frontiers

Author's Proof

Please check your proof carefully and mark all corrections in the appropriate place.

Annotate your corrections on-screen using the Adobe Reader PDF editing tools, save and upload as Author's Proof Corrections.
Login → Journal Tab → Manage Articles → Author's Proof → Select Article → Press Enter Production Forum → Manuscript → Upload Files

Alternatively, the corrections can be listed by referring to the specific line number in the proof and be directly communicated in the Discussion Forum of your article.
Login → Journal Tab → Manage Articles → Author's Proof → Select Article → Press Enter Production Forum → Interactive Discussion → Enter Discussion Forum

Make sure to also answer all the queries thoroughly before submitting your comments, as failing to do so will cause delays.

Do not make any corrections by submitting a new manuscript file.

To ensure fast publication of your paper please return your corrections as soon as possible.

If you have any questions contact the Frontiers Production Office.

- Ensure to proofread the entire article, including figures and tables, captions, equations, citations, and references.
- Double-check the spelling of all author names, accuracy of affiliations and addresses.
- Verify that all the special characters are displayed correctly.
- Be sure that you have obtained permission for any reprinted material.
- Carefully reply to all of the author queries to avoid any production delays.

Author Queries Form

Query No.	Details required	Author's Reponse
Q1	Please cite "Feldman, 2004" inside the text.	

frontiers in
NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: xx March 2012
doi: 10.3389/fnins.2012.00035

The role of reward in dynamic decision making

Magda Osman*

Biological and Experimental Psychology Centre, School of Biological and Chemical Sciences, Queen Mary College, University of London, London, UK

Edited by:
Carlos Eduardo De Sousa, Northern Rio de Janeiro State University, Brazil

Reviewed by:
Jochen Ditterich, University of California, USA
Michael X. Cohen, University of Amsterdam, Netherlands

****Correspondence:***
Magda Osman, Biological and Experimental Psychology Centre, School of Biological and Chemical Sciences, Queen Mary College, University of London, Mile End Road, London E1 4NS, UK.
e-mail: m.osman@qmul.ac.uk

The present study investigates two aspects of decision making that have yet to be explored within a dynamic environment, (1) comparing the accuracy of cue-outcome knowledge under conditions in which knowledge acquisition is either through Prediction or Choice, and (2) examining the effects of reward on both Prediction and Choice. In the present study participants either learnt about the cue-outcome relations in the environment by choosing cue values in order to maintain an outcome to criterion (Choice-based decision making), or learnt to predict the outcome from seeing changes to the cue values (Prediction-based decision making). During training participants received outcome feedback and one of four types of reward manipulations: Positive Reward, Negative Reward, Both Positive + Negative Reward, No Reward. After training both groups of learners were tested on prediction and choice-based tasks. In the main, the findings revealed that cue-outcome knowledge was more accurate when knowledge acquisition was Choice-based rather than Prediction-based. During learning Negative Reward adversely affected Choice-based decision making while Positive Reward adversely affected predictive-based decision making. During the test phase only performance on tests of choice was adversely affected by having received Positive Reward or Negative Reward during training. This article proposes that the adverse effects of reward may reflect the additional demands placed on processing rewards which compete for cognitive resources required to perform the main goal of the task. This in turn implies that, rather than facilitate decision making, the presentation of rewards can interfere with Choice-based and Prediction-based decisions.

Keywords: dynamic, decision making, prediction, choice, reward

INTRODUCTION

The main objective of the present study is to build on the paradigms developed in the decision sciences in order to explore insights from work in the neurosciences on the role of reward. Based on the presentation of different types of reward outcomes, the present study examines the accuracy of cue-outcome knowledge when learning about a dynamic environment either through Choice-based decisions or Prediction-based decisions. A broader aim of this article is to elucidate the philosophical issues raised from work investigating decision making exclusively using behavioral techniques as compared to work using neuropsychological techniques.

Imagine a scenario in which we have recently installed a new energy monitoring system as a way of trying to reduce our fuel bill. In order to achieve this goal we need to learn about the relationship between cues (the devices in our home) and outcomes (energy use), while also taking into account our basic living requirements. We might decide that the best way to go about learning the cue-outcome relationships is by first choosing to make regular interventions on cues (varying which devices to use, varying the length of time of using the devices, and the time of use of various devices) and then examining their effects on the outcome (billing of fuel consumption). This is an example of Choice-based decision making in which cue-outcome relations are acquired via cue-intervention. Alternatively, by first monitoring the changes in cues (i.e., what devices are being used, and when) and then observing the changes in the outcome (energy use as indicated on the monitor) we might decide to predict the changes in the outcome from the changes in cue values. This is an example of Prediction-based decision making in which cue-outcome relations are acquired via estimates of the expected outcome value. Thus, both Choice-based decision making and Prediction-based decision making are methods of acquiring cue-outcome knowledge.

In order to achieve the intended goal, which is to ultimately to reduce our fuel bill, we would need to implement cue-outcome knowledge (acquired by either method – prediction/choice) in order to decide how we might change our future behavior to reduce our energy consumption. By implementing cue-outcome knowledge, over time we would be able to track the relative success of our decisions (positive reward, i.e., discovering that there was a decrease in the fuel bill) and the relative failure of our decisions (negative reward, i.e., discovering that there was an increase in the fuel bill). This form of updating, often referred to as reinforcement learning/reward learning is a way of associating rewards to the outcomes of decisions, which in turn influences how cue-outcome knowledge is implemented and modified.

What the above example illustrates is that, when we try to learn what variables that cause changes in a dynamic environment, we need to learn about cue-outcome relations, and we can do this through Choice-based decision making or Prediction-based decision making. Choice-based decision making involves refining the decisions that will help utilize the value functions associated with an outcome in order to reduce the discrepancy between a target (goal) and the outcome (Wörrgötter and Porr, 2005). Alternatively, we can learn what variables generate changes in a dynamic environment via Prediction-based decision making. This involves a process that refines the decisions that will determine the expected value function associated with an outcome (Wörrgötter and Porr, 2005). Either form of decision making will enable an incremental build-up of cue-outcome knowledge through a series of decision (prediction or choice). This means that future actions reflect the process of adapting and updating the cumulative changes experienced in the environment (Osman et al., 2008; Osman, 2008a,b, 2010a).

