Temporal Difference Modeling of the Blood-Oxygen Level Dependent Response During Aversive Conditioning in Humans: Effects of Dopaminergic Modulation

Mahesh Menon, Jimmy Jensen, Irina Vitcu, Ariel Graff-Guerrero, Adrian Crawley, Mark A. Smith, and Shitij Kapur

Background: The prediction error (PE) hypothesized by the temporal difference model has been shown to correlate with the phasic activity of dopamine neurons during reward learning and the blood-oxygen level dependent (BOLD) response during reward and aversive conditioning tasks. We hypothesized that dopamine would modulate the PE related signal in aversive conditioning and that haloperidol would reduce PE related activity, while an acute dose of amphetamine would increase PE related activity in the ventral striatum.

Methods: Healthy participants took an acute dose of amphetamine, haloperidol, or placebo. We used functional magnetic resonance imaging (fMRI) to measure the BOLD signal while they carried out an aversive conditioning task, using cutaneous electrical stimulation as the unconditioned stimulus (US) and yellow and blue circles as conditioned stimulus (CS+ and CS−, respectively).

Results: Prediction error related BOLD activity was seen only in the ventral striatum in the placebo subjects. The subjects given amphetamine showed a wider network of PE related BOLD activity, including the ventral striatum, globus pallidus, putamen, insula, anterior cingulate, and substantia nigra/ventral tegmental area. Haloperidol subjects did not show PE related activity in any of these regions.

Conclusions: Our results provide the first demonstration that the modulation of dopamine transmission affects both the physiological correlates and PE related BOLD activity during aversive learning.

Key Words: Amphetamine, aversive conditioning, fMRI, haloperidol, reward learning

Studies in animals (Schultz 1998) have shown that phasic firing of midbrain dopamine neurons into the striatum signals the presence of appetitive stimuli. The dopamine reward signal is supplemented by activity in neurons in the striatum, frontal cortex, and amygdala, which process specific reward information but do not emit a global reward prediction error signal. The link between dopaminergic neurons and reward prediction is best demonstrated in conditioning tasks where a previously neutral stimulus (conditioned stimulus [CS]), through its pairing with the unconditioned reward stimuli (US), comes to predict the onset of the US. As the animals learn the task, the onset of the phasic dopamine response shifts from the onset of the US to the onset of the CS. This firing pattern can be modeled using the temporal difference (TD) model (Sutton and Barto 1990), a machine-learning approach to reinforcement learning that predicts the amount of future expected reward associated with events. When an event occurs, the difference between actual and predicted reward is signaled as a prediction error (PE), and this error is used to update the predictions accordingly. Unlike other models of learning (e.g., Rescorla-Wagner model, Rescorla and Wagner, 1972), which also use prediction errors, the TD model considers the temporal ordering of events and the shift in prediction error signal from the US onset to the CS onset.

The TD algorithm has been used in human functional magnetic resonance imaging (fMRI) studies to examine the brain activation patterns involved in the prediction of reward (McClure *et al.* 2003; O'Doherty *et al.* 2004, 2006) and punishment (Seymour *et al.* 2004). McClure *et al.* (2003) and O'Doherty *et al.* (2003) used an appetitive reward-learning paradigm (delivery of juice) and showed that the fMRI blood-oxygen level dependent (BOLD) response in the striatum correlated with the PE predicted by the TD model. Seymour *et al.* (2004) examined the pattern of brain responses in humans during a second-order pain-learning task and found that the PE correlated with the BOLD activity seen in the ventral striatum (VS) and the insula over the course of this higher-order learning paradigm. The results indicate that the prediction error generated by the TD algorithm correlates with the BOLD response seen in the ventral striatum in both appetitive and aversive learning. When the stimuli are aversive and particularly pain related, other regions such as the insula, the orbitofrontal cortex, and the anterior cingulate cortex also show activity in a manner predicted by the TD algorithm (Seymour *et al.* 2005). Consistent with this, a recent study (Jensen *et al.* 2006) found that in a design where cues were conditioned to both appetitive and aversive cues, the prediction error related BOLD response in the ventral striatum signals the salience rather than the valence of the cues, i.e., there is an increased BOLD activity in the striatum both to the cues signaling aversive and appetitive events, while differential activity was observed in the anterior insula and orbitofrontal cortex, respectively. This is also in keeping with data suggesting that these dopamine neurons predict and detect not only rewards but also signal salient, alerting and motivating events (Horvitz 2000; Redgrave *et al.* 1999; Young and Rees 1998).

