on extensive learning histories that have accumulated slowly over time. The X-system contains our implicit theories and expectations that allow us to interface smoothly with the world, seamlessly promoting our goals and avoiding our foes.

The C-System

Sometimes our existing expectations in the X-system fail us. The failure can occur in one of two ways, but the result is the same in both cases: Another mechanism besides the X-system is needed to guide behavior. Sometimes the task or stimulus is novel, and consequently the X-system has no preexisting well-learned representations that can assimilate the incoming information. At other times the task or stimulus is familiar, but one's goal or the constraints of the current context render the X-system's habitual response inappropriate. For instance, egalitarian individuals may have negative racial stereotypes activated in the X-system in response to a member of that race. Given egalitarian motives, the X-system's response to the target is undesirable, leading to the involvement of controlled processing in order to override or suppress the X-system's response (Galinsky, Martorana & Ku, this volume; Monteith, 1993; cf. Moskowitz, Gollwitzer, Wasel, & Schaal, 1999). In each case, the X-system's generalities are not prepared to deal with the specific situation at hand.

In contrast to the X-system's efficiency with social phenomena that conform to its generalities, the C-system is critical for handling the exceptions to the rules (McClelland, McNaughton, & O'Reilly, 1995). The C-system is composed of three neurocognitive mechanisms that work closely together: anterior cingulate cortex, prefrontal cortex, and the medial temporal lobe (including the hippocampus). The first two regions are responsible for detecting the need for top-down control and for implementing control, respectively. The medial temporal lobe stores information about past episodes to the extent that they required controlled processing (Brewer, Zhao, Desmond, Glover, & Gabrieli, 1998; Craik & Tulving, 1975; Otten, Henson, & Rugg, 2001; Wagner et al., 1998), presumably so that this information can facilitate processing the next time the special circumstances arise. The details of how these structures perform their functions have been described at length elsewhere (Lieberman et al., 2002b). Consequently, I will limit myself to some of the conclusions posited by the reflection-reflexion model.

First, the anterior cingulate cortex is sensitive to many forms of conflict and error in the X-system that generally indicate that the X-system is unable to advance one's current goals (Botnivick, Braver, Barch, Carter, & Cohen, 2000) and thus serves as an *alarm system* that alerts prefrontal cortex that control is needed. The anterior cingulate is sensitive to a variety of conflict and error signals including pain (Baciu et al., 1999; Ladabaum, Minoshima, & Owyang, 2000; Rainville, Duncan, Price, Carrier, & Bushnell, 1997), distress vocalizations (Lorberbaum et al., 1999; see Shaver & Mikulincer, this volume, for relevant attachment work), cognitive errors (Bush, Luu, & Posner, 2000), and tasks that require the overriding of a habitual response based on current goals (Barch, Braver, Sabb, & Noll, 2000; Carter et al., 2000). Although the anterior cingulate is clearly involved in recruiting controlled processes, it would be inaccurate to rigidly classify the computations of the anterior cingulate as controlled (or automatic). Instead, the anterior cingulate serves as an example of why the existing automaticity–control dichotomy is in need of revision. On the one hand, like a typical alarm clock, the anterior cingulate is a "set it and forget it" alarm. The anterior cingulate is sensitive to numerous kinds of errors simultaneously without requiring conscious intent or effort (Benedict et al., 2002). On the other hand, the sensitivity of the cingulate can be weighted more heavily toward some discrepancies than others based on current goals (Carter et al., 1998, 2000; Kropotov, Crawford, & Polyakov, 1997; Sawamoto et al., 2000), and its sensitivity can be undermined by cognitive load (Frankenstein, Richter, McIntrye, & Remy, 2001; Petrovic, Petersson, Ghatan, Stone-Elander, & Ingvar, 2000). This last finding, that the sensitivity of the anterior cingulate can be blunted by cognitive load, is of particular interest because it suggests that the standard account of cognitive load effects or controlled processing is only half of the story. Typically, cognitive load is thought to undermine the use of controlled processing resources, but the reflection-reflexion model suggests that much of the time cognitive load has its impact by making the anterior cingulate less likely to sound the alarm that controlled processing resources are needed. In other words, under cognitive load, maladaptive decisions constructed in the X-system will often go unchecked.

When the anterior cingulate does sound the alarm, it is usually for the benefit of prefrontal cortex, which is activated during numerous processes that are associated with controlled processing including working memory load (Braver et al., 1997), propositional reasoning (Goel & Dolan, 2000; Waltz et al., 1999), causal inference (Lieberman et al., 2002c), linguistic constructions (Bookheimer, in press), goal generation (Milner, 1963), and hypothesis formation (Christoff & Gabrieli, 2000). Cohen and colleagues have posited that many of these prefrontal processes function to override the X-system or bias it to function temporarily in more contextually appropriate ways (MacDonald, Cohen, Stenger, & Carter, 2000; Miller & Cohen, 2001).