While neuropsychological research has made considerable advances in understanding the ways in which rewards are processed under different conditions (i.e., when the rewards occur and how often), very little work has focused on comparing the effects of different types of rewards on Prediction-based and Choice-based decision making, particularly in task environments that involve dynamic decision making (hereafter DDM) of the kind described in the example. Similarly, only recently has there been any work in the Judgment and decision making domain which directly compares the accuracy of cue-outcome knowledge gained via Prediction-based and Choice-based decision making in a dynamic environment (Osman and Speekenbrink, in press).

Osman and Speekenbrink (in press) showed that generally cue-outcome knowledge acquired either through Prediction-based or Choice-based decision making was sufficiently flexible to enable successful transfer to tests of choice and prediction. Moreover, these findings are generally consistent with reinforcement learning models that would claim that prediction errors are the source of cue-outcome learning, which can be generated either through Choice or Prediction. The key issue, and the focus of the present study, is to bring together the work from the decision sciences and the neuropsychological domain in order to investigate an unexplored question: What are the effects of different types of rewards on cue-outcome learning (i.e., Prediction-based, Choice-based decision making) in a DDM environment?

Broadly, both Prediction-based decisions and Choice-based decisions should lead to an estimate of what will happen to the outcome following a change in a cue variable, in other words a prediction is generated. Moreover, Reinforcement learning/Reward based learning models (Montague et al., 1996; Schultz et al., 1997) also claim that cue-outcome knowledge is acquired via error-based learning, that is, an error (prediction error) is generated by a comparison between an action (cue-intervention) and the actual outcome that occurs (reward; i.e., Choice-based decision). Alternatively an error can occur based on a comparison between an expected outcome from a choice and the actual outcome (i.e., Prediction-based decision). Thus, prediction errors are the source of learning – or fine tuning cue-outcome knowledge, and this is because the magnitude of the deviation between prediction/cue-intervention and the actual outcome indicates the accuracy of cue-outcome knowledge. The models predict that changes in the rate of learning reflect changes in the reward outcomes (i.e., success or failure of a decision reflected in the outcome itself).

Reinforcement learning models have enjoyed much success in the neuropsychological domain in which there is amassing evidence that the processing of rewards corresponds to phasic activity of mid-brain dopamine neurons (Schultz et al., 1997; Schultz, 2006; Rutledge et al., 2009). The pattern of activation of these neurons differs according to the different types of reward outcomes that occur. That is, dopaminergic neurons show short phasic activation in the presence of unexpected rewarding outcomes (e.g., presentation of food, presentation of money), and in the course of learning the phasic response shifts to indicators (i.e., cues) of rewarding outcomes (e.g., lights, tones, smiley faces, money). Similarly, in the presence of unexpected negative outcomes (e.g., loss of reward) there is a corresponding decrease in activation (Hollerman and Schultz, 1998). In addition, event-related brain potential (ERP) studies have reported that performance feedback generates ERP waveforms that are typically observed as a negative-going component peaking between 250 and 300 ms after feedback is presented (Holroyd and Coles, 2002; Hajcak et al., 2007; Peterson et al., 2011). The amplitude of the feedback negativity is determined by the impact of phasic dopamine signals (Holroyd and Coles, 2002). The amplitude of feedback negativity indicates the interaction between feedback valence and expectedness, so that unexpected negative feedback produces greater feedback negativity relative to unexpected positive feedback, which is typically associated with smaller negativity signals (Hajcak et al., 2007).

In addition, neuropsychological research on decision making has examined different properties of rewards (e.g., reward probabilities, reward structures; e.g., Daw et al., 2006; Behrens et al., 2007; Boorman et al., 2009; Jocham et al., 2009). Brain imaging data (O'Doherty, 2004; Sailer et al., 2007) has shown that there is greater brain activation in the orbital frontal cortex (OFC), caudate nucleus, and frontal polar areas when participants experience positive rewards (gains) rather than negative rewards (losses). This suggests that reward outcomes themselves are processed differently. Also, cortical activation can also reflect differences in reward probabilities, as well as changes in the reward probabilities over time (Cohen, 2006; Schultz, 2006; Sailer et al., 2007; Schultz et al., 2008). Moreover, during cue-outcome learning, activation increases in the OFC and putamen when experiencing losses, and activation decreases following gains; this is consistent with evidence from EEG studies (e.g., Cohen et al., 1996) and fMRI studies (e.g., Cohen et al., 2008).

Two recent neuropsychological studies contrasting Prediction-based learning (making judgments of expected rewards from actions, alternatively Prediction-based decision making) with action-based learning (choosing a cue that will bring about a reward, alternatively Choice-based decision making) suggest that there may in fact be underlying neurological differences between these two forms of learning (Hajcak et al., 2007; Peterson et al., 2011). The task in Hajcak et al.'s (2007) ERP study involved selecting from four doors the one which was likely to have a prize behind it (i.e., choice). Prior to each choice participants were told the objective probability of reward [i.e., the prize is behind 1 ($P = 0.25$), 2 ($P = 0.50$), or 3 ($P = 0.75$) doors]. The key manipulation involved participants guessing (i.e., predict) "yes" or "no" that they would win just before their choice (Experiment 1), or just after their choice (Experiment 2). Hajcak et al. (2007) found that consistent with reinforcement models, there was no difference between the two conditions based on behavioral measures of prediction and choice. There was however an effect on the correspondence between feedback negativity amplitude and subjective estimates of success. Feedback negativity tracked predictions of outcomes after people made their choices, but not before. It was speculated that the process of actively making a selection involved estimating the success of each choice, and then selecting the option with the highest subjective reward outcome. Thus, this evaluative method strengthened and stabilized predictions, whereas before a choice was made the prediction was based on few evaluations of the expected outcomes, and therefore weakened the strength of the predictions.

