



Published in final edited form as:

Neuroimage. 2011 May 1; 56(1): 315–322. doi:10.1016/j.neuroimage.2011.01.068.

Human Reversal Learning Under Conditions of Certain versus Uncertain Outcomes

Anna-Maria D’Cruz, M.Phys., M.A.^{a,b}, Michael E. Ragozzino, Ph.D.^{a,b}, Matthew W. Mosconi, Ph.D.^a, Mani N. Pavuluri, M.D., Ph.D.^{a,c}, and John A. Sweeney, Ph.D.^{a,b}

Anna-Maria D’Cruz: adcruz@psych.uic.edu; Michael E. Ragozzino: mrago@uic.edu; Matthew W. Mosconi: mmosconi@psych.uic.edu; Mani N. Pavuluri: mpavuluri@psych.uic.edu; John A. Sweeney: jsweeney@psych.uic.edu

^aCenter for Cognitive Medicine, 912 S.Wood St., Suite 235, Chicago, IL 60612, USA

^bDepartment of Psychology, 1007 W.Harrison St., Chicago, IL 60607, USA

^cPediatric Brain Research and Intervention Center, 1747 W.Roosevelt Rd., Chicago, IL 60608, USA

Abstract

Reversal learning tasks assess behavioral flexibility by requiring subjects to switch from one learned response choice to a different response choice when task contingencies change. This requires both the processing of negative feedback once a learned response is no longer reinforced, and the capacity for flexible response selection. In 2-choice reversal learning tasks, subjects switch between only two responses. Multiple choice reversal learning is qualitatively different in that at reversal, it requires subjects to respond to nonreinforcement of a learned response by selecting a new response from among several alternatives that have uncertain consequences. While activity in brain regions responsible for processing unexpected negative feedback is known to increase in relation to the hedonic value of the reward itself, it is not known whether the uncertainty of reinforcement for future response choices also modulates these responses. In an fMRI study, 15 participants performed 2- and 4-choice reversal learning tasks. Upon reversal in both tasks, activation was observed in brain regions associated with processing changing reinforcement contingencies (midbrain, ventral striatum, insula), as well as in neocortical regions that support cognitive control and behavioral planning (prefrontal, premotor, posterior parietal, and anterior cingulate cortices). Activation in both systems was greater in the 4- than in the 2-choice task. Therefore, reinforcement uncertainty for future responses enhanced activity in brain systems that process performance feedback, as well as in areas supporting behavioral planning of future response choices. A mutually facilitative integration of responses in motivational and cognitive brain systems might enhance behavioral flexibility and decision making in conditions for which outcomes for future response choices are uncertain.

Keywords

reversal learning; human; fMRI; cognitive flexibility; reinforcement learning; reward processing

Corresponding author: John A. Sweeney, Ph.D., Phone: (312)-413-9205, Fax: (312)-413-8837.

Conflicts of interest: Dr. Sweeney has consulted to Pfizer and receives grant support from Janssen. Dr. Pavuluri is on the speaker’s bureaus for Ortho-McNeil and Bristol Myers Squibb.

Introduction

The ability to shift choice patterns under changing reinforcement contingencies represents a type of cognitive flexibility that is essential for daily living. Reduced cognitive flexibility is characteristic of several psychiatric and neurologic disorders, including schizophrenia (Floresco et al., 2009), obsessive-compulsive disorder (Rosenberg et al., 1997), autism (Mosconi et al., 2009), Parkinson's disease (Cools et al., 2007), and Huntington's disease (Lawrence et al., 1999). Clarifying the component processes that support cognitive flexibility and their neural substrates is thus important not only for neurocognitive models of behavioral control but for laying a foundation for clinical research as well.

Reversal learning tasks are a widely used tool for assessing behavioral flexibility (Ghahremani et al., 2010; Gläscher et al., 2009; Ragozzino et al., 2009). Such tasks assess the ability to acquire a behavioral strategy using performance feedback, and to change or “reverse” that response set to an alternative option when the previously rewarded response is no longer reinforced. In a standard 2-choice reversal learning task, participants are presented with two response options (Boulougouris and Robbins, 2009; McAlonan and Brown, 2003; Palencia and Ragozzino, 2006). After participants learn to choose the correct response over multiple trials using performance feedback, the reinforcement contingency is changed without warning. At this point, participants receive unexpected feedback that the previously reinforced response is incorrect, cueing them to switch to an alternative response. In these tasks, once one response is no longer correct, the alternative response is certain to be the correct choice. Previous functional imaging studies of 2-choice tasks have documented that a distributed network of brain regions subserves reversal learning including dorsolateral and ventrolateral prefrontal cortex, orbitofrontal cortex, and anterior cingulate cortex, as well as dorsal and ventral striatum (Ghahremani et al., 2010; O'Doherty et al., 2003). Studies of non-human primates and rodents provide parallel findings, and have clarified the neurotransmitter systems involved in reversal learning (Dias et al., 1996; Kim and Ragozzino, 2005).