In summary, the C-system is designed to sense the floundering of the X-system and to intervene when appropriate. Of course, in the modern world the C-system is activated much of the time, regardless of the X-system's preparedness. That is to say, although the C-system may have evolved to come to the X-system's rescue, the C-system has clearly taken on a life of its own in a world in which nearly every external cue is designed to evoke some degree of C-system processing (Deacon, 1997). Moreover, the rationalist tradition of Western society looks down upon the use of intuition (Bruner, 1957; Haidt, 2001; Hogarth, 2001; Lieberman, 2000a), and consequently, people may tend to rely on C-system processing even when X-system processing would suffice.

HOW THE X- AND C-SYSTEMS MAKE DECISIONS

To this point, it is clear that the reflection-reflexion model posits that the brain is designed to give the X-system first crack at making most judgments and decisions. Quite literally, the structures of the X-system receive incoming information before the structures of the C-system (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001; Iwata, LeDoux, Meeley, Arneric, & Reis, 1986; LeDoux, Ruggiero, & Reis, 1986). If the X-system is unable to generate a useful solution, the C-system is activated so that its arsenal of decision-making tools can be used. It is still unclear from this discussion how decision making actually differs in the two systems. One major difference that has long been accepted is that X-system processes operate in parallel, whereas C-system processes are exclusionary, operating one at a time (McClelland, Rumelhart, & Hinton, 1988; Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Wegner & Bargh, 1998). Although this difference has usually been thought of in terms of processing speed, it also has enormous consequences for the nature of processing itself (Sloman, 1996; Smith &

DeCoster, 1999). Because X-system processes operate in parallel, and because many of the neurons in the X-system project to and receive projections from the same neurons, neurons are simultaneously influenced by the neurons they are influencing. In other words, neurons in the X-system are highly interdependent, acting much like people do according to balance theory (Heider, 1958; Read, Vanman, & Miller, 1997). This process of parallel constraint satisfaction in the X-system, described in detail elsewhere (Kunda & Thagard, 1996; Lieberman et al., 2002b; Read et al., 1997; Shultz & Lepper, 1995; Spellman & Holyoak, 1992), creates a pattern-matching function that yields both judgment shifts related to cognitive dissonance reduction and similarity-based decision making that is contextually inappropriate (Donovan & Epstein, 1997; Sloman, 1996).

Reflexive Judgment Shifts in Dissonance Reduction

Shultz and Lepper (1995) gave a radical reinterpretation of standard cognitive dissonance effects. Their computational model suggests that cognitions are interdependently represented; thus, changes in any one cognition in a network of cognitions or the introduction of a new cognition can have a ripple effect on others. Importantly, these ripple effects occur simply because of the distributed representational structure of the mind, rather than as a result of conscious awareness of a conflict between cognitions combined with effortful attempts to make the cognitions more consonant with one another, as other formulations have suggested (Brehm & Cohen, 1962; Elliott & Devine, 1994; Festinger, 1964).

Two recent studies (Lieberman, Ochsner, Gilbert, & Schacter, 2001b) have provided empirical support for the notion that this sort of judgment shift does not depend on C-system processes. In one study, patients with anterograde amnesia performed a variation of Brehm's (1956) free-choice paradigm. In this task, individuals rank 15 postcard-sized art prints and then are asked to choose whether they would prefer to have full-sized reproductions of the pair of prints they ranked 4th and 10th or the pair of prints they ranked 6th and 12th (participants are shown the pairs without any reminder of their previous rankings). In either case, participants are choosing one print they do not like very much (the 10th or 12th) and rejecting one print that they presumably do like (the 4th or 6th). At the end of the experiment, participants are asked to rerank the 15 prints based on how they are currently feeling. Normal healthy participants tend to rank the chosen prints higher and the rejected prints lower than they did originally. By "spreading the difference" between the accepted and rejected prints, the earlier choice becomes more consonant. Amnesics, in contrast to healthy adults, cannot form new memories. Thus, when amnesics are distracted as soon as they choose between the 4/10 and 6/12 pairs, they cannot explicitly recall that they have just made a dissonance-inducing

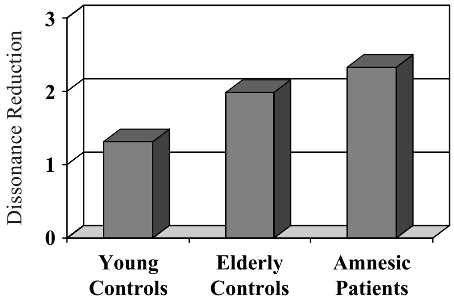

Figure 3.1. Dissonance reduction.

decision. Consequently, even if they feel physiological dissonance, amnesics will not have the counterattitudinal decision consciously accessible to which to attribute their physiological discomfort. Thus, if dissonance-based judgment changes require awareness of the conflict and the conscious effort to resolve the conflict, amnesic patients should not show any dissonance reduction. In fact, the amnesic patients showed just as much dissonance reduction as healthy elderly controls with natural declines in memory and twice as much dissonance reduction as college-age controls (see Figure 3.1), even though they evidenced no memory for having made a counterattitudinal decision moments earlier. In a second study, healthy college-age participants under cognitive load showed just as much dissonance reduction as participants who were not under cognitive load. These studies suggest, along with Shultz and Lepper (1995), that the C-system is not always involved in cognitive dissonance reduction.