Using a different design, Peterson et al.'s (2011) study also separated prediction from action using an incremental learning task. Participants were either free to select a cue (one of four pictures) that yielded the highest expected pay off (choice trials), or were instructed to select a particular cue (instructed trials). Generally, the findings from the neurophysiological data suggested that prediction error magnitudes were lower for choice trials compared to instructed trials, but that only in choice trials did the error magnitude became substantially lower over the course of learning. Peterson et al. (2011) claimed that expectations are in closer alignment with feedback when feedback itself results from actions that are under volitional control, and this is based on the speculation that in Choice-based trials people can actively choose the option with the highest payoff where as for instructed trials people do not have volitional control.

The implication of Peterson et al. (2011) and Hajcak et al.'s (2007) findings is that active choice (i.e., Choice-based decision making) is an important factor in reward learning, and may involve different neural activity as compared to non-choice-based decisions (e.g., prediction, classical conditioning), but that there is no corresponding difference in behavioral measures of choice and prediction. The main reason for focusing on Hajcak et al. (2007) and Peterson et al. (2011) studies is that both make strong claims about reward learning in choice-based and prediction-based decision making. Moreover, in both studies the claim is made that reward differentially effects neurological behavior associated with prediction and choice, but that there is no corresponding behavioral differences (i.e., performance on tests of prediction and choice are no different). The problem is that without directly testing prediction and choice under the same task environment, unless one first establishes the presence or absence of behavioral differences, there are no secure ground for claiming that there are neurological differences but not behavioral differences. It is not clear why there would be differences at the neurological level and not at the behavioral level, which poses a number of questions concerning the kinds of inferences that can be drawn from neurological data to behavioral data, and vice versa.

What can we infer about the relationship between brain and behavior given that the changes detected at the neurophysiological level do not correspond with any observable changes in behavior at the psychological level? These findings raise important issues with respect to making inferences about the neurological mechanisms that support different forms of decision making. First, although in Hajcak et al.'s (2007) study predictions were made either before or after choices, both decisions were made on each trial. A cleaner design would have been to block trials in which people either predicted the success of a choice, or actually made a choice. In this way a comparison of prediction only and choice only trials would be free from potential order effects which were not examined in the study. Peterson et al. (2011) did in fact separate the trials in which choices and non-choices were made, but since participants were not explicitly required to make a subjective judgment about expected reward, the critical comparison was not between prediction and choice, but between choice and no-choice. Peterson et al. (2011) argued that their method of estimating prediction error magnitude from their reinforcement learning model was a more sensitive method than simply relying on verbal reports. Taken together, these methodological factors may explain the reported differences in neural activity and the absence of a difference at a behavioral level. However, both EGGs studies of choice and prediction are consistent with behavioral findings from Osman and Speekenbrink's (in press) study showing that the accuracy of cue-outcome knowledge is similar regardless of whether it was gained through prediction or choice. Though crucially in Osman and Speekenbrink's study there was no presentation of rewards during learning, only outcome feedback. Thus, the issue remains, to what extent can we extrapolate from neuropsychological findings to behavioral findings given that the differences are only present neurologically?

These issues will be revisited in the Section "General Discussion," but for now the key point is that evidence suggesting that choice and prediction may in fact be supported by different neurological processes has been demonstrated in simple forced choice tasks. The methodological concerns raised here may limit the extent to which the findings can be generalized to more complex decision making contexts. Therefore, given that behavioral studies comparing prediction and choice-based decision making do not include reward manipulations along the lines of Peterson et al. (2011) and Hajcak et al.'s (2007), and given that both these studies are problematic, the aim of the present study is to: (1) address the methodological issues raised here, (2) explore the generalizability of their findings to a DDM task by incorporating reward manipulations, and (3) explore the generalizability of their findings to a task which is commonly described as cognitively demanding (Brehmer, 1992).

Previous studies using DDM tasks directly comparing the effects of learning via prediction and learning via Choice-based decisions have shown that accuracy of cue-outcome knowledge is unaffected by mode of learning (Osman and Speekenbrink, in press). However, in the DDM tasks used previously, only outcome feedback was presented. This is different from the typical reward outcomes used in choice tasks in the neuropsychological domain. These tasks tend to incorporate salient reward outcomes (i.e., tones, lights, smiley faces) which have been shown to impact on performance. Therefore, the DDM task used in the present study incorporated reward outcomes during learning. Participants received outcome feedback, and were also presented with information as to the relative success of their decisions over time (indicated by a thumbs up sign and a smiley face – positive feedback), and the relative failure of decisions over time (indicated by a thumbs down sign and a sad face – negative feedback). In addition, the present study incorporated experimental procedures from Peterson et al. (2011) study and Hajcak et al.'s (2007) studies to make the DDM task comparable to their studies. In the prediction-based learning condition participants were presented with pre-selected cues (akin to Peterson et al., 2011 study) and were given the opportunity of guessing what the outcome value would be on each trial (akin to Hajcak et al., 2007 study).

By incorporating these methodological features into the present study, the aim is to align Peterson et al. (2011) and Hajcak et al.'s (2007) tasks to a paradigm examining decision making processes which is commonly referred to as cognitively demanding (Osman, 2010a), and is often described as externally valid (Funke, 2001). In so doing, the present study examines Hajcak et al.'s (2007) and Peterson et al.'s (2011) claim that Choice-based decisions rather than Prediction-based decisions facilitate closer correspondence between subjective expectations and feedback. They propose that, compared with Prediction-based decisions, Choice-based decisions reflect a process of volitional control over an action. The action itself is informed by an evaluative process in which each choice option is weighted and the one with the highest subjective reward is selected. This in turn would suggest an advantage for those making Choice-based decisions rather than Prediction-based decisions. However, this generates a discernable difference in neurophysiological behavior, but not in behavioral measures of performance. A null effect is also predicted from a reinforcement learning perspective. If experiencing the effects of one's predictions or choices cumulatively in a dynamic environment leads to the same prediction error, then regardless of the mode of learning, cue-outcome knowledge should be equally accurate in Prediction-based and Choice-based learning conditions.