Animal and human studies indicate that the ventral striatum is sensitive to changes in reinforcement contingencies (Gregorios-Pippas et al., 2009). When response-outcome contingencies change during reversal learning, subjects expecting ongoing positive reinforcement for learned responses instead receive unexpected nonreinforcement. This feedback elicits a response known as the negative reward prediction error signal in the nucleus accumbens, which is manifested by phasic changes in dopamine signaling (Schultz et al., 1997). Human neuroimaging studies have shown an increase of activation in the ventral striatum in response to unexpected negative feedback (Gläscher et al., 2010; Rolls et al., 2008). Error processing signals reflecting a violation of expectancies have also been identified in the human anterior cingulate cortex (Baker and Holroyd, 2009), a candidate region for integrating feedback-related information with action planning, and updating expectations about response-outcome contingencies (Hayden and Platt, 2010; Hillman and Bilkey, 2010).

Shifting behavior to a newly adaptive response during reversal learning also requires inhibition of a previously learned response and the planning of a new response choice. Inhibition of previously rewarded responses is modulated by the orbitofrontal cortex (Budhani et al., 2007). The dorsolateral prefrontal cortex also has been implicated in decisions to inhibit learned responses (Kenner et al., 2010; Velanova et al., 2008). The selection and implementation of alternative motor responses is dependent upon the dorsomedial striatum (Balleine et al., 2007), motor cingulate (Picard and Strick, 1996), and the supplementary motor area (Matsuzaka and Tanji, 1996).

Reversal learning has been examined in multiple-choice as well as 2-choice tasks in rodents using radial-arm mazes and T-mazes respectively (Dong et al., 2005; Hauber and Sommer, 2009). Tasks with multiple response options are fundamentally different from 2-choice tasks because after the response contingency changes, subjects are uncertain about what the next correct response choice will be. Response-outcome uncertainty at reversal might impact function in brain regions supporting feedback processing and those involved in decision making. Responses in brain systems supporting the processing of negative feedback cues may be increased because of the greater relevance of response feedback for future behavior. Activation in brain regions involved in response selection and decision making may also be enhanced because of the need to consider alternative response choices with uncertain outcomes at the reversal. To date, there has been no direct comparison of 2-choice versus multiple response option reversal learning tasks in human neuroimaging research.

In the present study, participants performed reversal learning paradigms with two or four response choices during functional neuroimaging studies. Simple visual stimuli that differed only in their spatial location, and a deterministic feedback schedule (i.e., 100% accurate feedback presented on all trials), were used to minimize demands from other cognitive processes, and to parallel paradigms used in animal models. We examined fMRI activation during flexible transitions in behavioral set under conditions in which the next correct response choice was either certain (2-choice task) or was uncertain (4-choice task), and had to be selected from a trial-and-error search of several options. The primary study aim was to ascertain whether uncertainty of future outcomes at reversal modulated brain activation in the two brain circuits supporting processes crucial to reversal learning, namely ventrostriatal feedback processing systems, and dorsal frontal decision making systems.

Materials and Methods

Fifteen healthy young adults (9 females; mean age 25.4 years, $SD=4.2$ years) were recruited from a sample of community volunteers. Exclusion criteria included any history of significant medical, neurological or psychiatric illness. Participants abstained from caffeine and nicotine for at least 2 h prior to scan sessions. Written informed consent was obtained from all participants. Study procedures were approved by the Institutional Review Board at the University of Illinois at Chicago.

fMRI paradigms

2-choice reversal learning task—Participants were presented with two identical stimuli (one stimulus each on the left and right side of the display screen) and instructed to select the stimulus that was in the correct location by pressing a button corresponding to its location on screen (Fig. 1). Participants held a button box with four buttons placed on their torso with both hands, and used the two outer buttons to indicate their response choice (left hand for stimulus on the left, and right hand for right stimulus choice). Immediate feedback was provided in the form of check marks (correct) or crosses (incorrect), which appeared directly above the stimulus selected until the end of the trial.

Requirements to change response set were imposed by making the other stimulus location the correct response choice. In order to reduce the predictability of the reversal, and therefore the predictability of receiving negative feedback on a given trial, the correct location changed after a variable number (four to six) of consecutive correct responses. Each trial (including presentation of stimulus, participant response, and feedback presentation) lasted for 2.5 s, followed by a 500 ms inter-trial interval during which a blank screen was presented. 180 trials were presented over a fixed task duration of 9 min.

4-choice reversal learning task—In the 4-choice task, participants were presented with four identical stimuli placed along the horizontal axis of the display screen (Fig. 1). They were told to choose the stimulus that was in the correct location, this time using all four response buttons. Two buttons were assigned to each hand. Each of the four stimulus locations had an equal probability of being the correct stimulus choice. The 4- and the 2-choice tasks were similar, with the following two exceptions. First, in order to reduce demands on working memory imposed by having to keep track of which locations were previously determined to be incorrect response choices, feedback for incorrect choices remained on screen until a correct choice was made. This was not deemed necessary in the 2-choice task, for which new response choices following reversal were always correct. Second, this paradigm incorporated a predetermined rate of incorrect trials at the point of reversal. When the correct stimulus location changed, the new correct response choice could be at one of the three alternative locations. To ensure similar rates of non-reinforcement amongst participants at the reversal, the first choice was correct on 15% of trials, the second choice was correct on 33% of trials, and the third and final choice was always correct. The mean duration of feedback presentation was 2.14 s (SD=0.28 s) and 2.13 s (SD=0.29 s) in the 2- and 4-choice tasks respectively. The 2- and 4-choice tasks were presented in counterbalanced order across participants. There was no effect of the order of task presentation on brain activity or behavioral measures reported.

