DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Inhibition of Action, Thought, and Emotion: A Selective Neurobiological Review

Citation	Dillon, Daniel G., and Diego A. Pizzagalli. 2007. Inhibition of action, thought, and emotion: A selective neurobiological review. Applied and Preventive Psychology 12, no. 3: 99-114.
Published Version	doi:10.1016/j.appsy.2007.09.004
Accessed	February 17, 2015 10:49:56 PM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:3200668
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

(Article begins on next page)

Inhibition of Action, Thought, and Emotion: A Selective Neurobiological Review

Daniel G. Dillon and Diego A. Pizzagalli

Harvard University, Department of Psychology

Running head: NEUROBIOLOGY OF INHIBITION

Corresponding Address:

Diego A. Pizzagalli, Ph. D.

Department of Psychology

Harvard University

1220 William James Hall

33 Kirkland Street

Cambridge, MA 02138

Tel: (617) 496-8896

Fax: (617) 495-3728

e-mail: dap@wjh.harvard.edu

Abstract

The neural bases of inhibitory function are reviewed, covering data from paradigms assessing inhibition of motor responses (antisaccade, go/nogo, stop-signal), cognitive sets (e.g., Wisconsin Card Sort Test), and emotion (fear extinction). The frontal cortex supports performance on these paradigms, but the specific neural circuitry varies: response inhibition depends upon fronto-basal ganglia networks, inhibition of cognitive sets is supported by orbitofrontal cortex, and retention of fear extinction reflects ventromedial prefrontal cortex-amygdala interactions. Inhibition is thus neurobiologically heterogeneous, although right ventrolateral prefrontal cortex may support a general inhibitory process. Dysfunctions in these circuits may contribute to psychopathological conditions marked by inhibitory deficits.

Keywords: inhibition, executive function, emotion, psychopathology, fMRI

Inhibition is a key concept in psychology because so much of successful behavior depends on it: we need to inhibit distracting information in order to focus attention, inhibit irrelevant cues in order to retrieve particular memories, and inhibit habitual responses in order to make adaptive choices. Inhibitory successes and failures have real consequences, and the articles in this special issue attest to the fact that various forms of psychopathology are prominently characterized by inhibitory deficits. It is important to note, however, that inhibition is not unitary. Friedman & Miyake (2004), for example, conducted comprehensive analyses on a large dataset featuring several inhibitory tasks and found evidence for not one unique inhibitory process, but three: Prepotent Response Inhibition, Resistance to Distractor Interference (ignoring or filtering out task-irrelevant information), and Resistance to Proactive Interference (preventing previously relevant but now irrelevant information from intruding into memory) (for other ways of parsing the behavioral data on inhibition, see Harnishfeger, 1995; Nigg, 2000). On the basis of these results, they urged researchers to be more specific when referring to inhibition. In addition, inhibition's value as an explanatory construct with respect to certain paradigms has been questioned. For example, MacLeod, Dodd, Sheard, Wilson, & Bibi, (2003) investigated two phenomena widely believed to reflect inhibitory processes—negative priming and directed forgetting—and argued instead that these may primarily reflect a combination of routine memory retrieval and response conflict (in negative priming) and selective rehearsal (in directed forgetting). Considering the complexity of the behavioral research in this field, Aron (2007) argued that a neuroscientific approach may be particularly useful to researchers interested in inhibition. In particular, it may be possible to parse inhibition biologically by identifying brain regions that consistently and selectively participate in specific types of inhibitory tasks. Along this line of thought, demonstrating that increased activity in one brain region is consistently and

specifically tied to decreased activity in another would provide strong support for an inhibitory account.

This paper reviews the neurobiological substrates of inhibitory processes, and is organized into three main sections. We begin in Section I with inhibition of motor responses. Because there is little disagreement over the fact that humans (and non-human animals) can inhibit motor movements and there is a consistent literature on this research issue, response inhibition provides an excellent starting point. Section II covers cognitive inhibition, which is the topic addressed by the other papers in this issue. Cognitive inhibition is a broad concept that has been used to explain a wide variety of phenomena, including negative priming, Stroop interference, directed forgetting, and performance on the "think/no-think" memory paradigm (Anderson & Green, 2001) and the Wisconsin Card Sorting Test (WCST: Berg, 1948). Many of these phenomena have not been the focus of much neuroscientific study, and a review of all the relevant behavioral data is beyond the scope of this paper. Thus, we focus on the WCST, which has been widely investigated in the neuroscience literature. However, the WCST is a complex task that depends on many cognitive functions besides inhibition. Therefore, we also review findings from a paradigm that has successfully parsed cognitive inhibition into two components—attentional shifting and reversal learning (Dias, Robbins, & Roberts, 1996a, 1996b, 1997). Section III addresses extinction of conditioned fear, a form of emotion inhibition that is well-understood at both the behavioral and neural level.

All three sections feature a short introduction, description of the relevant paradigms, brief treatment of psychological mechanisms underlying performance, and a review of neuroscientific findings from work with non-human animals, investigations of patients with brain lesions, and neuroimaging experiments[1]. To preview the main conclusions, inhibition is generally supported

by top-down control mechanisms mediated by the frontal lobes. However, different forms of inhibition recruit distinct sectors of frontal cortex, including the dorsolateral, ventrolateral, orbitofrontal, and ventromedial prefrontal cortex (Figure 1), and the neural structures involved in inhibition vary accordingly. For example, fear extinction depends upon interactions between the ventromedial prefrontal cortex (VMPFC) and the amygdala (Quirk, 2006), but neither of these structures is critical to inhibition of motor responses or cognitive sets. Notably, there may be an exception to this rule. The right ventrolateral PFC (VLPFC)—also known as the inferior frontal cortex and encompassing Brodmann areas 44, 45, and 47/12 (Petrides & Pandya, 2002)—has been implicated in inhibition of both motor responses and cognitive sets, thus this region may support a general inhibitory process (for review, see Aron, Robbins, & Poldrack, 2004). The paper concludes with a brief summary and proposals for future directions, with a particular focus on experimental studies of psychopathology (Section IV).

I. Inhibition of Behavioral Responses

Response inhibition encompasses a variety of processes aimed at controlling motor behavior, particularly suppression of unwanted, prepotent, or reflexive actions. As it is widely accepted that motor movements can be withheld or withdrawn, response inhibition is a non-controversial concept (Aron, 2007). Furthermore, Friedman & Miyake (2004) found evidence for Prepotent Response Inhibition as a basic inhibitory process. In their analysis, the antisaccade, stop-signal, and Stroop paradigms loaded heavily on Prepotent Response Inhibition. However, others (e.g., Nigg, 2000) have argued that the Stroop task is more closely tied to a facet of cognitive inhibition (resistance to interference) than to response inhibition. Moreover, the go/nogo task (which was not considered in Friedman & Miyake, 2004) has been widely used in

neuroscientific studies of response inhibition (Aron, Robbins, et al., 2004). Therefore, we concentrate on data from the antisaccade, go/nogo, and stop-signal tasks. Each of these paradigms features a prepotent motor response that the participant must inhibit on a subset of trials. Successful performance permits the investigation of brain regions that support control over motor activity. Relative to healthy controls, individuals with schizophrenia exhibit deficits in saccade inhibition (e.g., Fukushima et al., 1988), while individuals with attention deficit hyperactivity disorder (ADHD) perform poorly on go/nogo and stop-signal tasks (e.g., Durston et al., 2003). Thus, response inhibition deficits may serve as endophenotypes for these conditions (Almasy & Blangero, 2001; Aron & Poldrack, 2005; Hutton & Ettinger, 2006).

Studying Response Inhibition in the Laboratory: the Antisaccade, Go/NoGo, and Stop-signal Tasks

The standard *antisaccade task* features two trial types: prosaccade and antisaccade (Hallett, 1978). Trials include presentation of an instructional cue indicating the trial type (prosaccade, antisaccade), a period of central fixation, and the sudden appearance of a lateral target. On prosaccade trials the participant moves his or her eyes from fixation towards the target as quickly as possible. By contrast, on antisaccade trials participants are to rapidly direct their gaze towards the direction opposite the target. Correctly executed antisaccades are hypothesized to engage two processes: inhibition of reflexive saccades towards the target and generation of voluntary saccades away from it.

In the antisaccade task there is a substantial preparatory interval between presentation of the instructional cue and the target. No such preparatory interval exists in standard *go/nogo tasks*; instead, participants respond to frequent go stimuli while withholding responses to

infrequently presented nogo stimuli. For example, in a recent study participants viewed a stream of letters and responded to every letter but "X"—the nogo stimulus—with a button press; when "X" was presented, the response needed to be withheld (Menon, Adelman, White, Glover, & Reiss, 2001). Slower reaction times (RTs) on successful nogo trials relative to go trials, as well as frequent errors of commission, demonstrate the difficulty of inhibiting the prepotent go response.

Although the go/nogo paradigm minimizes the preparatory interval relative to the antisaccade task, a critique of the paradigm is that on successful nogo trials the response is omitted entirely rather than withdrawn, raising the possibility that response inhibition may be confounded with selective attention (needed to discriminate between the go and nogo stimuli) and response selection as opposed to inhibition (Rubia, Smith, Brammer, & Taylor, 2003). An arguably more pure test of response inhibition is the *stop-signal task* (Logan & Cowan, 1984; Logan, Cowan, & Davis, 1984). The stop-signal task retains go trials but does not feature nogo stimuli. Instead, individual go trials are occasionally interrupted by a stop signal indicating that the ongoing response should be halted (e.g., on critical trials the go stimulus is presented and the participant begins to execute a button press, but then the stop signal is presented and the participant must cancel the button press). Inhibitory difficulty can be modulated by varying the interval between presentation of the go and stop stimuli, referred to as the stop-signal delay (SSD). When SSD is short, stopping is easier; when SSD is long, stopping is more difficult. By analyzing both the SSD associated with stopping successes and failures and the reaction time on go trials, it is possible to calculate the latency of the inhibitory process, referred to as the stop-signal reaction time (SSRT: Logan et al., 1984). Shorter SSRTs are associated with more efficient inhibition.

Psychological Processes Underlying Inhibition of Motor Responses

Performance in response inhibition paradigms has been explained via a race model and neurocognitive models of executive control, which are complementary. According to the race model, performance in the stop-signal task reflects the outcome of a contest between independent go and stop processes: whichever reaches a threshold value first determines the behavioral outcome (Logan et al., 1984). In the antisaccade task, the race is between reflexive processes underlying rapid orientation towards the lateral target and controlled processes supporting inhibition (Massen, 2004; Munoz & Everling, 2004). A prediction of the race model is that consistently delaying either the stop or go process should allow the other to reach threshold first. This hypothesis was supported by a study which showed that increasing the latency of correct antisaccades led to an increase in antisaccade errors, presumably because the prosaccade process reached threshold first on a larger number of trials (Massen, 2004). Notably, RT on prosaccade trials was not affected, supporting a corollary hypothesis of the race model—namely, that the stop and go processes operate in parallel and do not interfere with each other.

The race model highlights the competition between volitional/controlled processes and prepotent/reflexive processes that must be inhibited. Neurocognitive models posit that this competition is supported by interactions between executive mechanisms in the frontal lobes and posterior cortical/subcortical regions devoted to stimulus processing and motor responses (Miller & Cohen, 2001). A benefit of neurocognitive models is that they can provide insight into the mechanisms supporting volitional control. For example, effective performance in the antisaccade task depends on the ability to maintain a task goal ("look opposite the target") in the face of the competing tendency to orient towards the target (Nieuwenhuis, Broerse, Nielen, & Jong, 2004).

According to neurocognitive models, if the task goal is adequately represented in working memory, an inhibitory signal is sent from the frontal lobes to oculomotor regions and the saccade is inhibited. By contrast, failures of executive control—or "goal neglect"—should lead to failures of saccade inhibition. Psychological studies have found support for this hypothesis. For example, high working memory loads generated via a secondary *n*-back task disrupt saccade inhibition, leading to increased antisaccade errors relative to low memory load conditions (Mitchell, Macrae, & Gilchrist, 2002). Similarly, individuals with shorter working memory spans are more prone to antisaccade errors than individuals with longer spans (Unsworth, Schrock, & Engle, 2004), and both healthy aging and schizophrenia—each of which is associated with impaired frontal function—are associated with increased antisaccade errors (Nieuwenhuis et al., 2004; see also Minas & Park, 2007). These results make a point which might be particularly important for studies on psychopathology: failed attempts at response inhibition need not necessarily reflect a specific deficit in inhibitory mechanisms. Instead, they may be due to failures of executive control, that is, failure to maintain task goals and rapidly recruit the inhibitory mechanisms that underlie the stop process. These kinds of executive deficits are not specific to inhibition and would presumably be apparent in other contexts.