EXPERIMENT 1

The experiment is designed to address the following empirical question: *Are there behavioral differences between Choice-based and Prediction-based dynamic decision making under reward based learning?* To answer this, the present study employed a DDM paradigm that incorporated a reward based structure similar to the simple choice tasks used in the neuropsychological domain discussed above. In one version of the DDM task, from trial to trial participants were required to learn the probabilistic cue-outcome associations by using the cue values to predict the outcome value (Prediction-based learners). The other version involved the same cue-outcome task structure, but in this case participants were required to control the outcome value by manipulating the cue values to reach and maintain a specific outcome value (Choice-based learners). To match the two versions as closely as possible, the learning histories experienced by both types of learners were identical, but the critical difference between the two was that Choice-based learners set the cue values (choice under volition), whereas the cue values were preset for Prediction-based learners (non-volitional cue manipulation). This was achieved by using a yoked design. In this way, Prediction-based learners were matched to Choice-based learners' learning trials, and so the cue-outcome values that were experienced were identical to those chosen by Choice-based learners. To examine the effects of the different modes of learning on the accuracy of cue-outcome knowledge, all participants were presented with two tests of control, and two tests of prediction.

METHODS

Participants

Ninety-six graduate and undergraduate students from University of London volunteered to participate in the experiment for reimbursement of £5. The assignment of participants to the four conditions was semi-randomized. There were a total of eight groups (Choice-based learning Positive Reward, Choice-based learning Negative Reward, Choice-based learning Both Positive + Negative Reward, Choice-based learning No Reward, and Prediction-based learning Positive Reward, Prediction-based learning Negative Reward, Prediction-based learning Both Positive + Negative Reward, Prediction-based learning No Reward), with 12 participants in each. Pairs of participants (Choice-based learners and yoked Prediction-based learners) were randomly allocated to one of the four types of reward based conditions (Positive Reward, Negative Reward, Both Positive + Negative Reward, No Reward). Participants were tested individually.

DESIGN

The experiment used a 2×4 design. It included two between subject manipulations, namely learning mode (Prediction-based vs. Choice-based) and type of reward (Positive Reward, Negative Reward, Both Positive + Negative Reward, No Reward). Success of learning performance was measured using two types of tests (Control Test 1, 2; Predictive Tests 1, 2).

The task environment consisted of the following: Positive cue $= x_1$, Effect of positive cue $= b_1 = 0.65$, Negative cue $= x_2$, Effect of negative cue $= b_2 = -0.65$. Random perturbation $= e_t$, (the random perturbation component, is normally distributed, with a mean of 0), Outcome value $= y(t)$, Previous outcome value $= y(t-1)$. Thus, there were three cues and one outcome. One of the cues increased the outcome, and one of the cues decreased the outcome. The third cue had no effect on the outcome. More formally, the task environment can be described as in the following equation

$$y(t) = y(t-1) + 0.65\ x_1(t) - 0.65\ x_2(t) + e(t)$$

in which $y(t)$ is the outcome on trial t, x_1 is the positive cue, x_2 is the negative cue, and e a random noise component, normally distributed with a zero mean and SD of 8[1]. The null cue x_3 is not included in the equation as it had no effect on the outcome.

The DDM task included a total of 112 trials, divided into two phases. The structure of the entire experiment was as follows: Learning phase (40 trials), Test Phase – Two tests of Controlling the Outcome (20 trials each) interleaved with Two test of Predicting Cue and Outcome values (16 trials each). The order of presentation of the tests was as follows, Control Test 1, Prediction Test 1, Control Test 2, Prediction Test 2.

[1]The assignment of noise to the system was first piloted in order to generate High variance (16 SD) and low variance (4 SD). Osman and Speekenbrink (in press) includes two studies which varied the random perturbation component, In Experiment 1, 16 SD was found to be difficult as reflected in choice performance and predictive performance, while 4 SD was considerably easier. In Experiment 2, 8 SD was moderately difficult, and on this basis was chosen in order investigate the effects of reward on Choice-based and Prediction-based learning in the present study.

BEHAVIORAL TASK

The visual layout of the screen, cover story, and the main instructions were identical for Prediction-based and Choice-based learning groups. Participants were presented with a story about a newly developed incubator designed especially for babies with an irregular state of health (a global measure based on heart rate, temperature, blood pressure)[2]. Using this type of context ensured that participants were highly motivated to learn the task. Choice-based learners were informed that as a trainee maternity nurse they would be trying to regulate the health of a newborn girl called "Molly." They would be regulating the levels of three parameters (air pressure, oxygen, and humidity) with the aim of maintaining a specific safe healthy state. The system was operated by varying the cue values which would affect the baby's state of health. Prediction-based learners were assigned the same role, but instead they were told that they would see the nurse regulating the incubator parameters and that their role would be to predict the subsequent change in a global measure of health. The screen included three cues which were labeled (air pressure, oxygen, and humidity), and the outcome (healthy state) which was presented in two ways, as a value in the middle right of the screen, and also on a small progress screen in which a short trial history (five trials long) of outcome values was presented. Both Prediction-based and Choice-based learning groups were shown the current state of health, new value of the state of health after manipulation and the target value of the healthy state. Prediction-based learners were also shown the result they predicted in the form of a dashed line on the progress screen. The task was self-paced. **Figure 1** shows an example of the environment participants were required to interact with.

Rewards

Rewards based stimuli were presented during the learning phase only. The rewards did not correspond to money or points, but rather they were simple characters that indicated an increase (smiley face and a thumbs up sign) or decrease (sad face and a thumbs down sign) in performance. Participants in the No Reward (No Reward) condition received no reward, only outcome feedback.

[2]It was made clear to participants at the start of this experiment, that they were taking part in a simulation, and that there was no real baby in an incubator.

Outcome feedback was provided in the form of a value that changed on a progress screen indicating graphically the difference between the target value and the achieved outcome value (for Choice-based learners), or the predicted outcome value and the achieved value (for the Prediction-based learners). In addition the outcome value and target value were also listed on the side of the progress screen.