Causal Inference in the X- and C-Systems

When the X-system is presented with a causal inference problem or any problem that requires propositional logic, the X-system transforms the problem into a similarity problem. That is, the X-system tends to resolve propositions by estimating the similarity between the antecedent and the consequent (Sloman, 1996). Thus, in the famous “Linda problem” (Donovan & Epstein, 1997; Tversky & Kahneman, 1983), individuals are given a description of Linda that is highly consistent with her being a feminist without actually specifying that she is one. Participants are then asked whether it is more likely that Linda is (a) a bank teller or (b) a bank teller and a feminist. The correct answer is a because of the principle of conjunction (i.e., the conjunction of any two events is less or equally likely than either event alone); however, most people choose b. This choice makes sense if the individuals are engaged in similarity-based pattern-matching because the description of Linda *is* more similar to b than a. The overall structure of the question can be formed as “If someone has background x, is she then more likely to be a or to be b?" Computing the similarity of the antecedent to each of the two possible consequents will yield the conjunction fallacy.

Unlike the X-system, the C-system is capable of testing for various kinds of asymmetric and unidirectional relationships between variables rather than just for bidirectional association. The C-system can represent statements such as "If p then q" without concluding p from the presentation of q, but the X-system cannot. This qualitative difference was previously used to explain why the X-system, but not the C-system, would be prone to the correspondence bias (Lieberman et al., 2002b). For though it may be true that "If a person is dispositionally friendly, that person is more likely to smile," it is an inferential error to infer that a person is dispositionally friendly because he or she has smiled. Nevertheless, the X-system presented with a smile will activate the representation for the corresponding trait of friendliness and draw this inference.

Whereas the X-system is designed to learn and use social and affective generalities, the C-system evolved to handle those *situations* in which the X-system fails (McClelland et al., 1995). Thus, the C-system needs to be able to represent individual situations that are exceptions to the rules and characterize how the same objective stimulus should be processed differently, depending on the situation. To facilitate this, the neural structure of prefrontal cortex is quite different from that of the X-system. Whereas representations in the X-system are massively distributed such that similar representations blur into one another, prefrontal and hippocampal representations are relatively sparse and separated from one another such that the activation of any one representation does not automatically activate semantically related representations (O'Reilly, Braver, & Cohen, 1999). Consequently, in the C-system, asymmetric relationships between representations can be produced and maintained without the interference of similarity effects (Holyoak & Hummel, 2000).

A recent functional magnetic resonance imaging (fMRI) study tested the reflection-reflexion hypothesis that asymmetric judgment processes depend on the C-system, whereas purely associative decision processes do not. In this study (Lieberman et al., 2002c), participants were shown word pairs one word at a time. Half of the time, participants judged whether there was a causal relation between the words (e.g., *fire* → *smoke*), and the other half of the time, participants judged whether there was an associative relation between the words (e.g., *bread* → *butter*). Even though the two trial types were equated for reaction time, suggesting that the two tasks were similarly difficult, only the causal inference trials activated prefrontal cortex. This suggests that the C-system is not just for more effortful decisions, but for qualitatively different kinds of decisions – in particular, those that require asymmetric processing of discrete symbols rather than computing similarity-based overlap between features.

REFLECTIVE PROCESSES DISRUPT REFLEXIVE PROCESSES

One of the basic principles of automaticity contends that genuinely automatic processes cannot be disrupted by controlled processes (Bargh, 1999). Controlled processes may frequently set automatic processes in motion, but the ability to run to completion once started is one of the hallmarks of automaticity theory dating back over a century (James, 1890). A number of judgment and decision-making studies conducted by Wilson and colleagues (Wilson, Dunn, Kraft, & Lisle, 1989; Wilson et al., 1993) highlight a possible exception to this rule. In Wilson's work, participants are required to generate careful introspective accounts of choices that would otherwise be made intuitively. Using a variety of outcome criteria, Wilson has found that introspection leads to poorer decisions. In one study (Wilson et al., 1993), participants were asked to provide ratings of five posters; some were artistic prints by Van Gogh and Monet, and others were humorous posters with captions. Only 5% of control participants preferred the humorous posters, whereas 36% of participants asked to introspect on the basis for their preference chose the humorous poster. Additionally, participants were allowed to take a copy of their preferred print home. When contacted weeks later, participants who had been in the introspection condition expressed less satisfaction with their earlier choice than did the control participants. Thus, introspecting on the reasons for our preferences changes our preferences temporarily, leading to outcomes that are less satisfactory in the long run.

Wilson and Schooler's (1991) explanation of these effects would suggest that introspective processes rely on the C-system and that the C-system sometimes generates less adaptive decision criteria than the X-system. In the poster study, most people probably have X-systems that prefer the artistic posters, but the X-system does not provide logical proofs for its output (Lieberman, 2000a). Thus, when people need to provide logical support for their preferences, they turn to the C-system, which allows for the development of logical arguments. Unfortunately, the C-system of individuals untrained in critiquing art is probably not sensitive to the same factors that drive the intuitive responses of the X-system. Instead, the C-system constructs relatively simplistic principles of preference based on those dimensions that can be easily articulated. This is easier to do for the humorous prints, which are quite simple and direct, if not of enduring aesthetic worth, compared with the complex appeal of real works or art.