Neurobiological Mechanisms of Response Inhibition

　　Antisaccade task. The antisaccade task is attractive for neuroscientific investigations of response inhibition because the neural networks underlying saccade generation are well-understood (Figure 2; for more extensive reviews, see Hikosaka, Takikawa, & Kawagoe, 2000; Hutton & Ettinger, 2006; Munoz & Everling, 2004). Most important for this review is the fact that saccade generation is supported by interactions involving multiple sectors of the frontal

lobes (including the frontal eye fields (FEF), supplementary eye fields (SEF), and the

dorsolateral PFC (DLPFC)), the basal ganglia (including the caudate, putamen, and substantia

nigra), and the superior colliculus (SC), which influences saccade execution via connections with

the midbrain.

Consistent with race models, outcomes in the antisaccade task depend upon the relative

activity levels of two populations of neurons in the SC, saccade and fixation neurons (Munoz &

Everling, 2004). Whether or not a saccade occurs is determined by which of these two classes of

neurons exceeds a critical activity threshold first. The sudden appearance of the visual target will

prompt a rapid increase in the activity of saccade neurons. Therefore, it is hypothesized that

correct antisaccade performance depends on the baseline activity of saccade neurons being

suppressed below the baseline activity of fixation neurons, such that target appearance does not

push the activity of saccade neurons past threshold first.

Suppression of the baseline activity of saccade neurons is believed to stem from the

inhibitory influence of other neural structures. A series of studies involving patients with damage

to the frontal lobes implicates the DLPFC as the source of those inhibitory signals (Pierrot-

Deseilligny, Rivaud, Gaymard, & Agid, 1991; Pierrot-Deseilligny, Muri, Ploner, Gaymard,

Demeret, & Rivaud-Pechoux, 2003; Ploner, Gaymard, Rivaud-Pechoux, & Pierrot-Deseilligny,

2005). Damage to this region (or the white matter tracts that connect it to the basal ganglia)

yields increased errors on antisaccade trials. In addition, functional magnetic resonance imaging

(fMRI) studies of healthy individuals report greater DLPFC activation during antisaccades as

opposed to prosaccades (e.g., Ford, Goltz, Brown, & Everling, 2005; Matsuda et al., 2004).

Other possible sources of inhibitory signals are the SEF and the FEF (Munoz & Everling, 2004).

Electrophysiological recording studies in monkeys have demonstrated increased pre-target

activity in SEF fixation neurons preceding correct antisaccades relative to both incorrect

antisaccades and correct prosaccades (Amador, Schlag-Rey, & Schlag, 2004; Schlag-Rey,

Amador, Sanchez, & Schlag, 1997). Human fMRI studies have obtained similar results,

reporting increased pre-target activity in both the SEF (Ford et al., 2005) and FEF (Cornelissen

et al., 2002; Ford et al., 2005; O'Driscoll, Alpert, Matthysse, Levy, Rauch, & Holzman, 1995)

for correct antisaccades versus incorrect antisaccades and correct prosaccades.

Finally, the basal ganglia are critical to saccade generation and inhibition (for review, see

Hikosaka et al., 2000). The substantia nigra, one of the major output structures of the basal

ganglia, tonically inhibits the SC and prevents it from exciting midbrain saccade generators.

However, the caudate can inhibit the substantia nigra, disinhibiting the SC and leading to a

saccade. By contrast, a second neural circuit passing through other sectors of the basal ganglia,

including the globus pallidus and the subthalamic nucleus, can excite the substantia nigra,

increasing inhibition of the SC and preventing saccades.

Research on schizophrenia implicates basal ganglia dysfunction in impaired antisaccade

performance. Compared to healthy controls, individuals with schizophrenia (e.g., Fukushima et

al., 1988; Sereno & Holzman, 1995), first-degree relatives of schizophrenics (Clementz,

McDowell, & Zisook, 1994; Crawford, Sharma, Puri, Murray, Berridge, & Lewis, 1998), and

healthy participants with elevated levels of schizotypy (e.g., O'Driscoll, Lenzenweger, &

Holzman, 1998) generate increased numbers of antisaccade errors. Functional neuroimaging has

linked these deficits to decreased recruitment of the caudate, putamen, and globus pallidus

(Crawford, Puri, Nijran, Jones, Kennard, & Lewis,1996; Raemaekers et al., 2002; Raemaekers,

Ramsey, Vink, van den Heuvel, & Kahn, 2006). Although impairments in saccade inhibition are

not specific to schizophrenia (Brownstein et al., 2003; Munoz & Everling, 2004), these data

suggest that the antisaccade task may be sensitive to neural deficits implicated in the disorder (Hutton & Ettinger, 2006).

Go/NoGo and Stop-signal tasks. Inhibition of manual motor responses in go/nogo and stop-signal tasks also depends upon the interaction of frontal and basal ganglia control regions with motor output structures, including the thalamus and primary motor cortex (Figure 3; for review, see Band & van Boxtel, 1999). Lateral PFC regions appear to support inhibition in these paradigms. In monkeys, electrical potentials elicited by nogo stimuli were recorded from both the DLPFC and VLPFC regions, and electrically stimulating these regions approximately 100 ms after presentation of the go stimulus resulted in complete cancellation or dramatic delay of the go response (Sasaki, Gemba, & Tsujimoto, 1989; see also Sakagami, Tsutsui, Lauwereyns, Koizumi, Kobayashi, & Hikosaka, 2001). Similarly, an fMRI study of macaques found that relative to go trials, nogo trials elicited strong activity in bilateral VLPFC (Morita, Nakahara, & Hayashi, 2004).

 Convergent findings from human research suggest that the right VLPFC is especially critical to inhibition of motor responses. A noteworthy study administered the stop-signal task to patients with unilateral lesions of either right or left frontal regions (Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; see also Aron, Monsell, Sahakian, & Robbins, 2004). Compared to both normal controls and patients with left frontal damage, patients with right frontal damage exhibited increased SSRTs, a behavioral proxy of inefficient inhibition. Furthermore, the size of lesions in the right VLPFC was positively correlated with SSRT; notably, no other region in either hemisphere showed this relationship.

Functional neuroimaging studies reveal that nogo stimuli consistently elicit activity in a network of primarily right lateralized regions, including the VLPFC (Garavan, Hester, Murphy,

Fassbender, & Kelly, 2006; Garavan, Ross, & Stein, 1999; Konishi, Nakajima, Uchida, Kikyo, Kameyama, & Miyashita, 1999; Liddle, Kiehl, & Smith, 2001; Menon et al., 2001), and the stop-signal task consistently reveals activity in the right VLPFC (Chevrier, Noseworthy, & Schachar, 2007; Rubia et al., 2001; Rubia et al., 2003). Interactions between the right VLPFC and subcortical structures may underlie response stopping. A recent fMRI study observed right VLPFC and subthalamic nucleus (STN) activation on successful stop trials, and activity in these regions was correlated across participants (Aron & Poldrack, 2006). Furthermore, shorter SSRTs were associated with greater activation in the right VLPFC and STN on stop trials. Confirming the importance of right VLPFC to response inhibition, in a study with healthy controls Chambers et al. (2006) used transcranial magnetic stimulation to temporarily deactivate three cortical regions just prior to performance of the stop-signal task: right VLPFC, right DLPFC, and right parietal cortex. Only deactivation of the right VLPFC impaired stop-signal performance, leading to increased SSRT and increased errors of commission.

Finally, a recent fMRI study demonstrated right VLPFC activity during a modified version of the antisaccade task (Chikazoe, Konishi, Asari, Jimura, & Miyashita, 2007). As noted earlier, the classic antisaccade task involves establishment of a preparatory set prior to antisaccade execution, while the go/nogo and stop-signal tasks minimize preparation and put stronger demands on inhibition at the time of response execution. To address this issue, Chikazoe et al. (2007) modified the antisaccade task so that the preparatory period was minimized and demands on inhibition at the time of response execution were maximized. With these modifications, right VLPFC activation was observed on successful antisaccade trials versus control saccade trials.

Summary

Response inhibition has been studied with the antisaccade, go/nogo, and stop-signal tasks, each of which requires inhibition of a prepotent motor response. Performance on these tasks is well-modeled as a race between reflexive/prepotent go processes and volitional/controlled stop processes. Neurobiologically, response inhibition depends upon the interaction of frontal control systems with the basal ganglia and motor output regions. Although a variety of frontal regions are recruited by these tasks, right VLPFC activity has been directly tied to inhibitory control across multiple paradigms. Dysfunction in fronto-basal ganglia circuits has been observed in forms of psychopathology associated with deficits in response inhibition, including schizophrenia (e.g., Raemaekers et al., 2002, 2006) and ADHD (e.g., Aron & Poldrack, 2005; Casey et al., 1997; Nigg & Casey, 2005).

II. Inhibition of Cognitive Sets

It is relatively easy to infer when response inhibition has occurred: a motor response is withheld or withdrawn. By contrast, cognitive inhibition is often used to refer to a considerably more diverse and complex group of processes. For example, Joormann, Yoon, and Zetsche (2007; see also Joormann, 2004) argue that depression is associated with deficits in cognitive inhibition related to selective attention, working memory, and episodic memory. Specifically, depressed individuals have difficulty disengaging attention from emotionally negative material, inhibiting representations of negative material in working memory, and resisting their propensity to selectively retrieve negative memories from long-term storage. These phenomena are important and well-documented. However, as Joorman et al. acknowledge, whether or not they truly reflect inhibitory deficits is more controversial.

An example of this controversy is directly addressed in the papers by Dorahy (2007) and Minas and Park (2007), which review negative priming research as it applies to dissociative identity disorder and schizophrenia, respectively. Negative priming refers to the fact that if a target stimulus served as a distractor on the preceding trial, the latency to respond to it in the current trial is increased (for reviews, see May, Kane, & Hasher, 1995; Tipper, 2001). Most explanations of negative priming invoke an inhibitory mechanism: during selective attention tasks, target representations are amplified and distractor representations are inhibited, thus when a stimulus that was a distractor becomes a target, its representation begins in an inhibited state and processing is slowed. Based on this proposal, the negative priming paradigm is widely used as a test of inhibitory functions.

Competing hypotheses argue that negative priming does not depend on inhibition. For example, as Dorahy reviews, a theory emphasizing episodic retrieval proposes that distractors are initially given a "do not respond" tag (Neill & Valdes, 1992). When the same stimuli are presented as targets, automatic retrieval of the "do not respond" tag causes conflict, and resolving this conflict slows responding. This hypothesis thus explains negative priming without postulating an inhibitory mechanism. Minas and Park describe the feature mismatch account developed by Park and Kanwisher (1994), which is based on the fact that in many negative priming paradigms a perceptual characteristic serves to distinguish targets from distractors (e.g., distracting words are printed in red, while target words are printed in white). The feature mismatch account proposes that stimuli are encoded along with their perceptual characteristics, such that when a former distractor is presented as a target, there is conflict between the old perceptual features that are retrieved from memory (e.g., word was printed in red) and the new perceptual features being presented (e.g., word is now printed in white). This mismatch causes

conflict which takes time to resolve, and, again, this hypothesis accounts for negative priming without recourse to inhibition.

Supporting these hypotheses, in several studies MacLeod and colleagues have provided data suggesting that negative priming may be more closely tied to routine memory retrieval and conflict resolution than to inhibition (reviewed in MacLeod et al., 2003; see also MacDonald & Joordens, 2000). They have also critically analyzed data from "think/no-think" (Anderson & Green, 2001), directed forgetting (MacLeod, 1999), and lexical decision (Meyer & Schvaneveldt, 1976) tasks, and in each case have provided convincing alternatives to inhibitory explanations (MacLeod et al., 2003). It is important to note that a rapprochement may be possible: inhibitory processes may be more critical during stimulus encoding, while conflict resolution may be more critical during retrieval (Tipper, 2001). Aron (2007) argues that neuroscientific data may help resolve this controversy: demonstrating that increased activity in one brain region consistently causes decreased activity in another would provide compelling support for an inhibition account.