MRI image acquisition

MRI studies were performed using a 3.0 T whole body scanner with a standard quadrature coil (Signa, General Electric Medical System, Milwaukee, WI). Functional images were acquired using a single shot gradient-echo echo-planar imaging sequence (15 axial slices; TR=1000 ms; TE=25 ms; flip angle=90°; slice thickness=5 mm; gap=1 mm; acquisition matrix=64×64; voxel size=3.12 mm×3.12 mm×5 mm; field of view (FOV)=20×20 cm²; 540 images). This protocol provided a field of view typically extending from the dorsal neocortex through most of the cerebellum, and therefore covered the neocortical and striatal regions of primary interest. Anatomical images were acquired with a 3D volume inversion recovery fast spoiled gradient-recalled at steady state pulse sequence (120 axial slices; flip angle=25°; slice thickness=1.5 mm; gap=0 mm; FOV=24×24 cm²).

Image preprocessing and analysis

Event-related fMRI analyses were carried out using FSL 4.1.0 (FMRIB Software Library; Smith et al., 2004) within the FEAT tool (fMRI Expert Analysis Tool). Brain Extraction Tool (BET) software was used to remove non-brain tissue from each participant's structural images (Smith, 2002). MCFLIRT motion correction was applied to functional datasets (Jenkinson et al., 2002). A high-pass temporal filter with a cut-off of 100 ms was applied to the data. Spatial smoothing was conducted using a Gaussian kernel of full-width half-maximum 6 mm. Functional data were registered first to the high-resolution structural scan and then transformed into standard MNI (Montreal Neurological Institute) space using the MNI152 template.

Modeling of activation responses

The time of onset of performance feedback, which immediately followed response choices, was used to identify trial-wise events of interest in the functional time-series data. Non-reinforcement at reversal elicited negative feedback processing, and cued the need for a new response choice. To control for differences in the number of reversals completed by each participant in the 2- and 4-choice tasks, the minimum number of sets completed for each participant in either of these two tasks (which did not differ significantly across participants) was used for both tasks; subsequent trials in the task with the greater number of sets

completed were dropped. The mean number of reversals was 26 and 23 in the 2-choice and 4-choice tasks respectively. Imaging results were similar when data from all reversals from all participants were included in the analyses. The following epochs of the time-series data were modeled in both the 2- and 4-choice reversal learning tasks: (1) the first instance of non-reinforcement for a learned response (indicating that participants' previous response set was no longer correct), and (2) expected reinforcement of correct responses (i.e. reinforcement of the second consecutive correct response and all later correct responses in a set). A double-gamma hemodynamic response function was applied to each model.

In order to examine brain activation related to processing unexpected non-reinforcement and planning a behavioral reversal, differences in response to unexpected non-reinforcement and expected reinforcement were contrasted separately for the 2- and 4-choice tasks. To identify differences in reversal learning when the outcomes of new response choices were certain and uncertain, the within-subjects difference in response to unexpected non-reinforcement relative to ongoing positive reinforcement was contrasted between the 2-choice and 4-choice tasks. For all fMRI analyses, Z-statistic images were thresholded at $Z > 2.5$ and evaluated for statistical significance using a cluster threshold that preserved an experiment-wise Type 1 error rate of $p < .05$ using Gaussian random field theory.

Results

Imaging results

Non-reinforcement of learned responses relative to expected reinforcement of correct responses led to similar regions of activation in both the 2-choice and 4-choice tasks. Significant activation in both tasks at reversal trials was seen bilaterally in the ventral striatum, midbrain, thalamus, anterior insula, dorsal anterior cingulate, orbitofrontal cortex, dorsolateral prefrontal cortex, posterior parietal cortex, motor cingulate, frontal eye fields, lateral extrastriate, visual cortex, and in the right superior temporal gyrus (Table 1; Fig. 2).

Activation during 2- and 4-choice reversal learning tasks

Comparison of 4- versus 2-choice reversal learning—The 4- versus 2-choice reversal learning tasks were compared to identify brain regions that were preferentially involved when reversing a learned response to an alternative response with uncertain versus certain outcomes (Table 2; Fig. 3). The increase in activity during unexpected non-reinforcement of a learned response relative to ongoing reinforcement of learned responses was greater in the 4-choice than 2-choice task in the bilateral ventral striatum, midbrain, thalamus, anterior parietal cortex, motor cingulate, frontal eye fields, and precuneus.

Supplementary analyses

We also conducted analyses aiming to identify brain areas in which response to ongoing reinforcement was greater than that for unexpected non-reinforcement. In both the 2- and 4-choice tasks, bilateral subgenual anterior cingulate cortex, posterior cingulate cortex and ventromedial prefrontal cortex were activated more during expected reinforcement than in non-reinforced trials. Additional regions were active in the 4-choice task, including the left orbitofrontal cortex, middle temporal gyrus, lateral extrastriate cortex, right insula, and bilateral hippocampus. Direct comparison of where unexpected non-reinforcement elicited greater activation relative to ongoing expected reinforcement in the 2-choice than 4-choice task showed significant activation in bilateral posterior parietal cortex, right frontal pole, right posterior middle frontal gyrus, right amygdala, and left middle temporal gyrus. These effects were due to modest activation in the 2-choice task and somewhat greater deactivation in the 4-choice task (see Appendix A for supplementary materials, Fig. A.1).