Participants in the positive reward condition (Positive Reward) observed a picture of a smiley face and a thumbs up on trials in which the discrepancy between their achieved outcome value and the target value was smaller than the previous trial (for Choice-based learners), or the discrepancy between expected and actual outcome was smaller than the previous trial (for the Prediction-based learners). Participants in the negative reward condition (Negative Reward) observed a picture of sad face and a thumbs down on trials in which discrepancy between the achieved outcome and target outcome was greater than the previous trial (for Choice-based learners), again a similar logic was applied to Prediction-based decisions (for the Prediction-based learners). Participants in Positive + Negative reward condition (Both-Rewards) received positive and negative rewards on trials adhering to the conditions specified above. Rewards were only presented during the learning phase. During the Test phase, for control tasks all participants received outcome feedback, and for tests of prediction no feedback was presented.

Learning phase

Choice-based learners. During each trial participants had to interact with the system by changing the value of the cues using a slider corresponding to each. Each slider had a scale that ranged from 0 to 100. On the start trial, the cue values were set to "0," the outcome value was 178, the target value throughout was 62, and a safe range (±10 of the target value) was given. When participants made their decision they clicked a button labeled "Submit" which deactivated the cues and revealed on the progress screen the effects of their decisions on the outcome. The effects on the outcome value were cumulative from one trial to the next, and so while the cue values were returned to "0" on the next trial, the outcome value was retained from the previous trial. After completing the learning phase, participants then proceeded to the test phase.

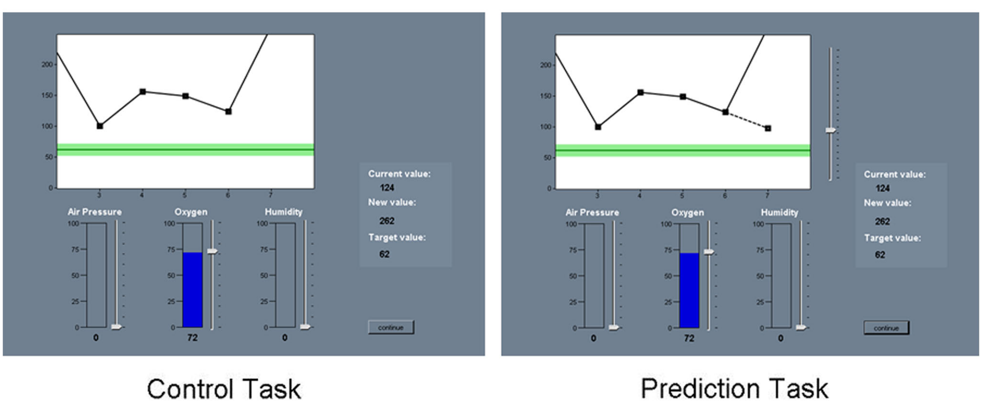

FIGURE 1 | Screen shots of a control-learning trial and a predict-learning trial.

Prediction-based learners. The procedure was identical to Choice-based learners, with the following exceptions. Once presented with the cue values, they predicted the outcome value by adjusting a slider that was placed alongside the outcome progress screen; this would move a line on the progress screen to indicate the outcome value. Once they made their decision, they clicked a button labeled "Submit," which deactivated the outcome value slider and revealed the actual outcome value as well as their predicted outcome value. The button "Continue" was then pressed to proceed to the next trial. The start of the next trial triggered the outcome value slider to become activated and the presentation of new cue values. The predicted value of the previous trial was omitted from the progress screen, but the trial history of the last five actual outcome values remained.

Test phase

Control tests. After the learning phase, all participants were examined on their ability to control the system to a criterion (outcome value = 62, and safe range ±10 of the target value). Test 1 involved the same procedure that the Choice-based learners were following during the learning phase, but consisted of only 20 trials. For the Prediction-based learners this was the first occasion they could manipulate the cues. To examine the ability to control the system to a different goal, all participants were then presented with Test 2 in which they followed the same procedure as Test 1, with the following exceptions. In the Test 2 participants were informed that they needed to be even more careful in reaching and maintaining the outcome value (outcome value = 74), and that staying within the safe range (±5 of the target value) was of particular importance. The starting value of Test 1 was 178, and was set to 156 in Test 2. In the Test 2 Choice-based learners and Prediction-based learners had no experience of the new criterion value, and so they would have to base their decisions on acquired knowledge of the system in order to control the new outcome value.

Predictive tests were designed to examine explicit cue-outcome knowledge. Each test included 16 trials which were divided in the following way. Participants were required to predict the value of a cue (Positive, Negative, Null) based on the given value of the outcome and the other cues (e.g., predicting the Positive cue value, based on the values of the Negative, Null, and Outcome values), or they were required to predict the outcome value given the value of the other three cues. Participants were not told that the test involved a mixture of eight old trials and eight new trials. Old trials were divided accordingly: 2 × Positive cue value, 2 × Negative cue value, 2 × Null cue value, 2 × Outcome value). These trials were randomly selected from the initial learning phase (for Choice-based learners these were trials that they had generated themselves, for Prediction-based learners these were the same yoked learning trials in which they predicted the outcome value). The 8 new trials were divided accordingly: 2 × Positive cue value, 2 × Negative cue value, 2 × Null cue value, 2 × Outcome value. Neither group had prior experience of them. All participants were presented with the same set of new trials; these were predetermined prior to the experiment. The presentation of the 16 trials in each set of Predicting Cue and Outcome values Tests was randomized. For each trial the predictive value was recorded along with the response time.

Dependent measures

Predictive performance was measured by an error score $S_{\mathrm{p}}(t)$ calculated as the absolute difference between predicted and expected outcome values:

$$S_{\mathrm{p}}(t) = \left| P(t) - y(t-1) - 0.65\ x_1(t) + 0.65\ x_2(t) \right|,$$

in which $P(t)$ is a participant's prediction on trial t. We chose to compare predictions to expected rather than actual outcomes as the latter are subject to random noise.