Researchers are beginning to examine the neural correlates of performance on various tasks thought to involve cognitive inhibition (e.g., Anderson, Ochsner, Kuhl, Cooper, Robertson, Gabrieli, et al., 2004; Depue, Curran, & Banich, 2007; Egner & Hirsch, 2005), and the body of knowledge in this area is small but growing. Rather than attempt to survey the scattered offerings, we concentrate on a larger body of work involving paradigms that manipulate rule-based stimulus response associations, referred to as cognitive sets (Buchsbaum, Greer, Chang, & Berman, 2005). Cognitive sets are typically established and maintained on the basis of positive feedback for correct responses. On critical trials, however, the previously correct response is no longer rewarded. In this case, the participant must switch from the old set to a new one; failure to do so results in perseverative errors. One hypothesis is that these types of switches depend on

cognitive inhibition of the old set, but set-switching likely involves many cognitive processes besides inhibition. Therefore, below we review research from a task that has successfully decomposed set-switching into simpler component processes (Dias, Robbins, & Roberts, 1996a, 1996b, 1997).

Studying Cognitive Inhibition in the Laboratory: the Wisconsin Card Sort Test, Dimensional Shifts, and Visual Discrimination Reversals

The Wisconsin Card Sort Test (WCST) is a classic test of cognitive flexibility (Berg, 1948), and successful performance appears to depend on the ability to inhibit prior cognitive sets. In the WCST, participants are given a deck of cards and asked to sort them according to four reference cards. All the cards depict geometric shapes that vary in form, color, and number: any of these dimensions can be used as the basis for sorting. Importantly, participants are not informed of the sorting rule and must deduce it by trial-and-error, using feedback provided by the experimenter. Over time, healthy participants deduce the rule (e.g., "sort by color") and respond accordingly. However, after 10 successful trials the experimenter changes the rule without warning (e.g., to "sort by number"). Effective behavior is hypothesized to depend on inhibiting the old cognitive set so that the new rule can be identified and used to guide responding.

The WCST is complex—in addition to cognitive inhibition, it makes demands on learning, selective attention, set-switching, and error correction. To reduce this complexity, new paradigms probe some of these component processes more directly (Figure 4; Dias, Robbins, & Roberts, 1996a, 1996b, 1997). In the paradigm developed by Dias and colleagues, trials begin with the presentation of two compound visual stimuli, one on the left and one on the right

(Figure 4a). Each stimulus consists of one or more lines of varying orientation overlaid on a different polygon—for example, a triangle (left) and a square (right), each overlaid with a unique pattern of lines. Based on feedback presented after each trial, the participant learns to attend to one dimension (e.g., polygons) while ignoring the other (e.g., lines), and also learns that a particular exemplar from the attended dimension (e.g., triangle) constitutes the correct stimulus (Figure 4b).

Once the participant has learned to attend to the correct stimulus, three manipulations are possible. First, in an intra-dimensional shift (Figure 4c), novel stimulus pairs are presented and reward feedback is transferred from one exemplar to another within the same dimension (e.g., from triangle to diamond). Second, in an extra-dimensional shift (Figure 4d), rewards are shifted to an exemplar from the other dimension (e.g., from the triangle to an exemplar from the line stimuli). Third, in a visual discrimination reversal (also simply called a reversal), reward feedback is shifted from one member of a stimulus pair to the other (e.g., from the compound stimulus on the left to the compound stimulus on the right; Figure 4e).

Psychological Processes Supporting Inhibition of Cognitive Sets

Successful performance on the WCST depends on multiple psychological processes. First, the correct stimulus-response rule must be learned and held in working memory. Second, upon receipt of either positive or negative feedback, the contents of working memory are monitored and updated (Monchi, Petrides, Petre, Worsley, & Dagher, 2001). Receipt of positive feedback supports maintenance of ongoing behavior, while negative feedback signals a need to shift set. Set-shifting is hypothesized to involve inhibiting the old set, attending to previously ignored stimulus dimensions, and forming new stimulus-response associations.

Most errors in the WCST are perseverative in nature, which suggests inhibitory deficiencies (Demakis, 2003; Sullivan, Mathalon, Zipursky, Kersteen-Tucker, Knight, & Pfefferbaum, 1993). However, the paradigm developed by Dias and colleagues involving dimensional shifts and discrimination reversals has revealed that perseverative errors in this type of paradigm may stem from two sources: failures of selective attention versus failures to update stimulus-reward associations following a reversal (Dias et al., 1996b, 1997). While both of these types of failures yield perseveration, the latter is more clearly related to inhibitory function and has been directly related to the orbitofrontal cortex (OFC).

Neurobiological Mechanisms of Cognitive Inhibition

Wisconsin Card Sort Test. The WCST is sensitive to frontal lobe dysfunction. In a classic study, Milner (1963) tested patients who had undergone cortical excisions as part of treatment for epilepsy. Patients with DLPFC damage were markedly impaired on the WCST, committing an increased number of perseverative errors relative to patients with damage to other frontal or temporal regions, and a recent meta-analysis confirmed that frontal lesions (as opposed to posterior lesions) are differentially associated with perseverative errors on the WCST (Demakis, 2003).

A meta-analysis of neuroimaging studies examined activations elicited across various stages of the WCST (Buchsbaum et al., 2005). A bilateral pattern of fronto-parietal activity was revealed, consistent with recruitment of fronto-parietal attention networks in the task (e.g., Woldorff, Hazlett, Fichtenholtz, Weissman, Dale, & Song, 2004). More focused investigations by Monchi and colleagues have identified dissociable roles for the DLPFC and VLPFC in the WCST (Monchi et al., 2001; Monchi, Petrides, Doyon, Postuma, Worsley, & Dagher, 2004).

Specifically, DLPFC was activated by both positive and negative feedback, while the VLPFC was only activated by receipt of negative feedback. The DLPFC activations are hypothesized to reflect this region's role in monitoring the contents of working memory, which would be updated upon reception of both kinds of feedback (Petrides, 2000). By contrast, selective activation of VLPFC by negative feedback is consistent with a role for this region in inhibition during set-shifting, given its well-established role in response inhibition. Furthermore, the caudate was also activated by negative feedback, consistent with the larger role for fronto-basal ganglia circuitry in inhibitory functions (e.g., Nigg, 2000). Supporting this hypothesis, a series of studies by Konishi and colleagues revealed that right VLPFC activation observed during set-shifting in the WCST overlapped with a right VLPFC region identified in a go/nogo study (Konishi et al. 1998, 1999), consistent with a general inhibitory role for this region.

Dimensional shifts and visual discrimination reversals. Data from the WCST provide some support for the conclusion that inhibition of cognitive sets is supported by a fronto-basal ganglia network. This tentative conclusion has been refined and extended by a series of studies targeting the neural correlates of dimensional shifts and visual discrimination reversals in marmosets. In an initial investigation, marmosets learned to selectively attend to one of two dimensions (polygons versus lines; Figure 4) and to reliably select exemplars within that dimension to obtain a food reward (Dias et al., 1996a). After training, the experimental group received excitotoxic lesions to the PFC, including lateral and OFC regions. Compared to control animals, the experimental group showed no deficits on either reacquisition of visual discrimination or on performance of intra-dimensional shifts. However, they required many more trials to successfully complete extra-dimensional shifts and also made significantly more perseverative errors during discrimination reversals. This result indicates that two component

processes implicated in the WCST—namely, shifting attention from one perceptual dimension to another and reversing a pre-existing stimulus-response association—are supported by discrete PFC regions.

In a subsequent study, separate lateral PFC and OFC lesions were made to dissociate their unique effects (Dias et al., 1996b). Neither group exhibited difficulties in reacquiring visual discriminations or performing intra-dimensional shifts. However, compared to controls, the lateral PFC lesion group was significantly impaired on extra-dimensional shifts (but unimpaired on reversals), while the OFC group was significantly impaired on reversals (but unimpaired on extra-dimensional shifts). These results were interpreted as supporting a dissociation between attentional processing recruited during extra-dimensional shifting (supported by lateral PFC), and affective processing underlying the substitution of one stimulus-reward association for another during reversals (supported by OFC). This interpretation is consistent with the fact that the DLPFC is implicated in a number of executive functions, including attentional shifting, while the OFC is connected to limbic and striatal regions associated with emotional information processing and reward (Rolls, 1996, 2000). A third study further extended these findings by demonstrating that the lateral and OFC lesions used in these studies do not disrupt learning *per se*, but are specifically tied to inhibitory control of attentional and affective processing (Dias et al., 1997). In this study, lesions were made to the lateral PFC and OFC before training. The acquisition of visual discriminations was not affected, but specific deficits in extra-dimensional shifting and reversals, respectively, were observed once again.

A conceptually related investigation of humans with damage to the DLPFC or OFC revealed similar results (Hornak et al., 2004). The task involved choosing one of two stimuli on each trial; monetary rewards and punishments were differentially associated with the two stimuli.

Neither the DLPFC group nor the OFC group had difficulty learning the task. However, after a certain number of trials the stimulus-outcome contingencies were reversed such that the previously rewarded stimuli became associated with punishment and vice-versa. As in the studies with marmosets, this reversal revealed severe deficits in humans with OFC lesions, who showed perseverative responding even after receiving large monetary punishments. As a group, the patients with DLPFC lesions did not show the same deficit. However, a subset of DLPFC patients were severely impaired and performed as poorly as patients with OFC damage. Post-test questioning revealed that these patients were inattentive to visual signals associated with monetary rewards and punishments that were provided to facilitate performance. Thus, at least in some cases, DLPFC lesions were again associated with attentional failures.

Collectively, these studies indicate that lateral PFC and OFC make differential contributions to tasks demanding cognitive flexibility, such as the WCST. Lateral PFC regions support attentional shifts between perceptual dimensions. By contrast, the OFC is recruited by discrimination reversals, which require a change in stimulus-response mapping. Critically, cognitive inhibition is more directly assessed by reversals than by extra-dimensional shifts, since the previously rewarded exemplar is still present and must be ignored (Hampshire & Owen, 2006). Therefore, these data suggest that the OFC is more directly involved in cognitive inhibition than lateral PFC regions. Notably, neuroimaging experiments and studies with brain damaged patients indicate that successful reversals depend in part on connections between the OFC connections and the striatum (e.g., Cools, Clark, & Robbins, 2004; Cools, Ivry, & D'Esposito, 2006). In addition, there is some evidence that patients with OFC damage make an increased number of perseverative errors on the WCST (Freedman, Black, Ebert, & Binns, 1998), as would be expected based on these findings.

Summary

Flexible behavior depends on the ability to efficiently use selective attention and working memory in the face of distracting information. Deficits in these abilities have been associated with depression (Joormann et al., 2007), schizophrenia (Minas & Park, 2007), and dissociative identity disorder (Dorahy, 2007), among other psychopathological conditions. However, whether these deficits are specifically related to inhibitory failures is controversial. Inhibition of previously rewarded cognitive sets is thought to be important for successful performance on the WCST, but neuroscientific research reveals that performance on the WCST depends upon a large number of brain regions, including the VLPFC, DLPFC, parietal lobes, and basal ganglia. Activity in many of these regions may reflect processes unrelated to inhibition. New paradigms designed to tease apart these component processes reveal that extra-dimensional shifting depends upon the integrity of lateral PFC regions. By contrast, stimulus reversals—which make heavy demands on cognitive inhibition—depend upon the OFC. These findings are consistent with the hypothesis that the DLPFC is involved in attentional shifts while the OFC is more directly implicated in inhibitory and affective processes evoked by stimulus reversals (Hornak et al., 2004).

III. Inhibition of Emotional Responses

Emotion dysregulation is characteristic of a variety of forms of psychopathology (American Psychiatric Association, 1994), and dysregulated fear responses play a prominent role in phobias, panic disorder, and post-traumatic stress disorder (Barlow, 2002). By studying extinction, researchers have made substantial progress in understanding the psychological and

neural mechanisms underlying the inhibition of conditioned fear responses (Quirk, 2006).