This account is relatively agnostic as to whether the X-system continues to contribute its two cents to the preference judgments. Rather, this account seems to suggest that if the X-system processes do occur, they are passed over in favor of the C-system's conclusions. Here is a case, as described in the introduction to this chapter, in which the methods of social cognition cannot provide a full account of the relationship between the automatic and controlled processes. Introspection manipulations certainly increase the amount of C-system processing that occurs, but what is the consequence of C-system processing on the X-system processing that *would have occurred* had the C-system not been engaged? How can we distinguish between (a) the C-system outputs being selected over the ongoing X-system outputs and the possibility that (b) engaging the C-system prevents the X-system from producing its output in the first place (Gilbert, 1999)? Using fMRI, we can assess whether C-system processes merely override X-system processes, such that X-system activations are still present but do not covary with the resulting outputs, or whether the presence of C-system processes are associated with reduced X-system activity suggesting disruption.

In two fMRI studies of affective processing, an increase in prefrontal activity has been associated with diminished amygdala activity (Hariri, Bookheimer, & Mazziotta, 2000; Lieberman, Hariri, & Bookheimer, 2001a), suggesting that engaging the C-system can disrupt the X-system. In one of these studies (Lieberman et al., 2001a), African American and Caucasian participants were shown African American and Caucasian targets. In one task (“Match”), participants saw a pair of faces in addition to the target and had to choose which face of the pair was of the same race as the target face (e.g., choosing the African American face from the pair to match the African American target). In this condition, we replicated other fMRI studies of automatic stereotyping (Cunningham, Johnson, Gatenby, Gore, & Banaji, 2001; Hart et al., 2000; Phelps et al., 2000) and found amygdala activity to African American faces without any prefrontal activity. In the second task (“Label”), participants chose which of two race labels described the target. In the label condition, the minimal linguistic processing necessary to complete the task activated bilateral prefrontal cortex. Moreover, the amygdala activations were absent in this condition. Thus, when we activated the C-system ever so slightly, the X-system processing in the amygdala was disrupted. It is worth pointing out that although the linguistic labeling task presented the participants with fewer out-group faces than the perceptual matching task, there were still as many out-group faces presented in the linguistic task as in previous fMRI studies of automatic stereotyping. Moreover, we found a negative correlation between the responses of the prefrontal cortex and amygdala ($r = -.63$) during the Label condition. If the reduced amygdala activity was merely a consequence of not presenting a second African American face, there should not have been a strong negative correlation between prefrontal and amygdala activity (see Zárate & Stoever, this volume, for other neural approaches to stereotyping).

Although the results of this fMRI study may be somewhat provocative, this study is hardly a clean test of C-system processes disrupting X-system processing. Perceptual matching is certainly more reflexive than linguistic labeling, but neither comes from the standard arsenal of social cognition paradigms. To provide such a test, a subliminal mere exposure study was conducted in conjunction with a cognitive load manipulation (Lieberman & Jarcho, 2002). This provides a strong test of the hypothesis because subliminal processes are widely accepted as automatic or reflexive. Subliminal processes clearly do not involve effort, intention, or awareness, and because subliminal processes do not require controlled processing resources to operate, if cognitive load interrupts the subliminal mere exposure effect, the only sensible conclusion is that reflective processes can disrupt genuinely reflexive processes.

In this study, participants were shown a series of irregular polygons for very brief periods of time (80 ms), followed on each trial by a pattern mask (60 ms) and a second of a fixation cross alone. Each polygon was randomly presented in one of four locations on the screen, each parafoveal relative to the fixation cross in the middle of the screen. These methods were based on previously established techniques for ensuring subliminal exposure (Bargh & Chartrand, 2000). Each of 15 polygons was presented 1, 5, or 10 times. Half of the of the participants were also engaged in a cognitive load task during the presentation of the polygons. These participants were required to keep track of the number of tones at a particular pitch from among a series of tones. After the initial exposure period, participants were shown pairs of polygons and asked which they liked more. Unknown to the participants, in each pair, one polygon was new and the other had been presented subliminally during the prior exposure period. As shown in Figure 3.2, participants who were not under cognitive load replicated the standard mere exposure effect, preferring the previously exposed polygons to the new polygons two to one. Participants under cognitive load, however, showed no evidence of a mere exposure effect at all, instead expressing equal preference for the old and the new polygons.

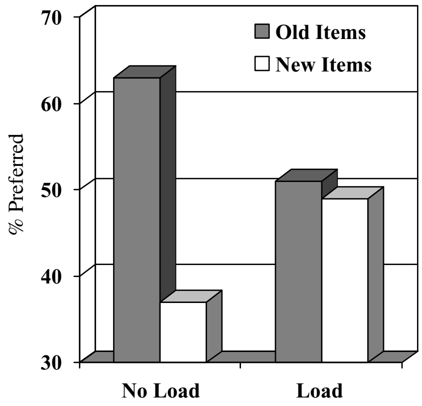

Figure 3.2. Mere exposure effects under conditions of cognitive load and no cognitive load.