Choice performance was measured as the absolute difference between the expected achieved and best possible outcome:

$$S_{\mathrm{c}}(t) = \left| G(t) - y(t-1) - 0.65\ x_1(t) + 0.65\ x_2(t) \right|,$$

in which $G(t)$ is the goal on trial t: either the target outcome if achievable on that trial, or the closest achievable outcome. To illustrate, choice performance was based on how much participants' cue manipulations deviated from the optimal cue settings (the same principle applies to predictive performance except the deviation was from expected outcome values on each trial). In the choice tasks used here, for a given (previous) outcome value and goal, the optimal cue settings define a line in a two-dimensional plane. For example, if the deviation between the previous outcome and goal is 50, then the optimal cue settings are all values for the positive cue x_1 and negative cue x_2 such that $50 = 0.65\ x_1 - 0.65\ x_2$, for instance a value of $x_1 = 77$ and $x_2 = 0$, or $x_1 = 78$ and $x_2 = 1$, $x_2 = 87$ and $x_2 = 10$, etc. Thus, choice performance was computed as the (shortest) distance between a participant's actual settings for these two cues and the line defining the optimal cue settings.

RESULTS

The participants' patterns of learning were first examined separately for Choice-based learners and Prediction-based learners. Comparisons between conditions could not be conducted at this stage as the optimality scores were incomparable (one based on the difference between achieved and best possible outcome value, and the other between predicted and expected outcome value). The Test Phase was the first occasion in which both conditions were directly compared for the participants' ability to reach and maintain the outcome to a specific criterion (Tests of Controlling the Outcome), and their ability to predict cue values from the state of the outcome, or predict the outcome from the pattern of cue values (Test of Predicting Cue and Outcome values).

Learning phase: choice-based learning

The learning phase was divided into two blocks of 20 trials each (Learning first half; Learning second half), and Control optimality scores were averaged across each block, for each participant. The following analyses were based on the mean error scores by block, presented in **Figure 2**. To examine the success of learning, 2 × 4 repeated measures ANOVA was conducted using Block (Learning first half; Learning second half) and Reward (No Reward, Both-Rewards, Positive Reward, Negative Reward). Overall, with more exposure to the task, Choice-based learners showed general improvements in their ability to control

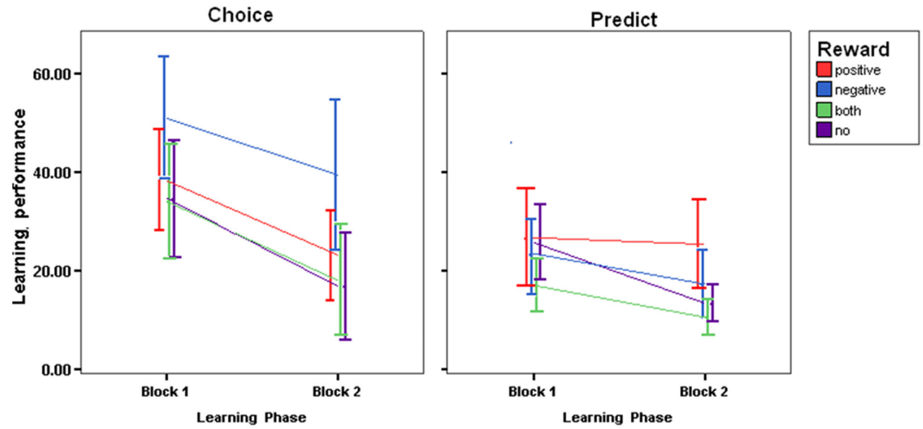

FIGURE 2 | Choice-based error scores and prediction-based error scores during the learning phase for all four reward groups (SE±).

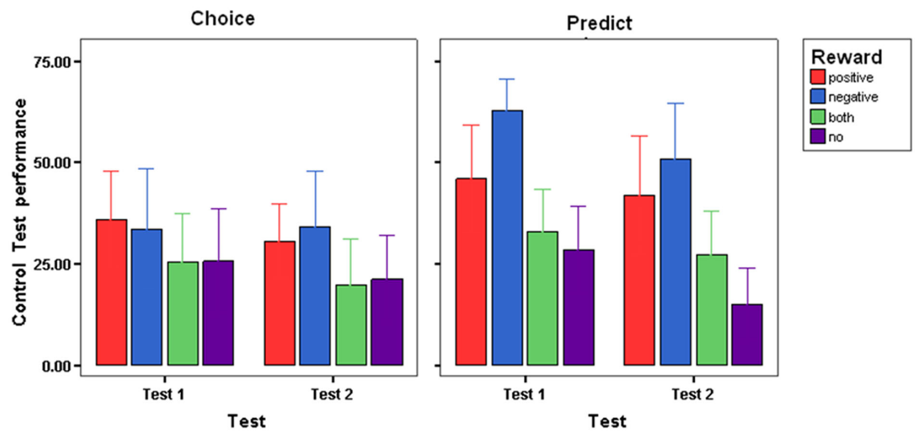

FIGURE 3 | Choice-error scores during the test phase control test 1, control test 2, for each reward group and condition (SE±).

the outcome to criterion as revealed by a main effect of Block [$F_{(1,44)} = 44.019$; $P < 0.0005$, $\eta = 0.527$]. There was a significant main effect of Reward [$F_{(2,44)} = 3.443$; $P < 0.05$, $\eta = 0.202$]. A Bonferroni *post hoc* tests revealed that Negative Reward led to poorer control performance as compared to those receiving Both-Rewards (19.147, $P < 0.05$) and compared to those receiving No Reward (19.389, $P < 0.05$).

Learning phase: prediction-based learning

In order to examine predictive accuracy during learning Predictive optimality scores were subjected to 2×4 repeated measures ANOVA with Block (Learning first half; Learning second half) and Reward (No Reward, Both-Rewards, Positive Reward, Negative Reward). The analysis revealed a main effect of Block [$F_{(1,44)} = 26.278$; $P < 0.001$, $\eta = 0.374$], confirming the pattern of behavior presented in **Figure 2** indicating that predictive accuracy improved with more practice. There was also a Block × Reward interaction [$F_{(3,44)} = 3.064$; $P < 0.05$, $\eta = 0.173$]. Bonferroni *post hoc* test failed to reach significance. There was also a significant main effect of Reward [$F_{(3,44)} = 3.010$; $P < 0.05$, $\eta = 0.170$]. Bonferroni *post hoc* tests revealed that receiving Positive Reward led to poorer predictive accuracy as compared to Both-Rewards (12.237, $P < 0.03$).