Below, we review evidence indicating that the VMPFC, amygdala, and hippocampus are critical

brain regions involved in fear extinction. Due to space limitations we must omit many important

details; interested readers are directed to more extensive reviews of the behavioral (Bouton,

2004) and neurobiological (Myers & Davis, 2007) literatures covering this topic.

Studying Extinction in the Laboratory

During the *acquisition* phase of fear conditioning experiments, a neutral stimulus (the to-

be conditioned stimulus, or CS) is paired with a noxious unconditioned stimulus (US), such as

an electric shock. Due to this pairing, the CS acquires the ability to elicit fear responses which

can be asssessed behaviorally (e.g., by measuring freezing behavior) and physiologically (e.g.,

by measuring increased skin conductance responses). During the *extinction* phase the CS is once

again presented alone. On early trials in this phase the CS elicits fear responses that then

progressively diminish in frequency and intensity. This reduced response to the CS constitutes

extinction.

Many experiments feature one or two variations on this basic theme. In *differential

conditioning paradigms* two CSs are presented. During acquisition, one (the CS+) is paired with

the US while the other (the CS-) is not (acting as a control condition): conditioning is measured

as the difference in response to the CS+ versus the CS-. In addition, it is valuable to distinguish

between short-term and long-term extinction processes. When the extinction phase is presented

at little or no delay after the acquisition phase, effects reflect short-term processes and constitute

within-session extinction (Quirk, Russo, Barron, & Lebron, 2000). By contrast, presentation of

the CS at a delay after the original extinction phase tests long-term memory for extinction learning (i.e., extinction retention).

Psychological Processes Underlying Extinction

Extinction depends on multiple psychological processes (Bouton, 2004; Myers & Davis, 2007), but the particular importance of associative learning mechanisms is supported by three phenomena: *spontaneous recovery, reinstatement,* and *renewal*. Spontaneous recovery refers to the fact that tests of long-term extinction often reveal substantial fear responding (for review, see Rescorla, 2004). This observation demonstrates that extinction is not supported by forgetting or unlearning of CS-US associations. Instead, it reflects inhibitory learning that suppresses the expression of the excitatory CS-US associations formed during acquisition. Spontaneous recovery suggests that the inhibitory extinction learning fades more rapidly than the excitatory conditioning learning, for reasons that are currently unclear.

The inhibitory hypothesis of extinction learning is also supported by reinstatement (Rescorla & Heth, 1975). Reinstatement refers to the fact that unsignaled US presentations, delivered after extinction, will restore the ability of the CS to elicit a fear response. Because the CS and US are only presented together during acquisition, reinstatement implies that excitatory CS-US associations must persist throughout extinction. It is important to note that reinstatement only occurs if the unsignaled US presentations are delivered in the context where reinstatement testing will take place (Bouton & Bolles, 1979). This finding demonstrates a critical principle: the response elicited by an extinguished CS is very sensitive to contextual manipulations.

The clearest examples of the context-dependency of extinction come from renewal studies (Bouton, 2004). In so-called ABA renewal paradigms, fear is acquired in context A and

extinguished in context B. When the CS is once again presented in context A, a robust

("renewed") fear response is observed. Fear renewal is not observed if the CS is tested in the

extinction context (e.g., an ABB paradigm would not reveal renewal). Renewal paradigms thus

highlight another important asymmetry with respect to conditioned fear: the excitatory

associations that underlie fear generally persist across contexts, while the inhibitory learning that

supports extinction is context-bound.

This asymmetry has been explained by positing that context serves as an "occasion-

setter" that facilitates the retrieval of a particular CS memory (Holland, 1992). On this account,

acquisition leads to a robust, excitatory CS-US association. During extinction, an inhibitory CS-

"no US" association is formed in a particular context. The occasion-setter hypothesis proposes

that context determines which of these two memories is retrieved and expressed (Bouton, 2004).

Specifically, the extinction context prompts retrieval and expression of the inhibitory

association, while other contexts lead to retrieval and expression of the excitatory association.

This hypothesis has considerable heuristic values since it can account for a wide range of

renewal effects.

Neurobiological Mechanisms of Extinction

Extinction is supported by neural systems involved in fear learning, inhibition, and

contextual processing, namely, the amygdala, the VMPFC, and the hippocampus, respectively

(Figure 5). Below we review both human and non-human animal studies that illustrate the

specific contributions made by these structures to extinction.

Amygdala. The amygdala is well-known for its role in the acquisition of conditioned fear

(for reviews, see Davis, 1994; LeDoux, 1995). During acquisition, sensory cortices transmit

information regarding the CS and US to the basolateral amygdaloid complex (BLA); this region

is crucial for the formation of excitatory CS-US associations. Expression of conditioned fear

depends on the amygdala's central nucleus (CE), which receives input from the BLA and

activates a number of brainstem and hypothalamic effector sites, resulting in the fear response.

From a molecular perspective, remarkable progress has been achieved to elucidate cellular and

molecular mechanisms (particularly those involving N-methyl-D-aspartate (NMDA) receptors

and glucocorticoids) implicated in both long-term memory of conditioned fear as well as

extinction learning, but these processes are beyond the scope of this review (the interested reader

is referred to Schafe, Nader, Blair, & LeDoux, 2001, and McGaugh & Roozendaal, 2002, for

excellent reviews).

The role of the amygdala in extinction has also emerged from functional neuroimaging

studies in humans. A handful of fMRI studies have demonstrated increased amygdala activation

to the CS+ during within-session extinction (Gottfried & Dolan, 2004; LaBar, Gatenby, Gore,

LeDoux, & Phelps, 1998; Milad, Wright, Orr, Pitman, Quirk, & Rauch, 2007), although one

study recorded a greater amygdala response to the CS- versus the CS+ (Phelps, Delgado,

Nearing, & LeDoux, 2004). Importantly, this effect was correlated with skin conductance

responses (SCRs) such that a larger amygdala response to the CS- (relative to the CS+)

correlated with a smaller conditioned response during extinction (Phelps et al., 2004).

An important goal for future work will be to find additional evidence for brain-behavior

relationships during extinction in humans. An exciting step in this direction has been made by

investigations examining the effects of the NMDA receptor partial agonist D-cycloserine (DCS).

Based on studies demonstrating that both systemic and intra-amygdala injections of DCS

facilitated extinction in rodents (Ledgerwood, Richardson, & Cranney, 2003; Walker, Ressler,

Lu, & Davis, 2002), Ressler and colleagues (Ressler et al., 2004) examined the effects of DCS on extinction in a clinical population. In a double-blind design, participants with acrophobia (an extreme and irrational fear of heights) received either single doses of DCS or placebo before undergoing two sessions of virtual exposure therapy in a virtual reality glass elevator. Outcome measures included skin conductance fluctuations and subjective ratings of distress during exposure, as well as self-reports of anxiety and avoidance of heights. Follow-up assessments were conducted one week and three months post-treatment.

No effects of DCS were observed during the first treatment session, indicating that the drug does not have anxiolytic effects. However, at every time point thereafter significantly better outcomes on virtually every measure were observed in the DCS group (versus the placebo group). Furthermore, at three months follow-up the DCS group reported exposing themselves to feared heights significantly more frequently than the control group, demonstrating that the effects generalized to the real world and were maintained long after treatment. These findings have since been conceptually replicated in a study of social anxiety disorder (Hofmann et al., 2006). As Ressler et al. (2004) point out, these studies showcase a new role for psychoactive drugs. Rather than being directed at presumed biochemical abnormalities in a patient population, DCS has been used to augment a learning process—extinction—that is critical for fear inhibition. An important issue to examine in the future will be whether DCS is also a useful adjunct to forms of psychotherapy which do not depend primarily on exposure.

Ventromedial PFC (VMPFC). The data reviewed above indicate that extinction involves the formation of inhibitory associations in the amygdala, but what neural structure is the source of the inhibition? A large body of evidence from rodent studies points to the VMPFC (Sotres-Bayon, Bush, & LeDoux, 2004). In an early study, Morgan, Romanski, and LeDoux (1993)

lesioned the VMPFC, established conditioned fear, and then conducted extinction sessions over several days. Compared to control animals, the VMPFC-lesioned group required significantly more days to extinguish fear responses to the CS, suggesting a loss of top-down inhibitory influence on the amygdala by the VMPFC.

Subsequent studies have revealed a more nuanced picture. Quirk et al. (2000) tested two groups of rodents: one group with extensive ("inclusive") VMPFC lesions (VMPFC-i group), and one group with lesions restricted to the rostral VMPFC (VMPFC-r group). In the VMPFC-r group, a section of caudal VMPFC referred to as infralimbic (IL) cortex was spared. Both within-session extinction (on Day 1) and extinction retention (on Day 2) were examined. No group differences in extinction were evident on Day 1. This important and surprising finding indicates that the VMPFC is not critical for short-term extinction. However, on Day 2 the VMPFC-i group showed virtually complete recovery of fear, whereas fear extinction was maintained in the other two groups. In other words, although they had displayed normal within-session extinction, the VMPFC-i group exhibited a complete failure of extinction retention. This result demonstrates that the VMPFC—particularly the IL cortex—is critical for long-term memory of extinction. To test this account, Milad and Quirk (2002) made electrophysiological recordings from the IL cortex. In agreement with the lesion data, spiking activity was not observed in response to the CS during acquisition or within-session extinction, but strong IL activity was recorded in response to the CS on Day 2 extinction. Furthermore, this activity was related to extinction retention: rats with increased IL activity demonstrated better memory for extinction.

Finally, Quirk, Likhtik, Pelletier, and Pare (2003) found that electrical stimulation of the VMPFC reduced the sensitivity of the central nucleus of the amygdala to inputs from the BLA

and the insula. The authors proposed that the VMPFC inhibits the amygdala via intercalated cells, which send inhibitory projections to the central nucleus. Increased excitatory input from the VMPFC to the intercalated cells thus yields increased inhibition of the central nucleus, which in turn results in reduced expression of fear (Figure 5).

Functional neuroimaging research indicates that the VMPFC's role in fear extinction has been conserved in humans. In humans, the rostral cingulate, subgenual cingulate, and medial OFC are generally considered to constitute the VMPFC. Three studies suggest a role for one or more of these regions in within-session extinction of conditioned fear. Two found that the VMPFC responded more strongly to the CS+ (versus the CS-) during within-session extinction (Gottfried & Dolan, 2004; Milad et al., 2007); this pattern of responding was also observed in the caudal OFC (Gottfried & Dolan, 2004). By contrast, another study identified two regions in the VMPFC—one in the subgenual cingulate, one in the medial gyrus—that responded more strongly to the CS- than the CS+, and in fact showed substantially decreased responding to the CS+ (Phelps et al., 2004).

Interestingly, none of these studies reported correlations between VMPFC activity and psychophysiological measures of within-session extinction. However, in one study extinction success on Days 1 and 2 (as measured by SCRs) was correlated with subgenual cingulate activity during a test of extinction recall given on Day 2 (Phelps et al., 2004). In a psychophysiological study, Milad and colleagues (Milad, Quinn, Pitman, Orr, Fischl, & Rauch, 2005) found that long-term extinction retention was positively correlated with the thickness of the medial OFC as measured by structural MRI. Furthermore, in a re-analysis of these data, this group showed that extinction retention fully mediated the link between medial OFC thickness and the personality trait of extraversion (Rauch, Milad, Orr, Quinn, Fischl, & Pitman, 2005).

Thus, a thicker medial OFC was associated with a better capacity to retain fear extinction, which in turn was associated with an extroverted personality. Finally, a recent fMRI study tested extinction recall in an ABB renewal paradigm (Milad et al., 2007). This study featured a paradigm in which conditioned fear to two CS+s was established but only one CS+ was extinguished (CS+E); the other was not (CS+U). During the test of extinction recall, significantly greater VMPFC activity was elicited by the CS+E as opposed to the CS+U.