From the Schizophrenia Program and PET Centre (MM, JJ, IV, AG-G, SK), Centre for Addiction and Mental Health; and Departments of Medical Imaging (AC) and Psychiatry (SK), University of Toronto, Toronto, Canada; Department of Psychiatry (JJ), Ullevål University Hospital, Oslo University, Oslo, Norway; and Discovery Medicine (MAS), Astra-Zeneca, Wilmington, Delaware.

Address reprint requests to Shitij Kapur, M.D., Ph.D., Centre for Addiction and Mental Health, 33 Russel Street, Toronto M5S 2S1, Canada; E-mail: Shitij_Kapur@camh.net.

Received June 27, 2006; revised October 14, 2006; accepted October 16, 2006.

0006-3223/07/$32.00
doi:10.1016/j.biopsych.2006.10.020

As the phasic firing pattern of the dopamine neurons in response to rewarding stimuli has been shown to correlate with the PE generated by the TD model in nonhuman primates (Schultz 1998) and this PE has been shown to correlate with BOLD activity in response to both aversive and appetitive conditioned stimuli in humans (Jensen *et al.* 2006; O'Doherty *et al.* 2003, 2006; Seymour *et al.* 2004), it is quite likely that BOLD response in the striatum is affected by the midbrain dopaminergic output. However, the exact role of dopamine in this process is yet to be examined. If it is the case that the BOLD response is related to the midbrain firing of dopaminergic neurons, we propose that an acute dose of amphetamine, an indirect dopamine agonist that causes increased dopamine release in the striatum (Jones *et al.* 1998), would result in increased prediction error related BOLD response in that region, as speculated by Montague *et al.* (1996), while haloperidol, which blocks postsynaptic striatal dopamine receptors, would result in reduced prediction error related BOLD activity in the striatum.

To test this hypothesis, we used fMRI to examine BOLD response in healthy volunteers who were given an acute dose of amphetamine, haloperidol, or a placebo and then took part in an aversive conditioning paradigm. We were particularly interested in looking at the activity in the midbrain dopamine regions, the substantia nigra (SN) and the ventral tegmental area (VTA), and the striatal regions that they project to (i.e., the ventral striatum and nucleus accumbens, caudate, putamen, and the globus pallidus), as these regions have all been shown to respond to the presence of salient stimuli (Zink *et al.* 2004) and in aversive and appetitive conditioning tasks (O'Doherty *et al.* 2003; Seymour *et al.* 2004). In addition, Seymour *et al.* (2005) also found activity in the insula, the anterior cingulate cortex, and orbitofrontal cortex in their aversive conditioning paradigms. We therefore included these as regions of primary interest and predicted that they would show PE related activity during the aversive conditioning task and that these activations would be enhanced and diminished by pharmacological enhancement and blockade of the dopamine system, respectively.

Methods and Materials

Participants

The study and the recruitment procedures were approved by Health Canada and the Research Ethics Board of the Centre for Addiction and Mental Health (CAMH). All subjects gave written informed consent.

The placebo group consisted of 17 subjects (12 male subjects, 5 female subjects), while the amphetamine and haloperidol groups each consisted of 15 subjects (amphetamine group: 7 male subjects, 8 female subjects; haloperidol group: 11 male subjects, 4 female subjects).