If, as these data suggest, C-system processes can disrupt genuinely automatic X-system processes, then it would seem likely that inducing introspection would lead not only to the selection of C-system outputs over X-system outputs, but also to the disruption of the X-system processes that might have led to better decisions. This proposed mechanism complements, rather than supplants, that of Wilson and colleagues. People certainly do shift their bases of judgment as Wilson has described. I am merely pointing out that introspection, in addition to overriding our intuitive preferences, can disrupt the initial production of those intuitive preferences.

INDIVIDUAL DIFFERENCES IN REFLECTIVE JUDGMENT PROCESSES

If the X- and C-systems contribute to qualitatively different kinds of decision making, then individual differences in the reactivity of the neural bases of these systems should lead to personality differences in the likelihood that each kind of decision process occurs and is effective. In recent studies, extraversion has been associated with more efficient functioning in the executive components or working memory (Gray & Braver, 2002; Lieberman, 2000b; Lieberman & Rosenthal, 2001), which in turn is associated with prefrontal cortex functioning (Braver et al., 1997; Smith & Jonides, 1999). This difference in working memory efficiency predicted differences in social judgment under cognitive load, such that extraverts were more accurate under cognitive load than introverts, whereas there was no difference in the absence of cognitive load across three studies (Lieberman & Rosenthal, 2001). This finding contradicts the stereotype of the introvert as the more rational and precise decision maker, carefully weighing different options.

It is important here to differentiate the capacity of the C-system to contribute to decision making from the likelihood that the C-system will be called upon to contribute to decision making. A second personality variable, neuroticism, may be an important determinant of this latter aspect of decision making. Recall that in the reflection-reflection model, the C-system evolved to deal with problems that the X-system fails to resolve. Whereas prefrontal cortex is essential to control exerted by the C-system, the anterior cingulate plays a major role in choosing when control will be called upon. Individual differences in the sensitivity of the anterior cingulate should determine the likelihood that a person will rely on the C-system in his or her decision making. Neurotics are hypothesized to have more sensitive anterior cingulates, and there is growing evidence that anxiety, a major components of neuroticism (Gray, 1991), is associated with the sensitivity of the anterior cingulate (Davidson, Abercrombie, Nitschke, & Putnam, 1999; Eisenberger & Lieberman, 2002). Thus, neuroticism and extraversion together should predict the likelihood and capacity of an individual to engage C-system processes in decision making.

This does not mean that questionnaire measures of extraversion and heuroticism will allow us to easily assess these decision-making differences across individuals. These questionnaires were initially meant to predict the peripheral consequences of individual differences in the central nervous system (Eysenck, 1967). That is, the questionnaires were constructed to predict skin conductance, which in turn was thought to be a consequence of cortical arousal. This has turned out to be like a game of telephone in which the original message was not entirely clear to begin with. The questionnaires predict psychophyiological responses imperfectly (Matthews & Gililand, 1999), and the psychophysiological responses are affected by multiple processes in the brain in addition to cortical arousal. Furthermore, there is no simple neural equivalent to the concept of cortical arousal (Neiss, 1988; Robbins, 1997), upon which much of the original was based (Eysenck, 1967; Hebb, 1955).

Today, however, we can measure the sensitivity of different neural structures that are critical for self-regulation and problem solving relatively directly with neuroimaging. Given this access, attempting to find which structures in the brain have activation patterns that track existing questionnaire measures of personality would be one more bad link in the telephone chain. Instead, we would do well to build our personality theories around the capacities of different neurocognitive structures and the computational consequences of varying the sensitivity of these structures. From here, then, it would make sense to develop self-report and behavioral measures that predict the reactivity of one or more neural systems, because neuroimaging is still a very expensive and time-consuming way to measure these individual differences. Tasks that tap the central executive component of working memory can serve as a quick index of at least some aspects of prefrontal function (Lieberman & Rosenthal, 2001). Naomi Eisenberger and I (Eisenberger & Lieberman, 2001) are currently developing a behavioral task that measures anterior cingulate function. In our pilot fMRI work, we have presented participants with reversible illusions such as a Necker cube and have coded the rate of oscillation (i.e., flipping between the two views of the cube). Though this work is preliminary, thus far it appears that the oscillation rate is a good predictor of anterior cingulate activity. Incidentally, personality psychologists of the 1930s used this same measure but abandoned it because it did not correlate well with their questionnaire methods (Guilford & Braly, 1931; McDougall, 1933). Because the current approach allows researchers more direct access to the neurocognitive individual differences, it is possible that this time around, this task will provide us with better predictive power. With these better tools, psychologists should be able to reinvigorate the study of personality and its consequences for individuals in the tasks of everyday life.

CONCLUSION

In this chapter, I have laid out two neurocognitive systems involved in reflective and reflexive processing. On the one hand, these two systems correspond to cognitive processing dichotomies that have been around for a generation or more: automatic versus controlled and implicit versus explicit. At the same time, these old dichotomies are limited in their ability to provide adequate treatment to the positive contributions of each half of the dichotomy. Instead, these formulations define the dichotomy as one or more continua, and thus, both ends of the continua are defined as the presence or absence of particular characteristics (effort, intention, awareness). By focusing on the neural basis of these systems, links can be made to the known computational characteristics of these systems, and these characteristics provide us with important clues as to why the two systems produce the outputs they do. Furthermore, current operationalizations of automaticity and control make it difficult to identify anything but the relative contributions of each system, rather than the absolute contributions of each. Finally, this operationalization assumes that the effects of automatic and controlled processes add and subtract linearly, without interaction effects. Neuroimaging allows the study of the ongoing interactions between the two systems.