In summary, the human and rodent literatures indicate that the VMPFC is critically involved in fear extinction. One apparent discrepancy concerns within-session extinction. Studies in rodents consistently reveal that VMPFC is not critical for within-session extinction, whereas human studies reveal VMPFC activity on such tests. It is not clear how to account for this difference, but the lack of correlations between within-session VMPFC activations and behavioral measures of conditioning suggests that this region may not actually be critical to within-session extinction in humans. Another possibility is that greater explicit awareness of CS-US contingencies in humans may lead to the deployment, during within-session extinction, of emotion regulation strategies that recruit VMPFC regions (Urry et al., 2006).

Hippocampus. In their test of context-dependent extinction retention, Milad et al. (2007) found significant hippocampal activation. Furthermore, activity in both the hippocampus and VMPFC was positively correlated with extinction retention. This study thus provides the first evidence that the human hippocampus and VMPFC work together to constrain fear expression in a contextually-sensitive fashion.

These findings are consistent with results from non-human animals. An important study conditioned rats in one context and conducted extinction training in a different context (Corcoran & Maren, 2001). After extinction, muscimol (a gamma-aminobutyric receptor agonist

that temporarily inactivates brain tissue) was injected in the dorsal hippocampus. Extinction retention was then tested, either in the context in which extinction training had been conducted, or in a different context. As expected, control animals injected with saline showed extinction retention when the extinction training and testing contexts were the same but showed fear renewal when these two contexts differed. By contrast, rats injected with muscimol displayed equivalent fear in both contexts. This finding indicates that the hippocampus is required for appropriate retrieval of contextual information relevant to expression of extinction; indeed, this brain region may be a critical contributor to occasion-setting as it relates to fear extinction. A subsequent study indicates that the hippocampus may also be important for the acquisition and consolidation of extinction (Corcoran, Desmond, Frey, & Maren, 2005). Specifically, muscimol injections in the dorsal hippocampus given after acquisition attenuated within-session extinction and prevented the consolidation of context-dependent extinction.

The importance of the hippocampus in emotion inhibition is suggested by evidence of hippocampal dysfunction in various forms of psychopathology, including depression and post-traumatic stress disorder (PTSD: Davidson, Jackson, & Kalin, 2000; Phillips, Drevets, Rauch, & Lane, 2003). In particular, chronic PTSD is associated with reductions in hippocampal volume (Bremner et al., 1995; Bremner et al., 1997). A long-standing question concerns the direction of causality in this relationship: are individuals with smaller hippocampal volumes more likely to develop PTSD following trauma exposure, or does PTSD drive a reduction in hippocampal volume? An important study by Gilbertson and colleagues (Gilbertson et al., 2002) supports the former hypothesis. They studied monozygotic twin pairs in which one twin was a Vietnam combat veteran and the other was not. Furthermore, the veterans were divided into two groups: those with PTSD, and those without. As in previous studies, veterans with PTSD had smaller

hippocampal volumes than those without. However, the critical finding concerned these men's twins, who had not been exposed to trauma—unaffected brothers of veterans with severe PTSD had significantly smaller hippocampal volumes than brothers of veterans without PTSD. In other words, small hippocampal volume appears to be a risk factor for the development of PTSD. This may be related to extinction: individuals with small hippocampal volumes may be impaired in their ability to either acquire and/or retrieve information that should help restrict the expression of fear to particular contexts.

Summary

Extinction of conditioned fear is a well-studied form of emotional inhibition. Whereas conditioned fear generalizes across contexts, extinction is remarkably context-dependent, as demonstrated by renewal studies. Neurobiologically, extinction is supported by new learning in the amygdala and appears to reflect the operation of inhibitory signals sent from VMPFC; the hippocampus is critical for the formation and retrieval of contextual information. These findings have considerable clinical relevance with respect to anxiety disorders. For example, the etiology of PTSD is well-modeled as a particularly intense fear conditioning episode, and in comparison with healthy controls, patients with PTSD demonstrate amygdala hyper-responsivity and attenuated recruitment of VMPFC regions during emotional provocation paradigms (reviewed in Rauch, Shin, & Phelps, 2006). Despite this overlap, relatively few studies have actually used fear conditioning and extinction paradigms in conjunction with patient populations—more studies of this kind are needed. A small number of studies has already successfully used DCS to augment extinction learning in patients with phobias (Ressler et al., 2004; Hofmann et al., 2006). Thus,

future research on the extinction of conditioned fear is expected to contribute both to basic science and to the understanding and treatment of psychopathology.

IV. Conclusions and Future Directions

Adaptive behavior in a fluctuating and unpredictable environment relies on flexible and accurate inhibition of prepotent responses, cognitive sets, and emotions. Various forms of inhibition have been described, including *response inhibition* (e.g., inhibition of prepotent or reflexive behavioral responses), *cognitive inhibition* (e.g., inhibition of irrelevant information), and *emotional inhibition* (e.g., inhibition of fear responses). The goal of the present review was to summarize and critically discuss the neural bases of inhibitory function through an integration of experimental tasks and approaches, including functional neuroimaging and lesion studies in humans and neurophysiological data in animals. Several important points emerged. First, although the prefrontal cortex plays a pivotal role in inhibitory functions, it is clear that specific facets of inhibition rely on partially non-overlapping neural pathways. Specifically, response inhibition, cognitive inhibition, and emotional inhibition are supported by a right-lateralized fronto-basal ganglia circuitry, the OFC, and interactions between the VMPFC and the amygdala, respectively. Accordingly, from both a psychological and neurobiological perspective, inhibition is a heterogeneous construct, and findings from the present review support recent taxonomic approaches to inhibition-related functions (Friedman & Miyake, 2004; Nigg, 2000). Critically, recent advances in experimental psychology and affective neuroscience have allowed researchers to "dissect" inhibitory functions and identify its critical sub-components, opening new avenues for a more precise characterization of various disorders featuring impairments in inhibition-related processes, including ADHD (e.g., Nigg & Casey, 2005), schizophrenia (e.g., Fukushima

et al., 1988), PTSD (e.g., Bremner et al., 1995, Rauch et al., 2006), depression (e.g., Goeleven, De Raedt, Baert, & Koster, 2006), and personality disorders (Nigg, Silk, Stavro, & Miller, 2005). In ADHD research, for example, this approach has allowed researchers to identify dysfunctions in response inhibition, but generally normative cognitive inhibition (see Nigg, 2000, for a review). Future research is warranted to evaluate whether dysfunctions in neural pathways subserving separable inhibition-related processes might serve as endophenotypes for various psychopathological conditions (Almasy & Blangero, 2001).

Second, the right VLPFC appears to be critically implicated in both response inhibition and cognitive inhibition, suggesting that this region supports a general inhibitory process (Aron, Robbins, et al., 2004; Konishi et al., 1999). This finding is intriguing, particularly when considering that the VLPFC is one of the last regions to develop ontogenetically (Pandya & Barnes, 1987). Consistent with this anatomical evidence, increases in cortical thickness (Sowell, Thompson, Leonard, Welcome, Kan, & Toga, 2004) and task-related functional activation (Rubia et al., 2006) have been described in VLPFC regions throughout development. Moreover, a recent study using diffusion tensor imaging to assess brain connectivity *in vivo* showed maturation of connections between right VLPFC and the basal ganglia between the age of 7 and 31 years; notably, enhanced connectivity correlated with improved recruitment of cognitive control in a go/nogo task (Liston et al., 2006). Collectively, these findings indicate that prolonged development of regions critically implicated in inhibition-related functions might provide a vulnerability window increasing the risk for specific forms of psychopathology.

Several critical issues should be investigated in future studies. First, recent evidence indicates that individual difference variables, including sex (e.g., Li, Huang, Constable, & Sinha, 2006; Garavan et al., 2006), age (e.g., Nielson, Langenecker, & Garavan, 2002) and genotypes

(e.g., Pezawas et al., 2005), modulate inhibition-related functions and underlying neural circuitries. A better understanding of the modulatory effects of these variables, particularly with respect to their role in increasing vulnerability to psychopathology, is needed. Second, our understanding of the contributions of various neurotransmitters (including serotonin, dopamine, and noradrenaline) to inhibitory-related functions is limited (for review, see Robbins, 2007). Early conceptualizations emphasized the role of serotonin in behavioral inhibition (e.g., Soubrié, 1986), but recent evidence indicates that other neurotransmitters (e.g., noradrenaline) are also critically involved (Chamberlain, Muller, Blackwell, Clark, Robbins, & Sahakian 2006). Clearly, a better understanding of the neurochemical correlates of inhibition promises to have important implications for pharmacological treatments of disorders characterized by inhibition-related dysfunctions (Lucki, 1998; Robbins, 2007).

A final theme emerging from the present review is that there is an acute need for increased research on cognitive inhibition. The basic phenomena that constitute response inhibition and fear extinction are relatively clear-cut and well-understood. By contrast, performance on many of the paradigms thought to tap cognitive inhibition—including negative priming and the WCST—may primarily reflect the contribution of other processes, including routine memory retrieval and conflict resolution (Aron, 2007; MacLeod et al., 2003; but see Tipper, 2001). Careful behavioral and neuroscientific research is needed to clarify this picture. The most powerful neuroscientific demonstrations of inhibitory effects will likely not rely solely on fMRI data. As Aron (2007) points out, the blood-oxygenation-level-dependent (BOLD) effect measured in most fMRI studies does not primarily or selectively reflect the spiking output of a brain region. In other words, decreased BOLD signal in a neural structure, while informative, does not necessarily imply inhibition of that structure. Complementary approaches, including

studies of populations with brain lesions, single-cell recording studies in non-human animals, and intra-cranial recordings in humans, will be necessary to arrive at a complete picture. Regardless of the exact mechanisms involved, many of the tasks hypothesized to assess cognitive inhibition are already useful for revealing deficits associated with various forms of psychopathology, as the other articles in this special issue illustrate.

Footnotes

[1]Although functional neuroimaging techniques have significantly improved our understanding of brain pathways implicated in inhibition, it is important to emphasize that—due to their correlational nature—these approaches cannot demonstrate whether particular brain regions are *necessary* for specific functions. This critical information can be derived from studies in experimental animals, studies investigating humans with focal brain lesions (Rorden & Karnath, 2004), as well as studies utilizing transcranial magnetic stimulation to induce transient and "virtual" lesions (Pascual-Leone, Walsh, & Rothwell, 2000). Throughout this review, information gathered from these different approaches will be integrated.

Acknowledgements

Figure Captions

Figure 1. Regions of prefrontal cortex (PFC) implicated in inhibition. *a*) Dorsolateral PFC

(blue) and ventrolateral PFC (orange). *b*) Ventromedial PFC (red) and orbitofrontal cortex

(green). Reprinted with permission from Davidson, Pizzagalli, Nitschke, & Putman (2002). (For

interpretation of the references to color in this figure legend, the reader is referred to the web

version of the article.)

Figure 2. Neural bases of antisaccades. Simplified fronto-basal ganglia-collicular loop

underlying saccade generation and inhibition (adapted from Munoz & Everling, 2004). Saccades

are controlled by the midbrain reticular formation, which receives projections from the superior

colliculus, SEF, and FEF. In addition, the DLPFC, SEF, FEF, and basal ganglia can influence

eye movements via their projections to the superior colliculus. Note that many structures and

connections have been omitted for simplicity.

Figure 3. Neural basis of response inhibition in the go/nogo and stop-signal tasks (adapted from

Band & van Boxtel, 1999). Manual responses are under the influence of two neural loops. The

primary loop (black lines) involves connections between cortical structures (including the

DLPFC and VLPFC), the basal ganglia, and the thalamus. This loop is directly implicated in

response selection and response inhibition. The secondary loop (gray lines) involves connections

between more restricted cortical regions, the cerebellum, and the thalamus, and is thought to

fine-tune activity in the first loop. Output from these loops is integrated at the level of primary

motor cortex, which projects to the spinal cord (heavy black line). Several connections and

cortical regions have been omitted for simplicity. DLPFC: dorsolateral prefrontal cortex; SMA:

supplementary motor area; VLPFC: ventrolateral prefrontal cortex.