In addition to examining the difference between unexpected nonreinforcement and expected reinforcement for each task, we also examined brain activation in response to non-reinforcement whilst covarying for effects in trials in which positive reinforcement was provided. This was done in order to identify regions where activation related to non-reinforcement was significant in both tasks, but not differentially in the 2- and 4-choice tasks. Noteworthy observations from this analysis include: 1) activation in the dorsal anterior cingulate cortex was observed in both 2- and 4-choice tasks in response to nonreinforcement; 2) significant activation in the posterior cingulate cortex was seen bilaterally only in the 4-choice task at nonreinforcement; and 3) activation in the orbitofrontal cortex was seen only in the 2-choice task, and only in the right hemisphere. See supplementary materials for additional findings and details.

Behavioral data

A repeated-measures ANOVA of response latency data with response type (1st incorrect response, 2nd+ correct response) and task (2- and 4-choice tasks) showed a main effect of response type ($F(1,14)=11.60$, $p=.004$) such that 2nd+ correct responses were modestly faster ($M=.335$ s) than 1st incorrect responses ($M=.360$ s). That first incorrect responses tended to be slower than 2nd+ correct responses suggests that participants began to anticipate a rule reversal after getting several trials in a set consecutively correct, and consequently slowed their reactions to stimuli later in a set.

Response times of less than 100 ms occurred in 10% and 2% of trials respectively in the 2- and 4-choice paradigms. Imaging results were similar when these speeded responses were excluded from the above analyses. Omissions, failures to maintain set (changing response choice before non-reinforcement) and perseverative errors to previously rewarded stimuli were rare, occurring in less than 1% of trials in both tasks.

Discussion

The present fMRI study of reversal learning was designed to examine the neural substrates of reinforcement learning and choice behavior when outcomes of future response choices were certain and uncertain. At the reversal of response-outcome contingency in the 2- and 4-choice tasks, brain areas were engaged that are associated with both processing unexpected non-reinforcement and with changing response set. Activation was observed in regions involved in feedback processing, such as ventral striatum, midbrain and insula, as well as in areas involved in visual attention and response planning, including frontal eye fields, motor cingulate, dorsolateral prefrontal cortex, posterior parietal cortex, and dorsal anterior cingulate cortex. Activation in ventral striatum, midbrain, and in specific frontal and parietal regions was significantly enhanced at reversal in the 4-choice relative to the 2-choice reversal learning task.

Previous studies have established that prediction error signals scale in relation to intrinsic properties of rewards such as their magnitude and hedonic value (for a review, see Schultz, 2010). The present findings complement and extend those observations by showing that when response outcomes of future behavioral options are uncertain and consequently when there is greater behavioral relevance of nonreinforcement of learned responses, the response in brain regions known to manifest performance-related error signals is enhanced. This increased processing in reward circuitry may have facilitated the enhanced activation in higher-order cognitive systems that are crucial for reversal learning, including dorsal frontal regions that support the updating of response-outcome relationships and flexible behavior (Ghahremani et al., 2010; O'Doherty et al., 2003).

Feedback processing

The ability to appropriately respond to unexpected nonreinforcement, which cues participants to change a learned behavior to a newly adaptive response set, is crucial for reversal learning. Animal studies have demonstrated that when reinforcement is expected but not received, there is a depression of phasic dopamine firing in the nucleus accumbens, referred to as the negative reward prediction error signal (Schultz et al., 1993). This signal is thought to propagate throughout ventral striatum, and prefrontal and cingulate cortices, via projections from dopaminergic neurons in the ventral tegmental area (VTA; Schultz et al., 1997). This process is believed to serve an important role in facilitating adaptive behavior based on performance feedback. The relevance of this mesencephalic-ventrostriatal-frontal circuitry for reversal learning is demonstrated by observations that lesions of the ventral striatum in both primates and rodents result in perseverative responding, i.e., repeated selection of previously rewarded responses, despite these responses no longer being reinforced (Clarke et al., 2008; Ferry et al., 2000).

In the present study, during reversal trials in both the 2- and 4-choice tasks when expected reinforcement was not received, increased activation was observed in the midbrain, ventral striatum, and anterior insula. It is important to note the close temporal proximity of initial response planning/enacting and subsequent feedback in our paradigm, which does not allow for a meaningful separation of the haemodynamic responses related to these two task events. However, our findings of activation at reversal in this circuitry are consistent with those of previous studies demonstrating the relevance of these areas processing reward-related motivational signals (Cools et al., 2009; Gläscher et al., 2010; Preuschoff et al., 2008). The midbrain activation encompassed the VTA, which is a source of dopaminergic projections to striatum and neocortex, and is consistent with foci of activation seen in previous human neuroimaging studies of feedback-based learning, including reversal learning (Aron et al., 2004; Ghahremani et al., 2010). However, as the activation included much of the central midbrain, there also may be a contributing role of activity in other monoamine systems. For instance, noradrenergic output from the locus coeruleus has been implicated in the response to novelty and the engagement of cognitive control, both of which are involved in reversal learning (Minzenberg et al., 2008; Palencia and Ragozzino, 2006). The dorsal raphe nuclei send serotonergic projections to both the striatum and frontal cortex, which also is known to modulate the response to performance feedback and reversal learning (Boulougouris et al., 2008; Evers et al., 2005).