The reflective-reflexive model based in neurocognitive systems and the techniques of neuroimaging can make great contributions to the understanding of judgment and decision making. The findings from this chapter, as well as the other existing findings in social cognitive neuroscience, are undoubtedly just the tip of the iceberg. If we are ever going to size up the entire iceberg, we will need an army of scientists who are bilingual in social cognition and cognitive neuroscience, in terms of both theory and methods. As it stands, Berlitz does not have a crash course in either language, so let me end where I began by recommending that you take a cognitive neuroscientist to lunch.

References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4*, 267–278.

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

Baciu, M. V., Bonaz, B. L., Papillon, E., Bost, R. A., Le Bas, J. F., Fournet, J., & Segebarth, C. (1999). Central processing of rectal pain: A functional MR imaging study. *American Journal of Neuroradiology, 20*, 1920–1924.

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience, 12*, 298–309.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.). *Dual-process theories in social psychology* (pp. 361–382). New York: Guilford Press.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253–285). New York: Cambridge University Press.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* Cambridge, MA: MIT Press.

Bartels, A., & Zeki, S. (2000). The neural basis of romantic love. *NeuroReport, 17*, 3829–3834.

Benedict, R. H. B., Shucard, D. W., Santa Maria, M. P., Shucard, J. L., Abara, J. P., Coad, M. L., Wack, D., Sawusch, J., & Lockwood, A. (2002). Covert auditory attention generates activation in the rostral/dorsal anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14*, 637–645.

Berns, G. S., Cohen, J. D., & Mintun, M. A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science, 276*, 1272–1275.

Bookheimer, S. Y. (in press). fMRI of language. *Annual Review of Neuroscience, 25*.

Botvinick, M.M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2000). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage, 5*, 49–62.

Brehm, J. W. (1956). Post-decision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology, 52*, 384–389.

Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. New York: Wiley.

Breiter, H. C., Gollub, R. L., & Weisskoff, R. M. (1997). Acute effects of cocaine on human brain activity and emotion. *Neuron, 19*, 591–611.

Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). Making memories: Brain activity that predicts how well visual experience will be remembered. *Science, 281*, 1185–1187.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*, 123–152.

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

Cacioppo, J. T., & Gardner, W. L. (1999). Emotions. *Annual Review of Psychology, 50*, 191–214.

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, 280*, 747–749.

Carter, C. S, MacDonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., Noll, D., & Cohen, J. D. (2000). Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Science, 97*, 1944–1948.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control theory approach to human behavior*. New York: Springer-Verlag.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33*, 541–560.

Christoff, K., & Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within human prefrontal cortex. *Psychobiology, 28*, 168–186.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268–294.

Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C, & Banaji, M. R. (2001). *An fMRI study on the conscious and unconscious evaluations of social groups*. Paper presented at the UCLA Conference on Social Cognitive Neuroscience, Los Angeles.

Davidson, R. J., Abercrombie, H., Nitschke, J. B., & Putnam, K. (1999). Regional brain function, emotion and disorders of emotion. *Current Opinion in Neurobiology, 9*, 228–234.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: W. W. Norton.

Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences, 3*, 172–178.

Dennett, D. C. (1984). *Elbow room: The varieties of free will worth wanting*. Cambridge, MA: MIT Press.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences, 22*, 491–569.

Dewey, J. (1910). *How we think*. Boston: D. C. Heath.

Donovan, S., & Epstein, S. (1997). Conjunction problem can be attributed to its simultaneous concrete and unnatural representation, and not to conversational implicature. *Journal of Experimental Social Psychology, 33*, 1–20.

Dreyfus, H. L. (1991). *Being-in-the-world: A commentary on Heidegger's being and time, division I.* Cambridge, MA: MIT Press.

Eisenberger, N. I., & Lieberman, M. D. (2002). The role of the anterior cingulate in neuroticism and social cognition. Unpublished raw data. University of California, Los Angeles.

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology, 67*, 382–394.

Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, IL: Charles C. Thomas.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience, 13*, 171–180.

Festinger, L. (1964). *Conflict, decision, and dissonance*. Stanford, CA: Stanford University Press.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGraw-Hill.

Frankenstein, U. N., Richter, W., McIntyre, M. C., & Remy, F. (2001). Distraction modulates anterior cingulate gyrus activations during the cold pressure test. *NeuroImage, 14*, 827–836.

Fridja, N. H. (1986). *The emotions*. New York: Cambridge University Press.

Gilbert, D. T. (1989). Thinking lightly about others. Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189–211). New York: Guilford Press.

Gilbert, D. T. (1999). What the mind's not. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 3–11). New York: Guilford Press.

Goel, V., & Dolan, R. J. (2000). Anatomical segregation of component processes in an inductive inference task. *Journal of Cognitive Neuroscience, 12*, 110–119.