Figure 4. Example trials from the test of dimensional shifts and discrimination reversals (adapted

from Dias et al., 1996). Trials feature two compound stimuli consisting of line exemplars

overlaid on polygon exemplars. Correct choices are indicated by a plus (+), incorrect choices are

indicated by a minus (-). *a*) Compound discrimination: The participant must first identify the

correct exemplar (e.g., the triangle) from the correct dimension (e.g., the polygons). *b*) Correct

performance requires retaining the selection rule across trials. *c*) *Intra-dimensional shift*: A new

exemplar (diamond) in the same dimension (polygons) becomes the correct stimulus. *d*) *Extra-*

dimensional shift: An exemplar from the other dimension (lines) becomes correct. *e*)

Discrimination reversal: Stimuli from the previous trial are retained, but the previously correct

stimulus becomes incorrect and vice-versa.

Figure 5. Neural mechanisms involved in the acquisition and extinction of conditioned fear.

During fear acquisition, sensory information regarding the CS+ and US enters the basolateral

amgydaloid (BLA) complex via the cortex and thalamus; the BLA is where CS-US associations

are formed. The BLA sends excitatory projections to the central nucleus (CE) of the amygdala.

The central nucleus controls fear expression via its projections to a number of effector sites.

These include the lateral hypothalamus (LH), periaqueductal gray (PAG), and reticularis pontis

caudalis (RPC), which are important for autonomic components of the fear response, freezing

behavior, and startle-potentiation, respectively. The VMPFC mediates extinction of conditioned

fear, possibly through its connections with intercalated cell masses (ITC). The VMPFC sends

excitatory projections (+) to the ITC, which in turn send inhibitory projections (-) to the CE.

Thus, the net effect of vmPFC activity is inhibition of both CE activity and the fear response.

The hippocampus also sends projections to the amygdala, and has been implicated in contextual

control of extinction.

Reference List

Almasy, L., & Blangero, J. (2001). Endophenotypes as quantitative risk factors for psychiatric disease: rationale and study design. *American Journal of Medical Genetics, 105,* 42-44.

Amador, N., Schlag-Rey, M., & Schlag, J. (2004). Primate antisaccade. II. Supplementary eye field neuronal activity predicts correct performance. *Journal of Neurophysiology, 91*, 1672-1689.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, *410*, 366-369.

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., et al. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, *303*, 232-235.

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *The Neuroscientist*, *13*, 1-15.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience, 6,* 115-116.

Aron, A. R., Monsell, S., Sahakian, B. J., & Robbins, T. W. (2004). A componential analysis of task-switching deficits associated with lesions of left and right frontal cortex. *Brain*, *127*, 1561-1573.

Aron, A. R., & Poldrack, R. A. (2005). The cognitive neuroscience of response inhibition: relevance for genetic research in ADHD. *Biological Psychiatry, 57,* 1285-1292.

Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. *Journal of Neuroscience, 26,* 2424-2433.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences, 8*, 170-177.

Band, G. P. H., & van Boxtel, G. J. M. (1999). Inhibitory motor control in stop paradigms: review and reinterpretation of neural mechanisms. *Acta Psychologica, 101,* 179-211.

Barlow, D. H. (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. (2nd ed.) New York: Guilford Press.

Berg, E. A. (1948). A simple objective treatment for measuring flexibility in thinking. *Journal of General Psychology, 39,* 15-22.

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory, 11,* 485-494.

Bouton, M. E., & Bolles, R. C. (1979). Contextual control of the extinction of conditioned fear. *Learning and Motivation, 10,* 455-466.

Bremner, J. D., Randall, P., Scott, T. M., Bronen, R. A., Seibyl, J. P., Southwick, S. M. et al. (1995). MRI-based measurement of hippocampal volume in patients with combat-related posttraumatic stress disorder. *American Journal of Psychiatry, 152,* 973-981.

Bremner, J. D., Randall, P., Vermetten, E., Staib, L., Bronen, R. A., Mazure, C. et al. (1997). Magnetic resonance imaging-based measurement of hippocampal volume in posttraumatic stress disorder related to childhood physical and sexual abuse—a preliminary report. *Biological Psychiatry, 41,* 23-32.

Brownstein, J., Krastoshevsky, O., McCollum, C., Kundamal, S., Matthysse, S., Holzman, P. S. et al. (2003). Antisaccade performance is abnormal in schizophrenia patients but not in their biological relatives. *Schizophrenia Research, 63,* 13-25.

Buchsbaum, B. R., Greer, S., Chang, W.-L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the Wisconsin Card-Sorting Task and component processes. *Human Brain Mapping, 25,* 35-45.

Casey, B. J., Castellanos, F. X., Giedd, J. N., Marsh, W. L., Hamburger, S. D., Schubert, A. B. et al. (1997). Implication of right frontostriatal circuitry in response inhibition and attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 36,* 374-383.

Chamberlain, S. R., Muller, U., Blackwell, A. D., Clark, L., Robbins, T. W., & Sahakian, B. J. (2006). Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science, 311*, 861-863.

Chambers, C. D., Bellgrove, M. A., Stokes, M. G., Henderson, T. R., Garavan, H., Robertson, I. H. et al. (2006). Executive "brake failure" following deactivation of human frontal lobe. *The Journal of Cognitive Neuroscience, 18,* 444-455.

Chevrier, A. D., Noseworthy, M. D., & Schachar, R. Dissociation of response inhibition and performance monitoring in the stop signal task using event-related fMRI. *Human Brain Mapping* (2007), doi: 10.1002/hbm.20355

Chikazoe, J., Konishi, S., Asari, T., Jimura, K., & Miyashita, Y. (2007). Activation of right inferior frontal gyrus during response inhibition across response modalities. *Journal of Cognitive Neuroscience, 19,* 69-80.

Clementz, B. A., McDowell, J. E., & Zisook, S. (1994). Saccadic system functioning among schizophrenic patients and their first-degree biological relatives. *Journal of Abnormal Psychology, 103,* 277-287.

Cools, R., Clark, L., & Robbins, T. W. (2004). Differential responses in human striatum and prefrontal cortex to changes in object and rule relevance. *Journal of Neuroscience, 24*, 1129-1135.

Cools, R., Ivry, R. B., & D'Esposito, M. (2006). The human striatum is necessary for responding to changes in stimulus relevance. *Journal of Cognitive Neuroscience, 18*, 1973-1983.

Corcoran, K. A., Desmond, T. J., Frey, K. A., & Maren, S. (2005). Hippocampal inactivation disrupts the acquisition and contextual encoding of fear extinction. *Journal of Neuroscience, 25,* 8978-8987.

Corcoran, K. A., & Maren, S. (2001). Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. *Journal of Neuroscience, 21,* 1720-1726.

Cornelissen, F. W., Kimmig, H., Schira, M., Rutschmann, R. M., Maguire, R. P., Broerse, A. et al. (2002). Event-related fMRI responses in the human frontal eye fields in a randomized pro- and antisaccade task. *Experimental Brain Research, 145,* 270-274.

Crawford, T. J., Puri, B. K., Nijran, K. S., Jones, B., Kennard, C., & Lewis, S. W. (1996). Abnormal saccadic distractibility in patients with schizophrenia: a 99m Tc-HMPAO SPET study. *Psychological Medicine, 26,* 265-278.

Crawford, T. J., Sharma, T., Puri, B. K., Murray, R. M., Berridge, D. M., & Lewis, S. W. (1998). Saccadic eye movements in families multiply affected with schizophrenia: the Maudsley family study. *American Journal of Psychiatry, 155,* 1703-1710.

Davidson, R. J., Jackson, D. C., & Kalin, N. H. (2000). Emotion, plasticity, context, and regulation: perspectives from affective neuroscience. *Psychological Bulletin, 126*, 890-909.

Davidson, R. J., Pizzagalli, D., Nitschke, J. B., & Putnam, K. (2002). Depression: Perspectives from affective neuroscience. *Annual Review of Psychology*, *53*, 545-574.

Davis, M. (1994). The role of the amygdala in emotional learning. *International Review of Neurobiology, 36,* 225-266.

Demakis, G. J. (2003). A meta-analytic review of the sensitivity of the Wisconsin Card Sorting Test to frontal and lateralized frontal brain damage. *Neuropsychology, 17,* 255-264.

Depue, B. E., Curran, T., & Banich, M. T. (2007). Suppression of emotional memories via a two-phase process. *Science, 317*, 215-219.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996a). Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behavioral Neuroscience, 110*, 872-886.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996b). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380,* 69-72.

Dias, R., Robbins, T. W., & Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *Journal of Neuroscience, 17,* 9285-9297.

Dorahy, M. J. A critical evaluation of cognitive inhibition in dissociative identity disorder as inferred by negative priming in the flanker task: limitations and the episodic retrieval alternative. *Applied and Preventive Psychology* (2007), doi: 10.1016/j.appsy.2007.09.005

Durston, S., Tottenham, N. T., Thomas, K. M., Davidson, M. C., Eigsti, I. M., Yang, Y. et al. (2003). Differential patterns of striatal activation in young children with and without ADHD. *Biological Psychiatry, 53,* 871-878.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, 8, 1784-1790,

Ford, K. A., Goltz, H. C., Brown, M. R. G., & Everling, S. (2005). Neural processes associated with antisaccade task performance investigated with event-related fMRI. *Journal of Neurophysiology, 94,* 429-440.

Freedman, M., Black, S., Ebert, P., & Binns, M. (1998). Orbitofrontal function, object alternation and perseveration. *Cerebral Cortex*, 8, 18-27.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of Experimental Psychology: General, 133,* 101-135.

Fukushima, J., Fukushima, K., Chiba, T., Tanaka, S., Yamashita, I., & Kato, M. (1988). Disturbances of voluntary control of saccadic eye movements in schizophrenic patients. *Biological Psychiatry, 23,* 670-677.

Garavan, H., Hester, R., Murphy, K., Fassbender, C., & Kelly, C. (2006). Individual differences in the functional neuroanatomy of inhibitory control. *Brain Research, 1105,* 130-142.

Garavan, H., Ross, T. J., & Stein, E. A. (1999). Right hemispheric dominance of inhibitory control: An event-related functional MRI study. *Proceedings of the National Academy of Sciences, 96,* 8301-8306.

Gilbertson, M. W., Shenton, M. E., Ciszewski, A., Kasai, K., Lasko, N. B., Orr, S. P. et al. (2002). Smaller hippocampal volume predicts pathologic vulnerability to psychological trauma. *Nature Neuroscience, 5,* 1242-1247.

Goeleven, E., De Raedt, R., Baert, S., & Koster, E. H. (2006). Deficient inhibition of emotional information in depression. *Journal of Affective Disorders, 93*, 149-157.

Gottfried, J. A., & Dolan, R. J. (2004). Human orbitofrontal cortex mediates extinction learning while accessing conditioned representations of value. *Nature Neuroscience, 7,* 1144-1152.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18,* 1279-1296.

Hampshire, A., & Owen, A. M. (2006). Fractionating attentional control using event-related fMRI. *Cerebral Cortex, 16*, 1679-1689.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175-204). San Diego, CA: Academic Press.

Hikosaka, O., Takikawa, Y., & Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological Reveiws*, *80*, 953-978.

Hofmann, S. G., Meuret, A. E., Smits, J. A. J., Simon, N. M., Pollack, M. H., Eisenmenger, K. et al. (2006). Augmentation of exposure therapy with d-cycloserine for social anxiety disorder. *Archives of General Psychiatry, 63,* 298-304.

Holland, P. C. (1992). Occasion setting in Pavlovian conditioning. In D.L. Medlin (Ed.), *The psychology of learning and motivation* (pp. 69-125). New York: Academic Press.

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E. T., Morris, R. G., Bullock, P. R. et al. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience, 16,* 463-478.

Hutton, S. B., & Ettinger, U. (2006). The antisaccade task as a research tool in psychopathology: A critical review. *Psychophysiology, 43,* 302-313.

Joormann, J. (2004). Attentional bias in dysphoria: The role of inhibitory processes. *Cognition and Emotion, 18*, 125-147.

Joormann, J., Yoon, K. L., & Zetsche, U. Cognitive inhibition in depression. *Applied and Preventive Psychology* (2007), doi: 10.1016/j.appsy.2007.09.002

Konishi, S., Nakajima, K., Uchida, I., Kameyama, M., Nakahara, K., Sekihara, K. et al. (1998). Transient activation of inferior prefrontal cortex during cognitive set shifting. *Nature Neuroscience, 1,* 80-84.