At reversal, greater activation in midbrain, ventral striatum and insula was seen in the 4-choice task compared to the 2-choice task. The neural response to unexpected non-reinforcement is known to scale with dimensions that are directly related to intrinsic properties of rewards themselves, such as the magnitude of a reward and its physical salience (for a review, see Schultz, 2010). Although reinforcement cues were identical in physical properties in the 2- and 4-choice tasks, the 4-choice task required the selection of a future response from among alternative response options whose outcomes were uncertain. Thus, our findings extend those of previous studies by suggesting a scaling of the feedback-related error signal with the uncertainty of outcomes of future response selection, and therefore with the behavioral relevance of negative feedback cues. The broader implication of this finding is that as the importance of non-reinforcement cues for planning future behavior increases, responses to negative feedback cues are enhanced in midbrain and ventrostriatal circuitry.

Behavioral flexibility

In addition to processing performance feedback, reversal learning requires the ability to inhibit prepotent learned response tendencies, and to select and engage in newly adaptive

behavior. Following the contingency change that signaled the need for behavioral reversal, we observed activation in regions known to be involved in flexible choice behavior, including the anterior cingulate, premotor, prefrontal and posterior parietal cortices.

Activation of dorsal anterior cingulate cortex in both 2- and 4-choice reversal learning tasks is consistent with findings from previous neuroimaging studies of reversal learning in humans (Budhani et al., 2007; Hampton and O'Doherty, 2007; Xue et al., 2008). The dorsal anterior cingulate cortex is believed to play an important role in integrating information related to reward and action selection (Hampton and O'Doherty, 2007; Hayden and Platt, 2010; Hillman and Bilkey, 2010), and in alerting subjects to unexpected events (Alexander and Brown, 2010; Yeung and Nieuwenhuis, 2009). Although activation of the anterior cingulate was observed in the 4- and 2-choice tasks, this region was not more active in the 4-choice task. Thus in this study, the uncertainty of future outcomes did not significantly affect the processing of non-reinforcement or cognitive demands for behavioral reversal in the dorsal anterior cingulate cortex. We also observed recruitment of the posterior cingulate cortex during reversal in the 4- but not the 2-choice task. Previous studies suggest a role for posterior cingulate as an integrative center for decision making processes, and in ordering response choices (Hayden et al., 2008; Woo and Lee, 2007). However, activity in this region was not significantly greater in the 4- than the 2-choice task.

As with the anterior cingulate cortex, activation in the dorsolateral prefrontal cortex was present in the 2- and 4-choice tasks at reversal, but was not more active in the 4-choice condition. These effects at reversal are consistent with many previous studies showing the importance of this region for inhibitory control and decision making (Kenner et al., 2010; Velanova et al., 2008; Xue et al., 2008). The cingulate region that was more active in the 4- relative to the 2-choice task was the motor cingulate cortex (Picard and Strick, 1996). This region supports the higher-order control of motor output, and has been shown to be engaged during reversal of conditioned associations (Paus et al., 1993). The frontal eye fields, another premotor area (Rosano et al., 2003), were also more active in the 4-choice condition at reversal. Like the motor cingulate, the frontal eye fields support motor planning functions, but in addition play an important role in visual attention and spatial cognition (Goldberg and Bruce, 1990). A similarly enhanced activation in the 4-choice condition was seen in the posterior parietal cortex and precuneus, which also support visual attention processes (Cavanna and Trimble, 2006; Merriam and Colby, 2005).

In the context of widespread interest in the interaction of cognitive and affective/motivational systems, the parallel increase in activation that we observed in dorsal cognitive systems and ventral reward processing systems is of considerable interest. Further research is needed to clarify whether this pattern of co-activation reflects an integrated response of these systems to task demands, or parallel but independent responses. However, there would be several potential advantages to an-interaction of these systems for reversal learning, such as the ability to dynamically alter attentional load based on the relevance of motivational cues, or to recruit higher-order executive strategies to evaluate changing response-outcome contingencies. Many studies have shown that the level of activation in dorsal premotor and parietal attention systems increases in relation to the behavioral relevance of external cues (Egner et al., 2008; Zenon et al., 2010). Motivational systems might enhance responses in dorsal premotor response planning and parietal attention systems in a bottom-up fashion, and thereby support choice behavior in uncertain circumstances. Such a process would require the scaling of motivational signals in relation to their relevance for future behavior, an effect we observed in the present study. Conversely, or perhaps in parallel, enhanced attention might facilitate processing of reward cues via top-down modulation.

Lastly, the present paradigm was designed to closely approximate 2-choice T-maze and 4-arm radial maze tasks that are widely used in studies of reversal learning in rodents. Thus, approaches like those taken in the present study may facilitate the development of translational methodologies between animal and human models, which are needed to enhance understanding of reversal learning from regional neurophysiology and neurochemistry to complex choice behavior, as well as to guide clinical research.