Participants were included if they were between 18 and 50 years of age, had no history of psychiatric illness as confirmed by the Hopkins Symptoms Checklist (Derogatis *et al.* 1974), had not taken any psychotropic medication in the previous 2 years, and had been free of any drugs or medication (except caffeine and nicotine) in the 2 weeks preceding the test. Urine drug screening and urine pregnancy tests (female subjects only) were carried out during the screening visit and on the day of the experiment. The dose of the drug was based on the participant's body weight (.04 mg/kg haloperidol, .4 mg/kg amphetamine). All tablets were overencapsulated to ensure that subjects were blinded to dosage and could not compare or identify their drug. The amphetamine participants were scanned 1 hour after drug intake, placebo subjects were scanned 2 hours after drug intake, and haloperidol subjects were scanned 5 hours after drug intake to maximize drug concentration in the bloodstream (Angrist *et al.* 1987; Midha *et al.* 1989).

Experimental Protocol

The paradigm was based on classical aversive conditioning using a 33% partial reinforcement schedule with yellow and blue circles as cues (CS+ and CS−, respectively). The aversive unconditioned stimulus used was a 200 msec cutaneous electrical stimulation applied to the left index finger. The intensity of the US was titrated individually, prior to entering the scanner, until it reached a level that the participants described as "unpleasant but tolerable." Cues appeared in a pseudorandom order for a duration of 5 sec each. The US immediately followed the offset of the CS+ circle in 33% ($CS+_{paired}$) of the trials. The CS− circle had no consequences. A fixation cross was presented between trials and a fixed intertrial interval of 8.8 sec was used. The subjects were told before the experiment began that one of the two circles would sometimes be followed by the electrical stimulation but were not informed about the reinforcement schedule or which of the cues were CS+ and CS−. The first four trials consisted of two $CS+_{paired}$ trials, each followed by a CS− trial. These were followed by 75 randomized trials: 15 CS+ presentations paired with the US ($CS+_{paired}$), 30 CS+ trials without US (CS+), and 30 CS− trials.

The paradigm lasted approximately 18 minutes. Immediately after the participants left the scanner, they completed a post-test questionnaire, where they were asked to identify which colored circle had been associated with the shock, subjective anxiety in response to the cues, and how frequently the shock was associated with each color. Participants were considered to have learned the task if they correctly identified that the shock occurred with the yellow circle and rated their anxiety on seeing the blue circle as less than (or equal to) that of seeing the yellow. Their subjective state was measured using the subjective state questionnaire (SSQ), which consists of descriptive adjectives (e.g., calm, alert, sleepy, etc) with visual analogue scales below them (labeled from "not at all" to extremely"), on which participants placed a vertical tick to mark the magnitude of their current mood or state. The SSQ consists of four factor-analyzed measures: pleasant stimulation, unpleasant stimulation, pleasant sedation, and unpleasant sedation.

As a measure of autonomic arousal, we measured the participant's skin conductance response during the conditioning task.

Apparatus

Electrical Stimulation. The electrical stimulation was delivered by a stimulating bar electrode (Chalgren Enterprises, Gilroy, California) using a conducting gel electrolyte. The electrode was connected to a Grass Instruments SD-9 stimulator (Grass-Telefactor, West Warwick, Rhode Island) via well-isolated coaxial cable leads through a radio frequency (RF) filter.

Stimulus Presentation. The E-prime software (Psychology Software Tools, Inc., Pittsburgh, Pennsylvania) controlled the stimulus presentations and triggered the stimulator.

Galvanic Skin Conductance. The galvanic skin conductance (GSR) was measured at 10 Hz using magnetic resonance imaging (MRI)-compatible silver/silver chloride (Ag/AgCl) electrodes, which were attached to the left middle and ring fingers. The output was continuously monitored by PowerLab 2/20 (AD Instruments, Castle Hill, Australia) via long, well-isolated cables passed through an RF filter.