Konishi, S., Nakajima, K., Uchida, I., Kikyo, H., Kameyama, M., & Miyashita, Y. (1999). Common inhibitory mechanism in human inferior prefrontal cortex revealed by event-related functional MRI. *Brain, 122,* 981-991.

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

Ledgerwood, L., Richardson, R., & Cranney, J. (2003). D-cycloserine facilitates extinction of conditioned fear as assessed by freezing in rats. *Behavioral Neuroscience, 117,* 341-349.

LeDoux, J. E. (1995). Emotion: clues from the brain. *Annual Review of Psychology, 46,* 209-235.

Li, C. S., Huang, C., Constable, R. T., & Sinha, R. (2006). Gender differences in the neural correlates of response inhibition during a stop signal task. *Neuroimage, 32,* 1918-1929.

Liddle, P. F., Kiehl, K. A., & Smith, A. M. (2001). Event-related fMRI study of response inhibition. *Human Brain Mapping, 12,* 100-109.

Liston, C., Watts, R., Tottenham, N., Davidson, M. C., Niogi, S., Ulug, A. M et al. (2006). Frontostriatal microstructure modulates efficient recruitment of cognitive control. *Cerebral Cortex,16,* 553-560.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: a theory of an act of control. *Psychological Review, 91,* 295-327.

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: a model and a method. *Journal of Experimental Psychology: Human Perception & Performance, 10,* 276-291.

Lucki, I. (1998). The spectrum of behaviors influenced by serotonin. *Biological Psychiatry, 44,* 151-162.

MacDonald, P. A., & Joordens, S. (2000). Investigating a memory-based account of negative priming: support for selection-feature mismatch. *Journal of Experimental Psychology, 26,* 1478-1496.

MacLeod, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review, 6,* 123-129.

MacLeod, C. M, Dodd, M. D., Sheard, E. D., Wilson, D. E., & Bibi, U. (2003). In opposition to inhibition. In B. H. Ross (Ed.), *The psychology of learning and memory*, Vol. 43 (pp. 163-214). San Diego, CA: Academic Press.

Massen, C. (2004). Parallel programming of exogenous and endogenous components in the antisaccade task. *Quarterly Journal of Experimental Psychology, 57,* 475-498.

Matsuda, T., Matsuura, M., Ohkubo, T., Ohkubo, H., Matsushima, E., Inoue, K. et al. (2004). Functional MRI mapping of brain activation during visually guided saccades and antisaccades: cortical and subcortical networks. *Psychiatry Research: Neuroimaging, 131,* 147-155.

May, C. P., Kane, M. J., & Hasher, L. (1995). Determinants of negative priming. *Psychological Bulletin, 118,* 35-54.

McGaugh, J. L., & Roozendaal, B. (2002). Role of adrenal stress hormones in forming lasting memories in the brain. *Current Opinion in Neurobiology, 12,* 205-210.

Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Human Brain Mapping, 12,* 131-143.

Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory structure, and mental processes. *Science*, *192*, 27-33.

Milad, M. R., Quinn, B. T., Pitman, R. K., Orr, S. P., Fischl, B., & Rauch, S. L. (2005). Thickness of ventromedial prefrontal cortex in humans is correlated with extinction memory. *Proceedings of the National Academy of Sciences, 102,* 10706-10711.

Milad, M. R., & Quirk, G. J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature, 420,* 70-74.

Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., & Rauch, S. L. (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biological Psychiatry, 62*, 446 – 454.

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

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology, 9,* 100-110.

Minas, R. K., & Park, S. Attentional window in schizophrenia and schizotypal personality: insight from negative priming studies. *Applied and Preventive Psychology* (2007), doi: 10.1016/j.appsy.2007.09.003

Mitchell, J. P., Macrae, C. N., & Gilchrist, I. D. (2002). Working memory and the suppression of reflexive saccades. *Journal of Cognitive Neuroscience, 14,* 95-103.

Monchi, O., Petrides, M., Doyon, J., Postuma, R. B., Worsley, K., & Dagher, A. (2004). Neural bases of set-shifting deficits in Parkinson's disease. *Journal of Neuroscience, 24,* 702-710.

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *Journal of Neuroscience, 21,* 7733-7741.

Morgan, M. A., Romanski, L. M., & LeDoux, J. E. (1993). Extinction of emotional learning: contribution of medial prefrontal cortex. *Neuroscience Letters, 163,* 109-113.

Morita, M., Nakahara, K., & Hayashi, T. (2004). A rapid presentation event-related functional magnetic resonance imaging study of response inhibition in macaque monkeys. *Neuroscience Letters, 356,* 203-206.

Munoz, D. P., & Everling, S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience, 5,* 218-228.

Myers, K. M., & Davis, M. (2007). Mechanisms of fear extinction. *Molecular Psychiatry, 12,* 120-150.

Neill, W. T., & Valdes, L. A. (1992). Persistence of negative priming: Steady-state or decay? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 565-576.

Nielson, K. A., Langenecker, S. A., & Garavan, H. (2002). Differences in the functional neuroanatomy of inhibitory control across the adult life span. *Psychology and Aging, 17,* 56-71.

Nieuwenhuis, S., Broerse, A., Nielen, M. M. A., & Jong, R. (2004). A goal activation approach to the study of executive function: An application to antisaccade tasks. *Brain and Cognition, 56,* 198-214.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bullettin, 126,* 220-246.

Nigg, J. T., & Casey, B. J. (2005). An integrative theory of attention-deficit/hyperactivity disorder based on the cognitive and affective neurosciences. *Development & Psychopathology, 17,* 785-806.

Nigg, J. T., Silk, K. R, Stavro, G., & Miller, T. (2005). Disinhibition and borderline personality disorder. *Development & Psychopathology, 17,* 1129-1149.

O'Driscoll, G. A., Alpert, N. M., Matthysse, S. W., Levy, D. L., Rauch, S. L., & Holzman, P. S. (1995). Functional neuroanatomy of antisaccade eye movements investigated with positron emission tomography. *Proceedings of the National Academy of Sciences, 92,* 925-929.

O'Driscoll, G. A., Lenzenweger, M. F., & Holzman, P. S. (1998). Antisaccades and smooth pursuit eye tracking and schizotypy. *Archives of General Psychiatry, 55,* 837-843.

Pandya, D. N., & Barnes, C. L. (1987). Architecture and connections of the frontal lobe. In E. Perecman (Ed.), *The frontal lobes revisited* (pp. 41-72). New York: IRBN Press.

Park, J., & Kanwisher, N. (1994). Negative priming for spatial locations: Identity mismatching, not distractor inhibition. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 613-623.

Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience—virtual lesion, chronometry, and functional connectivity. *Current Opinion in Neurobiology, 10,* 232-237.

Petrides, M. (2000). The role of the mid-dorsolateral prefrontal cortex in working memory. *Experimental Brain Research, 133,* 44-54.

Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience, 16*, 291-310.

Pezawas, L., Meyer-Lindenberg, A., Drabant, E. M., Verchinski, B. A., Munoz, K. E., Kolachana, B. S. et al. (2005). 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nature Neuroscience, 8,* 828-834.

Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: role of the amygdala and vmPFC. *Neuron, 43,* 897-905.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception II: Implications for major psychiatric disorders. *Biological Psychiatry, 54,* 515-528.

Pierrot-Deseilligny, C., Muri, R. M., Ploner, C. J., Gaymard, B., Demeret, S., & Rivaud-Pechoux, S. (2003). Decisional role of the dorsolateral prefrontal cortex in ocular motor behaviour. *Brain, 126,* 1460-1473.

Pierrot-Deseilligny, C. H., Rivaud, S., Gaymard, B., & Agid, Y. (1991). Cortical control of reflexive visually-guided saccades. *Brain, 114,* 1473-1485.

Ploner, C. J., Gaymard, B. M., Rivaud-Pechoux, S., & Pierrot-Deseilligny, C. (2005). The prefrontal substrate of reflexive saccade inhibition in humans. *Biological Psychiatry, 57,* 1159-1165.

Quirk, G. J. (2006). Extinction: new excitement for an old phenomenon. *Biological Psychiatry, 60,* 317-318.

Quirk, G. J., Likhtik, E., Pelletier, J. G., & Pare, D. (2003). Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *Journal of Neuroscience, 23,* 8800-8807.

Quirk, G. J., Russo, G. K., Barron, J. L., & Lebron, K. (2000). The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *Journal of Neuroscience, 20,* 6225-6231.

Raemaekers, M., Jansma, J. M., Cahn, W., Van der Geest, J. N., van der Linden, J. A., Kahn, R. S. et al. (2002). Neuronal substrate of the saccadic inhibition deficit in schizophrenia investigated with 3-Dimensional event-related functional magnetic resonance imaging. *Archives of General Psychiatry, 59,* 313-320.

Raemaekers, M., Ramsey, N. F., Vink, M., van den Heuvel, M. P., & Kahn, R. S. (2006). Brain activation during antisaccades in unaffected relatives of schizophrenic patients. *Biological Psychiatry, 59,* 530-535.

Rauch, S.L, Milad, M.R., Orr, S.P, Quinn, B.T., Fischl, B., & Pitman, R.K. (2005). Orbitofrontal thickness, retention of fear extinction, and extraversion. *Neuroreport, 16*, 1909-1912.

Rauch, S. L., Shin, L. M., & Phelps, E. A. (2006). Neurocircuitry models of posttraumatic stress disorder and extinction: human neuroimaging research—past, present, and future. *Biological Psychiatry, 60*, 376-382.

Rescorla, R. A. (2004). Spontaneous recovery. *Learning & Memory, 11,* 501-509.

Rescorla, R. A., & Heth, C. D. (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes, 1,* 88-96.

Ressler, K. J., Rothbaum, B. O., Tannenbaum, L., Anderson, P., Graap, K., Zimand, E. et al. (2004). Cognitive enhances as adjuncts to psychotherapy: use of d-cycloserine in phobics to facilitate extinction to fear. *Archives of General Psychiatry, 61*, 1136-1144.

Robbins, T. W. (2007). Shifting and stopping: fronto-striatal substrates, neurochemical modulation and clinical implications. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences, 362*, 917-932.

Rolls, E. T. (1996). The orbitofrontal cortex. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences, 351,* 1433-1443.

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

Rorden, C., & Karnath, H. O. (2004). Using human brain lesions to infer function: A relic from a past era in the fMRI age? *Nature Review Neuroscience 5*, 813-819.

Rubia, K., Russell, T., Overmeyer, S., Brammer, M. J., Bullmore, E. T., Sharma, T. et al. (2001). Mapping motor inhibition: conjunctive brain activations across different versions of Go/No-Go and Stop tasks. *Neuroimage, 13,* 250-261.

Rubia, K., Smith, A. B., Brammer, M. J., & Taylor, E. (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *Neuroimage, 20,* 351-358.

Rubia, K., Smith, A. B., Woolley, J., Nosarti, C., Heyman, I., Taylor, E. et al. (2006). Progressive increase of frontostriatal brain activation from childhood to adulthood during event-related tasks of cognitive control. *Human Brain Mapping, 27*, 973-993.

Sakagami, M., Tsutsui, K., Lauwereyns, J., Koizumi, M., Kobayashi, S., & Hikosaka, O. (2001). A code for behavioral inhibition on the basis of color, but not motion, in ventrolateral prefrontal cortex of macaque monkey. *Journal of Neuroscience, 21,* 4801-4808.

Sasaki, K., Gemba, H., & Tsujimoto, T. (1989). Suppression of visually initiated hand movement by stimulation of the prefrontal cortex in the monkey. *Brain Research, 495,* 100-107.

Schafe, G. E., Nader, K., Blair, H. T., & LeDoux, J. E. (2001). Memory consolidation of Pavlovian fear conditioning: a cellular and molecular perspective. *Trends in Neurosciences, 24,* 540-546.

Schlag-Rey, M., Amador, N., Sanchez, H., & Schlag, J. (1997). Antisaccade performance predicted by neuronal activity in the supplementary eye field. *Nature, 390,* 398-401.