Grant, I., & Adams, K. M. (1996). *Neuropsychological assessment of neuropsychiatric disorders*. New York: Oxford University Press.

Gray, J. A. (1991). Neural systems, emotion and personality. In J. Madden (Ed.), *Neurobiology of learning, emotion, and affect* (pp. 273–366). New York: Raven Press.

Gray, J. R., & Braver, T. S. (2002). Personality predicts working-memory-related activation in the caudal anterior cingulate cortex. *Cognitive, Affective & Behavioral Neuroscience, 2,* 64–75.

Griffin, D. W., & Ross, L. (1991). Subject construal, social inference, and human misunderstanding. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 319–359). San Diego: Academic Press.

Guilford, J. P., & Braly, K. W. (1931). An experimental test of McDougall's theory of extroversion-introversion. *Journal of Abnormal and Social Psychology, 25*, 382–389.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814–834.

Hariri, A. R., Bookheimer, S. Y., & Mazziotta, J. C. (2000). Modulating emotional responses: Effects of a neocortical network on the limbic system. *NeuroReport, 11*, 43–48.

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. *NeuroReport, 11*, 2351–2355.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223–233.

Hebb, D. O. (1955). Drives and the CNS (conceptual nervous system). *Psychological Review, 62*, 243–254.

Heidegger, M. (1927/1962). *Being and time*. New York: Harper & Row.

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

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

Hogarth, R. H. (2001). *Educating intuition*. Chicago: University of Chicago Press.

Holyoak, K. J., & Hummel, J. E. (2000). The proper treatment of symbols in a connectionist architecture. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines* (pp. 229–263). Mahwah, NJ: Erlbaum.

Iwata, J., LeDoux, J. E., Meeley, M. P., Arneric, S., & Reis, D. J. (1986). Intrinsic neurons in the amygdaloid field projected to by the medial geniculate body mediate emotional responses conditioned to acoustic stimuli. *Brain Research, 383*, 195–214.

James, W. (1890/1950). *The principles of psychology.* New York: Dover.

James, W. (1913). *Essays in radical empiricism*. New York: Longman, Green.

Jellema, T., & Perrett, D. I. (2001). *Coding of visible and hidden actions: A tutorial*. Manuscript in preparation.

Kropotov, J. D., Crawford, H. J., & Polyakov, Y. I. (1997). Somatosensory event-related potential changes to painful stimuli during hypnotic analgesia: Anterior cingulate cortex and anterior temporal cortex intracranial recordings. *International Journal of Psychophysiology, 27*, 1–8.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel constraint satisfaction theory. *Psychological Review, 103*, 284–308.

Ladabaum, U., Minoshima, S., & Owyang, C. (2000). Pathobiology of visceral pain: Molecular mechanisms and therapeutic implications v. central nervous system processing of somatic and visceral sensory signal. *American Journal of Physiology: Gastrointestinal and Liver Physiology, 279*, G1–G6.

Langer, E. J., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology, 36*, 635–642.

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. New York: Simon & Schuster.

LeDoux, J. E., Ruggiero, D. A., & Reis, R. J. (1986). Projections to the subcortical forebrain from anatomically defined regions of the medial geniculate body in the rat. *Journal of Comparative Neurology, 242*, 182–213.

Lieberman, M. D. (2000a). Intuition: A social cognitive neuroscience approach. *Psychological Bulletin, 126*, 109–137.

Lieberman, M. D. (2000b). Introversion and working memory: Central executive differences. *Personality and Individual Differences, 28*, 479–486.

Lieberman, M. D., Chang, G. Y., Chiao, J., Bookheimer, S. Y., & Knowlton, B. J. (2002a). *An event-related fMRI study of artificial grammar learning in a balanced chunk strength design.* Unpublished manuscript. University of California, Los Angeles.

Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002b). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 199–249). New York: Academic Press.

Lieberman, M. D., Hariri, A., Bookheimer, S. (2001a). *Controlling automatic stereotype activation. An fMRI study*. Paper presented at the second annual meeting of the Society for Personality and Social Psychology, February 4, San Antonio, TX.

Lieberman, M. D., Hariri, A., Jarcho, J. J., & Bookheimer, S. Y. (2002a). Two Captains, One Ship: A Social Cognitive Neuroscience Approach to Disrupting Automatic Affective Processes. Unpublished manuscript. University of California, Los Angeles.

Lieberman, M. D., & Jarcho, J. M. (2002). *When controlled processing disrupts genuinely automatic processes: Cognitive load and subliminal mere exposure effects.* Manuscript in preparation.

Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001b). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science, 12*, 135–140

Lieberman, M. D., & Rosenthal, R. (2001). Why introverts can't always tell who likes them: Multi-tasking and nonverbal decoding. *Journal of Personality and Social Psychology, 80*, 294–310.

Lieberman, M. D., Sellner, D., Waldman, M. D., Satpute, A. B., Tabibnia, G., Bookheimer, S. Y., & Holyoak, K. J. (2002c). *Asymmetric decision making: An fMRI study of causal vs. associative inference*. Manuscript in preparation.