Conclusions

Reversal learning when future choices have certain and uncertain outcomes has not previously been compared in humans. Our results suggest that whilst there exists a core network of mesencephalic, frontostriatal and parietal attentional systems supporting reversal learning, uncertain consequences of response choices enhance the magnitude of activity in several brain regions during behavioral reversals. The increased activity in midbrain-ventrostriatal feedback processing systems may play a role in modulating activity upstream in attention and motor planning areas, which could benefit the selection and planning of future behavior. Via such a mechanism, enhanced performance-related error signals could provide bottom-up drive to facilitate attentional, inhibitory and motor planning processes when outcomes of future response choices are uncertain. Studies of reversal learning thus offer a promising platform for advancing the understanding of the interaction of affective/motivational and cognitive brain systems, and its relevance for supporting flexible adaptive behavior.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding sources: NIH Autism Center of Excellence 1P50HD055751; MH080066, and Autism Speaks

References

- Alexander WH, Brown JW. Competition between learned reward and error outcome predictions in anterior cingulate cortex. *Neuroimage*. 2010; 49(4):3210–3218. [PubMed: 19961940]
- Aron AR, Shohamy D, Clark J, Myers C, Gluck MA, Poldrack RA. Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *J. Neurophysiol.* 2004; 92(2): 1144–1152. [PubMed: 15014103]
- Baker TE, Holroyd CB. Which way do I go? Neural activation in response to feedback and spatial processing in a virtual T-maze. *Cereb. Cortex (New York N. Y.)*. 2009; 19(8):1708–1722.
- Balleine BW, Delgado MR, Hikosaka O. The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* 2007; 27(31):8161–8165. [PubMed: 17670959]
- Boulougouris V, Robbins TW. Pre-surgical training ameliorates orbitofrontal mediated impairments in spatial reversal learning. *Behav. Brain Res.* 2009; 197(2):469–475. [PubMed: 18983877]
- Boulougouris V, Glennon JC, Robbins TW. Dissociable effects of selective 5-HT_{2A} and 5-HT_{2C} receptor antagonists on serial spatial reversal learning in rats. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* 2008; 33(8):2007–2019.
- Budhani S, Marsh AA, Pine DS, Blair RJR. Neural correlates of response reversal: considering acquisition. *Neuroimage*. 2007; 34(4):1754–1765. [PubMed: 17188518]
- Cavanna AE, Trimble MR. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*. 2006; 129(3):564–583. [PubMed: 16399806]
- Clarke HF, Robbins TW, Roberts AC. Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *J. Neurosci.* 2008; 28(43):10972–10982. [PubMed: 18945905]

- Cools R, Lewis SJ, Clark L, Barker RA, Robbins TW. L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* 2007; 32(1):180–189.
- Cools R, Frank MJ, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *J. Neurosci.* 2009; 29(5):1538–1543. [PubMed: 19193900]
- Dias R, Robbins TW, Roberts AC. Primate analogue of the Wisconsin card sorting test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behav. Neurosci.* 1996; 110(5):872–886. [PubMed: 8918991]
- Dong H, Csernansky CA, Martin MV, Bertchume A, Vallera D, Csernansky JG. Acetylcholinesterase inhibitors ameliorate behavioral deficits in the Tg2576 mouse model of Alzheimer's disease. *Psychopharmacology.* 2005; 181(1):145–152. [PubMed: 15778881]
- Egner T, Monti JMP, Trittschuh EH, Wieneke CA, Hirsch J, Mesulam M. Neural integration of top-down spatial and feature-based information in visual search. *J. Neurosci.* 2008; 28(24):6141–6151. [PubMed: 18550756]
- Evers EA, Cools R, Clark L, van der Veen FM, Jolles J, Sahakian BJ, et al. Serotonergic modulation of prefrontal cortex during negative feedback in probabilistic reversal learning. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* 2005; 30(6):1138–1147.
- Ferry AT, Lu XC, Price JL. Effects of excitotoxic lesions in the ventral striatopallidal–thalamocortical pathway on odor reversal learning: inability to extinguish an incorrect response. *Brain Res. Exp. Hirnforsch. Experimentation Cerebrale.* 2000; 131(3):320–335.
- Floresco SB, Zhang Y, Enomoto T. Neural circuits subserving behavioral flexibility and their relevance to schizophrenia. *Behav. Brain Res.* 2009; 204(2):396–409. [PubMed: 19110006]
- Ghahremani DG, Monterosso J, Jentsch JD, Bilder RM, Poldrack RA. Neural components underlying behavioral flexibility in human reversal learning. *Cereb. Cortex.* 2010; 20(8):1843–1852. [PubMed: 19915091]
- Gläscher J, Hampton AN, O'Doherty JP. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb. Cortex (New York N. Y.).* 2009; 19(2):483–495.
- Gläscher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron.* 2010; 66(4):585–595. [PubMed: 20510862]
- Goldberg ME, Bruce CJ. Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal. *J. Neurophysiol.* 1990; 64(2):489–508. [PubMed: 2213128]
- Gregorios-Pippas L, Tobler PN, Schultz W. Short-term temporal discounting of reward value in human ventral striatum. *J. Neurophysiol.* 2009; 101(3):1507–1523. [PubMed: 19164109]
- Hampton AN, O'Doherty JP. Decoding the neural substrates of reward-related decision making with functional MRI. *Proc. Natl Acad. Sci. USA.* 2007; 104(4):1377–1382. [PubMed: 17227855]
- Hauber W, Sommer S. Prefrontostriatal circuitry regulates effort-related decision making. *Cereb. Cortex.* 2009; 19(10):2240–2247. [PubMed: 19131436]
- Hayden BY, Platt ML. Neurons in anterior cingulate cortex multiplex information about reward and action. *J. Neurosci.* 2010; 30(9):3339–3346. [PubMed: 20203193]
- Hayden BY, Nair AC, McCoy AN, Platt ML. Posterior cingulate cortex mediates outcome-contingent allocation of behavior. *Neuron.* 2008; 60(1):19–25. [PubMed: 18940585]
- Hillman KL, Bilkey DK. Neurons in the rat anterior cingulate cortex dynamically encode cost-benefit in a spatial decision-making task. *J. Neurosci.* 2010; 30(22):7705–7713. [PubMed: 20519545]
- Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage.* 2002; 17(2):825–841. [PubMed: 12377157]
- Kenner NM, Mumford JA, Hommer RE, Skup M, Leibenluft E, Poldrack RA. Inhibitory motor control in response stopping and response switching. *J. Neurosci.* 2010; 30(25):8512–8518. [PubMed: 20573898]
- Kim J, Ragozzino ME. The involvement of the orbitofrontal cortex in learning under changing task contingencies. *Neurobiol. Learn. Mem.* 2005; 83(2):125–133. [PubMed: 15721796]