Test phase: control

Control optimality scores were averaged across participants in each group for each of the two Tests of Controlling the Outcome and are presented in **Figure 3**. An ANOVA using Condition (Choice-based learners, Prediction-based learners) and Reward (No Reward, Both-Rewards, Positive Reward, Negative Reward) × Test (Control Test 1, Control Test 2) was conducted. Generally all participants improved in their control performance in Test 2 as compared to Test 1, suggesting the presence of practice effects, as revealed in a main effect of Test, [$F_{(1,88)} = 14.020$; $P < 0.0001$, $\eta = 0.137$]. A main effect of Condition suggested that Choice-based learners were more accurate in their control performance compared to Prediction-based learners [$F_{(1,88)} = 8.293$; $P < 0.005$, $\eta = 0.086$][3]. There was also a main effect of Reward [$F_{(3,88)} = 9.506$; $P < 0.0005$, $\eta = 0.245$]. To examine this further,

[3]Bonferroni correction was applied.

control optimality scores were collapsed across Test and Condition and Bonferroni tests were carried out on Feedback. The tests revealed those receiving No Reward during learning showed more accurate control performance as compared with Positive Reward (16.007, $P < 0.01$), and Negative Reward (22.756, $P < 0.001$). Also, receiving Negative Reward led to poorer control performance as compared to receiving Both-Rewards (18.87, $P < 0.001$). No other comparisons were significant. It appears that in tests of control, those receiving no reward during training tended to show the most accurate control performance.

Test phase: prediction

Tests of Predicting Cue values and Outcome values provided the opportunity to examine the extent to which the cue-outcome knowledge gained by Choice-based learners was sufficiently flexible to equivalent levels of accuracy as Prediction-based learners. Prediction optimality scores for Test 1 and Test 2 are presented in **Figure 4**. The scores were collapsed across the Tests, since an ANOVA with Test (Predictive Test 1, Predictive Test 2) × Condition (Choice-based learners, Prediction-based learners) and Reward (No Reward, Both-Rewards, Positive Reward, Negative Reward) failed to show any differences in patterns of predictive accuracy between tests. Cue (Positive, Negative, Outcome) × Familiarity (Old trials, New trials) × Condition (Choice-based learners, Prediction-based learners) × Reward (No Reward, Both-Rewards, Positive Reward, Negative Reward) were used as factors in an ANOVA. A main effect of Familiarity $[F_{(1,176)} = 21.464;\ P < 0.0005,\ \eta = 0.196]$ was significant. In general all participants were more accurate in their predictions for trials they had experienced previously during learning as compared to unfamiliar trials. There was a Familiarity × Cue interaction $[F_{(2,176)} = 3.902;\ P < 0.05,\ \eta = 0.042]$. Paired t-tests revealed that compared with new trials, there was greater predictive accuracy for old trials when predicting the value of the positive cue $[t(95) = 3.708,\ P < 0.0004]$ and the negative cue $[t(95) = 5.433,\ P < 0.00004]$. There was no difference in predictive accuracy between old and new trials when predicting the outcome. No other effects or interactions were significant.

GENERAL DISCUSSION

The main objective of this study was to investigate the following question: *Are there behavioral differences between Choice-based and Prediction-based dynamic decision making under reward based learning?* In general, the evidence from the present study corroborates the pattern of neuropsychological evidence from ERP studies (Hajcak et al., 2007; Peterson et al., 2011), but not the behavioral evidence from these studies. The present study shows that active involvement generates more accurate cue-outcome knowledge than non-volitional learning of cue-outcome relations. Though reward based learning led to differences in performance between Choice-based and Prediction-based learning, the effects of reward were unexpected. Compared to participants that were not presented with reward, on the whole the presentation of reward tended to impair learning and transfer of cue-outcome knowledge. Therefore, the findings demonstrate behavioral differences between Prediction-based and Choice-based decision making in a DDM task were the result of the presentation of reward.

More specifically, the findings from this study show that during learning Negative Reward severely impaired Choice-based performance, while Positive Reward severely degraded predictive accuracy. Moreover, Positive Reward and Negative Reward generally impaired performance in Learning and Test when compared with participants receiving No Reward or Both-Rewards. In addition, Choice-based learners showed an overall advantage in later tests of control. This suggests that volitional control over cue manipulations during learning facilitated later ability to control an outcome to different criteria. Moreover, Choice-based learning also facilitated successful transfer of cue-outcome knowledge to Predictive tests. The present discussion focuses on two main issues: (1) the detrimental effects of reward on decision making, and (2) the broad philosophical issues that are raised by neuropsychological research on choice and prediction.

WHY DID REWARD BASED FEEDBACK IMPAIR DDM?

Kluger and DeNisi's (1996, 1998) review of the effects of feedback on skill based learning (low level motor and perceptual learning as well as high level problem solving and decision making) suggest that unless the task is simple, feedback will lead to no

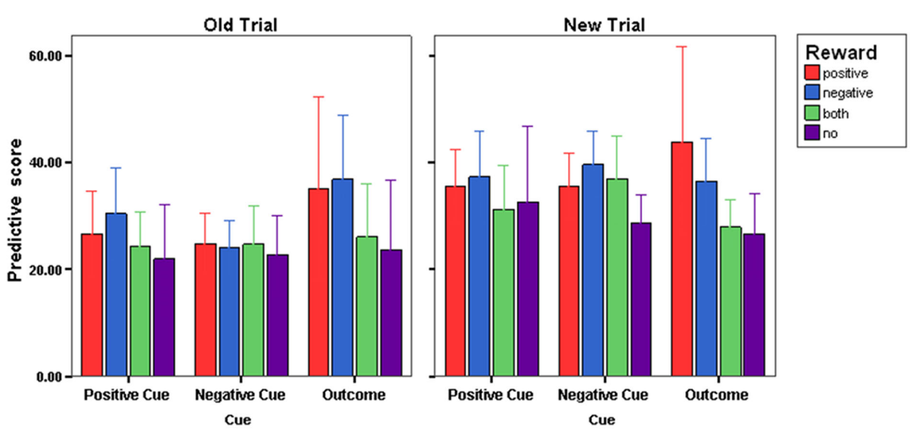

FIGURE 4 | Prediction error scores (SE±) during the test phase collapsed across prediction test 1 and prediction test 2, for each reward group.