Sereno, A. B., & Holzman, P. S. (1995). Antisaccades and smooth pursuit eye movements in schizophrenia. *Biological Psychiatry, 37,* 394-401.

Sotres-Bayon, F., Bush, D. E. A., & LeDoux, J. E. (2004). Emotional perseveration: An update on prefrontal-amygdala interactions in fear extinction. *Learning & Memory, 11,* 525-535.

Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience, 24,* 8223–8231.

Soubrié, P. (1986). Reconciling the role of central serotonin neurons in human and animal behavior. *Behavioral and Brain Sciences, 9,* 319-363.

Sullivan, E. V., Mathalon, D. H., Zipursky, R. B., Kersteen-Tucker, Z., Knight, R. T., & Pfefferbaum, A. (1993). Factors of the Wisconsin Card Sorting Test as measures of frontal-lobe function in schizophrenia and in chronic alcoholism. *Psychiatry Research, 46,* 175-199.

Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration. *Quarterly Journal of Experimental Psychology, 54A,* 321-343.

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30,* 1302-1321.

Urry, H. L., van Reekum, C. M., Johnstone, T., Kalin, N. H., Thurow, M. E., Schaefer, H. S. et al. (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. *Journal of Neuroscience, 26,* 4415-4425.

Walker, D. L., Ressler, K. J., Lu, K. T., & Davis, M. (2002). Facilitation of conditioned fear extinction by systemic administration or intra-amygdala infusions of d-cycloserine as assessed with fear-potentiated startle in rats. *Journal of Neuroscience, 22,* 2343-2351.

Woldorff, M. G., Hazlett, C. J., Fichtenholtz, H. M., Weissman, D. H., Dale, A. M., & Song, A. W. (2004). Functional parcellation of attentional control regions of the brain. *Journal of Cognitive Neuroscience, 16,* 149-165.
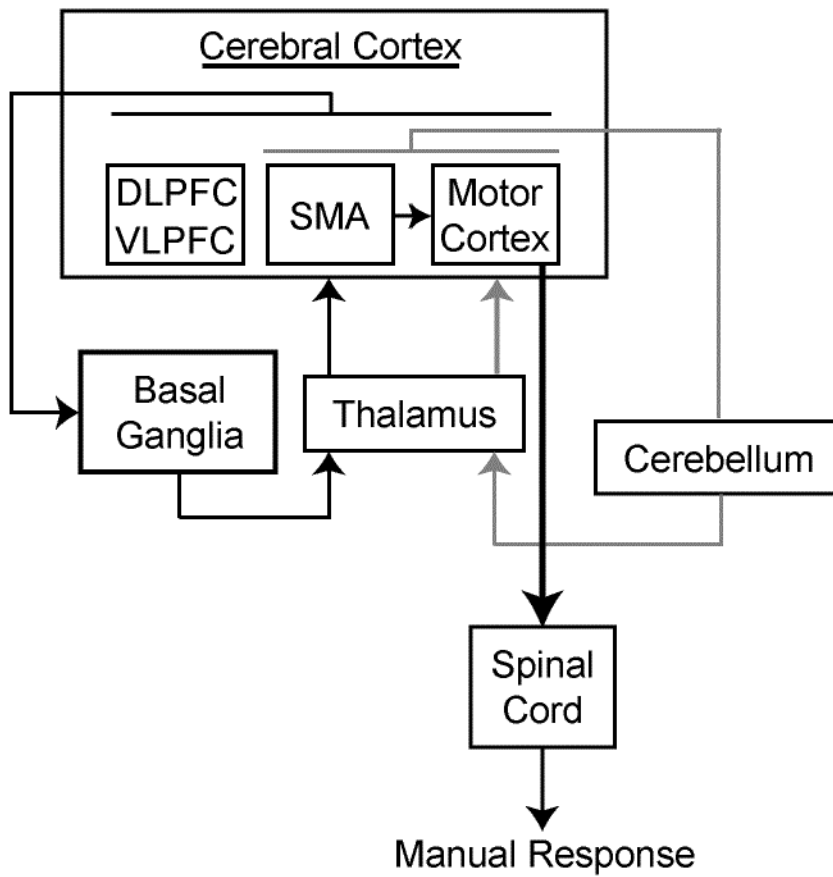
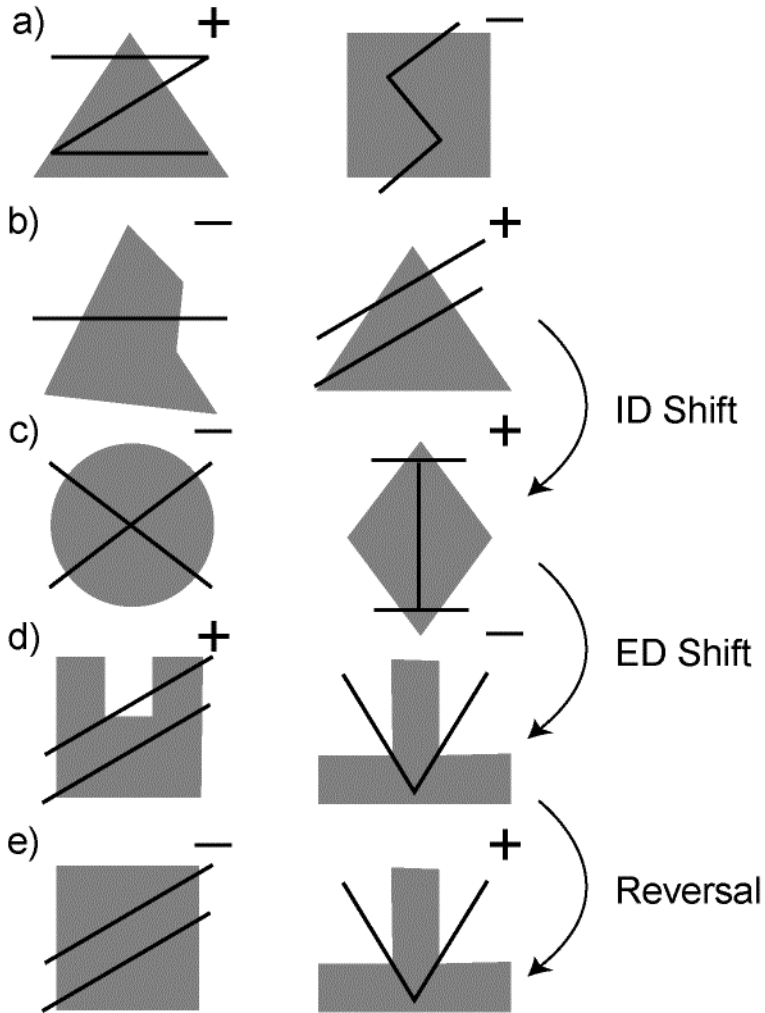Figure 1

a)

b)
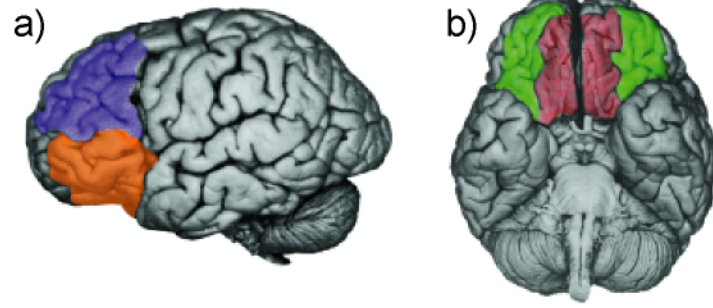
PFC: prefrontal cortex

Figure 2
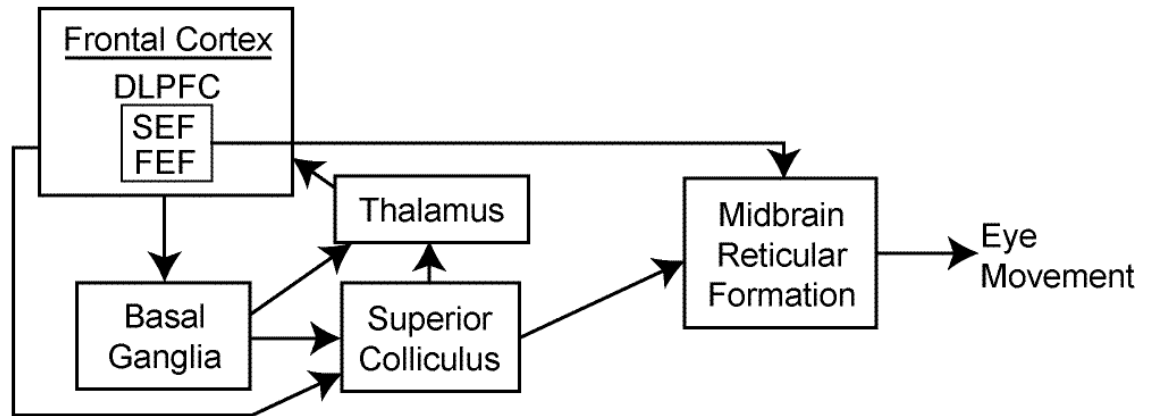
DLPFC: dorsolateral prefrontal cortex; FEF: frontal eye fields; SEF: supplementary eye fields
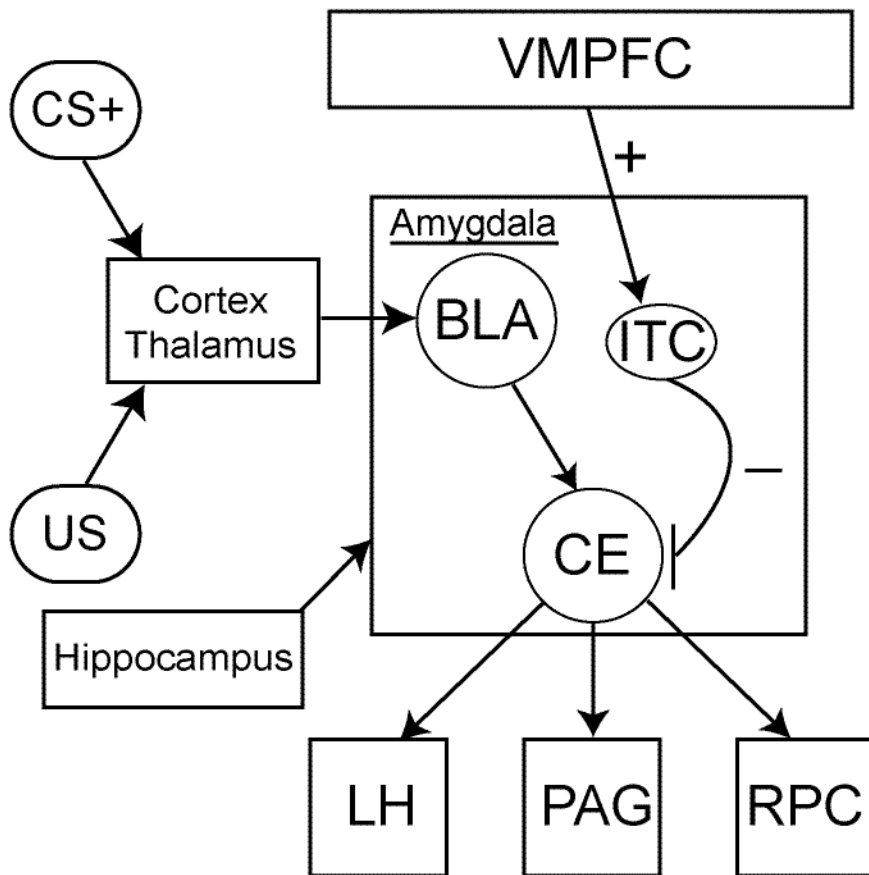
Figure 3

DLPFC: dorsolateral prefrontal cortex; SMA: supplementary motor area; VLPFC: ventrolateral prefrontal cortex.

Figure 4

ID: Intra-dimensional; ED: Extra-dimensional.

Figure 5

BLA: basolateral amygdaloid complex; CE: central nucleus;
CS+: conditioned stimulus; LH: lateral hypothalamus;
ITC: intercalated cell masses; PAG: periaqueductal gray; RPC:
reticularis pontis caudalis; US: unconditioned stimulus; VMPFC:
ventromedial prefrontal cortex.