- Lawrence AD, Sahakian BJ, Rogers RD, Hodges JR, Robbins TW. Discrimination, reversal, and shift learning in Huntington's disease: mechanisms of impaired response selection. *Neuropsychologia*. 1999; 37(12):1359–1374. [PubMed: 10606011]
- Matsuzaka Y, Tanji J. Changing directions of forthcoming arm movements: neuronal activity in the presupplementary and supplementary motor area of monkey cerebral cortex. *J. Neurophysiol*. 1996; 76(4):2327–2342. [PubMed: 8899607]
- McAlonan K, Brown VJ. Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behav. Brain Res*. 2003; 146(1–2):97–103. [PubMed: 14643463]
- Merriam EP, Colby CL. Active vision in parietal and extrastriate cortex. *Neuroscientist*. 2005; 11(5): 484–493. [PubMed: 16151048]
- Minzenberg MJ, Watrous AJ, Yoon JH, Ursu S, Carter CS. Modafinil shifts human locus coeruleus to low-tonic, high-phasic activity during functional MRI. *Science (New York, N.Y.)*. 2008; 322(5908):1700–1702.
- Mosconi MW, Kay M, D'Cruz AM, Seidenfeld A, Guter S, Stanford LD, et al. Impaired inhibitory control is associated with higher-order repetitive behaviors in autism spectrum disorders. *Psychol. Med*. 2009; (1–8)
- O'Doherty J, Critchley H, Deichmann R, Dolan RJ. Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J. Neurosci*. 2003; 23(21):7931–7939. [PubMed: 12944524]
- Palencia CA, Ragozzino ME. The effect of N-methyl-D-aspartate receptor blockade on acetylcholine efflux in the dorsomedial striatum during response reversal learning. *Neuroscience*. 2006; 143(3): 671–678. [PubMed: 17000053]
- Paus T, Petrides M, Evans AC, Meyer E. Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. *J. Neurophysiol*. 1993; 70(2):453–469. [PubMed: 8410148]
- Picard N, Strick PL. Motor areas of the medial wall: a review of their location and functional activation. *Cereb. Cortex*. 1996; 6(3):342–353. [PubMed: 8670662]
- Preusschoff K, Quartz SR, Bossaerts P. Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci*. 2008; 28(11):2745–2752. [PubMed: 18337404]
- Ragozzino ME, Mohler EG, Prior M, Palencia CA, Rozman S. Acetylcholine activity in selective striatal regions supports behavioral flexibility. *Neurobiol. Learn. Mem*. 2009; 91(1):13–22. [PubMed: 18845266]
- Rolls ET, McCabe C, Redoute J. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cereb. Cortex*. 2008; 18(3):652–663. [PubMed: 17586603]
- Rosano C, Sweeney JA, Melchitzky DS, Lewis DA. The human precentral sulcus: chemoarchitecture of a region corresponding to the frontal eye fields. *Brain Res*. 2003; 972(1–2):16–30. [PubMed: 12711074]
- Rosenberg DR, Averbach DH, O'Hearn KM, Seymour AB, Birmaher B, Sweeney JA. Oculomotor response inhibition abnormalities in pediatric obsessive compulsive disorder. *Arch. Gen. Psychiatry*. 1997; 54(9):831–838. [PubMed: 9294374]
- Schultz W. Dopamine signals for reward value and risk: basic and recent data. *Behav. Brain Funct*. 2010; 6(1):24. Retrieved from <http://www.behavioralandbrainfunctions.com/content/6/1/24>. [PubMed: 20416052]
- Schultz W, Apicella P, Ljungberg T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci*. 1993; 13(3):900–913. [PubMed: 8441015]
- Schultz W, Dayan P, Montague RR. A neural substrate of prediction and reward. *Science*. 1997; 275:1593–1599. [PubMed: 9054347]
- Smith SM. Fast robust automated brain extraction. *Hum. Brain Mapp*. 2002; 17(3):143–155. [PubMed: 12391568]
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004; 23 Supplement 1:S208–S219. [PubMed: 15501092]