Lorberbaum, J. P., Newman, J. D., Dubno, J. R., Horwitz, A. R., Nahas, Z., Teneback, C. C., Bloomer, C. W., Bohning, D. E., Vincent, D., Johnson, M. R., Emmanuel, N., Brawman-Mintzer, O., Book, S. W., Lydiard, R. B., Ballenger, J., & George, M. S. (1999). Feasibility of using fMRI to study mothers responding to infant cries. *Depression and Anxiety, 10*, 99–104.

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.

Matthews, G., & Gilliland, K. (1999). The personality theories of H. J. Eysenck and J. A. Gray: A comparative review. *Personality and Individual Differences, 26*, 583–626.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are compleimentary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*, 419–457.

McClelland, J. L., Rumelhart, D. E., & Hinton, G. E. (1988). The appeal of parallel distributed processing. In A. M. Collins & E. E. Smith (Eds.), *Readings in cognitive science: A perspective from psychology and artificial intelligence* (pp. 52–72). San Mateo, CA: Morgan Kaufmann.

McDougall, W. (1933). *Energies of men*. New York: Scribner's.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Miller, G. A., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart, & Winston.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology, 9*, 90–100.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neuroscience, 6*, 414–417.

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice reduction efforts. *Journal of Personality and Social Psychology, 65*, 469–485.

Morris, J. S., Ohman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Science, 96*, 1680–1685.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77*, 167–184.

Neiss, R. (1988). Reconceptualizing arousal: Psychobiological states in motor performance. *Psychological Bulletin, 103*, 345–366.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259.

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

O'Reilly, R., Braver, T. S., & Cohen, J. D. (1999). A biologically-based computational model of working memory. In A. Maykae & P. Shah (Eds.), *Models of working memory* (pp. 375–411). New York: Oxford University Press.

Otten, L. J., Henson, R. N. A., & Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding: Relationship between findings from across- and within-task comparisons. *Brain, 124*, 399–412.

Perrett, D. I., Jellema, T., Frigerio, E., & Burt, M. (2001, April). *Using "social attention" cues (where others are attending) to interpret actions, intentions and emotions of others*. Paper presented at the UCLA Conference on Social Cognitive Neuroscience, Los Angeles.

Petrovic, P., Petersson, K. M., Ghatan, P. H., Stone-Elander, S., & Ingvar, M. (2000). Pain-related cerebral activation is altered by a distracting cognitive task. *Pain, 85*, 19–30.

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

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 550–585). Hillsdale, NJ: Erlbaum.

Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., & Bushnell, M. D. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science, 277*, 968–971.

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and gestalt principles: (Re)Introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review, 1*, 26–53.

Robbins, T. W. (1997). Arousal systems and attentional processes. *Biological Psychology, 45*, 57–71.

Rolls, E. T. (1999). *The brain and emotion*. New York: Oxford University Press.

Sawamoto, N., Honda, M., Okada, T., Hanakawa, T., Kanda, M., Fukuyama, H., Konishi, J., & Shibasaki, H. (2000). Expectation of pain enhances responses to non-painful somatosensory stimulation in the anterior cingulate cortex and parietal operculum/posterior insula: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience, 20*, 7438–7445.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1–66.

Shultz, T. R., & Lepper, M. R. (1995). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review, 103*, 219–240.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*, 1–27.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3–22.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283*, 1657–1661.

Smith, E. R. (1996). What do connectionist and social psychology offer each other? *Journal of Personality and Social Psychology, 70*, 893–912.

Smith, E. R., & DeCoster, J. (1999). Associative and rule-based processing: A connectionist interpretation of dual-process models. In S. Chaiken & Y. Tropez (Eds.), *Dual-process theories in social psychology* (pp. 323–336). New York: Guilford Press.

Spellman, B. A., & Holyoak, K. J. (1992). If Saddam is Hitler then who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality and Social Psychology, 62*, 913–933.

Tabibnia, G. (2002). *Appetitive and aversive (e)motive systems: Contribution of functional neuroimaging studies*. Manuscript in preparation.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.

Wagner, A. D., Schacter, D. L., Rotte, M., Kootstaal, W., Maril, A., Dale, A. M., Rosen, B. R., & Buckner, R. L. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science, 281*, 1188–1191.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., Santos, M. M., Thomas, C. R., & Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science, 10*, 119–125.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 446–496). New York: Oxford University Press.

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. *Journal of Neuroscience, 18*, 411–418.

Whitehead, A. N. (1911). *An introduction to mathematics*. London: Williams and Norgate.

Wiener, N. (1948). *Cybernetics: Control and communication in the animal and the machine*. Cambridge, MA: MIT Press.

Wilson, T. D., Dunn, D. S., Kraft, D., & Lisle, D. J. (1989). Introspection, attitude change, and attitude–behavior consistency: The disruptive effects of explaining why we feel the way we do. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 287–343). New York: Academic Press.

Wilson, T. D., Lisle, D. J., Schooler, J. W., Hodges, S. D., Klaaren, K. J., & LaFleur, S. J. (1993). Introspecting about reasons can reduce post-choice satisfaction. *Personality and Social Psychology Bulletin, 19*, 331–339.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60*, 181–192.

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology, 47*, 237–252.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology, 10*, 109–120.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*, 1–28.