- Velanova K, Wheeler ME, Luna B. Maturation changes in anterior cingulate and frontoparietal recruitment support the development of error processing and inhibitory control. *Cereb. Cortex.* 2008; 18(11):2505–2522. [PubMed: 18281300]
- Woo SH, Lee KM. Effect of the number of response alternatives on brain activity in response selection. *Hum. Brain Mapp.* 2007; 28(10):950–958. [PubMed: 17094117]
- Xue G, Ghahremani DG, Poldrack RA. Neural substrates for reversing stimulus–outcome and stimulus–response associations. *J. Neurosci.* 2008; 28(44):11196–11204. [PubMed: 18971462]
- Yeung N, Nieuwenhuis S. Dissociating response conflict and error likelihood in anterior cingulate cortex. *J. Neurosci.* 2009; 29(46):14506–14510. [PubMed: 19923284]
- Zenon A, Filali N, Duhamel JR, Olivier E. Salience representation in the parietal and frontal cortex. *J. Cogn. Neurosci.* 2010; 22(5):918–930. [PubMed: 19366288]

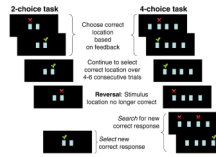


Figure 1.
Schematic presentation of 2- and 4-choice reversal learning tasks.

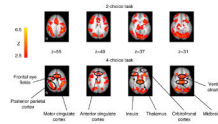


Figure 2. Activation observed in the 2-choice and 4-choice reversal learning tasks, for the contrast of unexpected non-reinforcement versus expected positive reinforcement of a learned response.

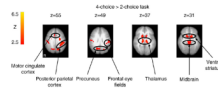


Figure 3. Regions for which significantly greater activation was observed in the 4-choice than in the 2-choice reversal learning task, for the contrast of unexpected non-reinforcement versus expected positive reinforcement of a learned response.

Table 1

Regions in which activation was greater in the 2-choice and 4-choice reversal learning tasks for unexpected non-reinforcement than for expected positive reinforcement of a learned response (all co-ordinates in standard MNI space).

Region	Hemi	2-choice Task			4-choice Task				
		Max t-value	Co-ordinates (MNI)			Max t-value	Co-ordinates (MNI)		
			x	y	z		x	y	z
Ventral striatum	R	3.35	14	22	-4	4.84	10	8	-8
	L	3.67	-12	14	-6	4.94	-12	6	-10
Midbrain	R	3.43	2	-20	18	5.27	6	-28	-12
	L	3.47	-4	-24	18	5.27	-6	-28	-18
Thalamus	R	4.12	12	-18	4	6.06	12	-14	4
	L	4.12	-12	-20	4	6.06	-12	-26	4
Anterior insula	R	4.31	34	22	-4	5.68	34	24	2
	L	4.28	-32	16	-6	6.06	-30	18	0
Anterior cingulate cortex, dorsal	R	4.49	0	14	42	5.65	0	14	42
	L	4.84	-4	12	42	5.78	-2	14	42
Orbitofrontal cortex	R	3.92	30	58	-16	2.81	24	52	0
	L	3.41	-26	50	-6	2.84	-22	50	0
Dorsolateral prefrontal cortex	R	4.52	40	40	26	4.46	26	12	48
	L	4.27	-38	34	16	4.88	-34	40	24
Posterior parietal cortex	R	4.92	44	-38	46	4.43	32	-46	36
	L	5.42	-42	-42	40	4.14	-40	-46	40
Motor cingulate	R	4.82	0	12	46	5.70	0	14	44
	L	5.23	-2	12	46	5.86	-2	14	44
Frontal eye fields	R	4.32	24	-8	50	5.31	22	-8	52
	L	4.43	-26	-12	56	5.96	-28	-4	46
Superior temporal gyrus	R	4.32	50	-42	12	3.70	56	-40	14
	L	--	--	--	--	--	--	--	--
Lateral extrastriate cortex	R	4.67	30	-68	-8	4.37	42	-64	12

Region	Hemi	2-choice Task				4-choice Task			
		Max t-value	Co-ordinates (MINI)			Max t-value	Co-ordinates (MINI)		
			x	y	z		x	y	z
	L	4.00	-40	-70	-12	3.53	-40	-80	12
Visual cortex	R	3.95	10	-82	-8	5.26	14	-86	-6
	L	4.01	-10	-82	-10	4.17	-16	-80	2

Table 2

Regions for which the increased activation for unexpected non-reinforcement relative to expected positive reinforcement of a learned response was greater in the 4-choice than the 2-choice reversal learning task (all co-ordinates in standard MNI space).

Region	Hemi	Max. t-value	MNI Co-ordinates of max. t-value		
			x	y	z
Ventral striatum	R	3.62	14	12	-8
	L	3.49	-14	10	-12
Midbrain	R	3.78	2	-32	-10
	L	3.87	-2	-32	-10
Thalamus	R	4.02	10	-18	4
	L	4.00	-10	-18	4
Posterior parietal cortex	R	3.86	48	-22	34
	L	3.88	-38	-34	34
Motor cingulate	R	3.74	6	10	46
	L	3.50	-2	16	44
Frontal eye fields	R	3.96	22	-4	50
	L	5.04	-24	-6	44
Precuneus	R	3.86	14	-60	24
	L	3.83	-16	-68	24