additional benefits in most cases, and in extreme cases impair learning (e.g., Hammond and Summers, 1972; Salmoni et al., 1984). They claimed that the effectiveness of feedback depends on the type of goal that that the learner is pursuing. More recently, Harvey (2011) has proposed a cognitive resources account as a way of explaining the differential effects on performance through feedback as a function of task difficulty. He proposes that tasks, such as DDM, are examples in which the knowledge needed to achieve success is not easily identified from the outset, and so the process of information search makes high demands on executive functions. As a result, the provision of feedback (e.g., cognitive feedback, reward outcomes) is problematic in these tasks for the reason that it is a source of additional information that needs to be processed in order to be usefully incorporated into the performance of the main task. The more demanding the task is, the more likely it is that feedback will interfere because processing feedback competes with performing the main task.

In fact, many have argued that DDM tasks are examples of complex problem solving tasks (Funke, 2010; Osman, 2010a), and have been used as methods of indexing IQ (Joslyn and Hunt, 1998; Gonzalez, 2005; Funke, 2010). Therefore, there are good grounds for assuming that the kind of decision making processing that goes on in DDM tasks is cognitively expensive. This is because decision making involves tracking cue-outcome relations in a dynamic environment. At any one time a decision maker is still uncertain as to the generative causes of changes in an observed outcome in a DDM task. The reason being that the observed changes to the outcome can result from endogenous influences (i.e., cue manipulations in the DDM task) or exogenous influences on those outcomes (i.e., functions of the system itself/noise), or a combination of both endogenous and exogenous influences.

It may be the case that feedback (cognitive feedback, reward outcomes) may impair decision making processes such as those involved in DDM tasks because additional processing resources are needed to evaluate feedback in order to use it to adapt and update decision making behavior (Harvey, 2011). For simple forced choice tasks (e.g., Hajcak et al., 2007; Peterson et al., 2011), the learner possess the relevant knowledge for making a decision from the outset, and learning simply reflects the efficiency in implementing that knowledge. Therefore, providing feedback in forced choice tasks does not compete with processing demands made from performing the main task. By extension, when contrasting the simple forced choice task used by Hajcak et al. (2007) and Peterson et al. (2011) and the DDM task in the present study, reward based learning may have adversely affected performance because DDM task is more cognitively demanding than forced choice tasks.

To explore this, separate analyses were conducted comparing the optimality scores of the Choice-based learning No Reward condition and the Prediction-based learning No Reward condition in the Control tests, and the findings revealed that there were no difference in performance between conditions [$F_{(1,22)} = 0.07$; $P = 0.785$, $\eta = 0.003$; see text footnote 3]. Furthermore, this result replicates the findings from Osman and Speekenbrink's (in press) study (Experiment 2). When the same analysis was conducted collapsing across the three remaining reward based conditions, more accurate performance was found for Choice-based learners receiving feedback as compared to Prediction-based learners receiving feedback, [$F_{(1,70)} = 9.47$; $P < 0.005$, $\eta = 0.119$]. Though caution should be exercised in drawing any firm conclusions from this result, it certainly is supportive of the proposal that in the case of DDM tasks, reward infers with DDM, more specifically, active based decision making in which cue-interventions are made. Moreover, the inference may result from the fact that DDM tasks are cognitively demanding and so processing rewards competes for the same limited resources available to perform the main task. This may also explain why the presentation of rewards does not appear to impair performance in forced choice tasks.

Clearly this has implications for reinforcement learning models (Schultz et al., 1997; Schultz, 2006), at two levels, given that fundamentally, Choice-based and Prediction-based decisions should lead to equivalent cue-outcome knowledge, why is it that a difference in performance at test was found? Second, reinforcement learning models would predict differential effects on performance based on different types of reward, but why is it that rewards differentially affected performance of Prediction-based and Choice-based conditions during the learning? In response to these issues, it might be worth considering the informational content of the outcome feedback for Choice-based and Prediction-based learners. On each trial during learning, outcome feedback could be used to indicate the deviation of the expected outcome value from the achieved outcome value (comparison 1 – prediction error) and the deviation of the achieved outcome value from the target value (comparison 2). This was the case in the present study and in Osman and Speekenbrink (in press). Osman and Speekenbrink's (in press) findings suggest that both Prediction-based and Choice-based learners were using comparison 1 and comparison 2 interchangeably during learning, because this enabled both Prediction-based and Choice-based learners to perform control and prediction tasks equally well at test. In the present study, the introduction of reward may have prevented Choice-based and Prediction-based learners from attended to both comparison 1 and 2. Instead the presence of reward made salient comparison 1 for Prediction-based learners, and made salient comparison 2 for Choice-based learners. This may have resulted in the advantage found in Choice-based learners in later tests of control. The equivalent cue-outcome knowledge found in Prediction-based and Choice-based learners in tests of prediction suggest that either comparison 1 or 2 generates sufficient cue-outcome knowledge to perform the test.

This would be consistent with the speculation that volitional control over setting the cue values during learning encouraged Choice-based learners to evaluate each cue-outcome relationship, whereas the evaluation process was not as exhaustive during Prediction-based learning (Hajcak et al., 2007; Peterson et al., 2011). The differential effects of reward on Prediction-based decisions and Choice-based decisions may reflect a difference in the magnitude of the effects of gains and losses for different types of decisions (Schultz et al., 1997; Sailer et al., 2007). However, this is still speculative and given that to date, no previous study has examined the effects of feedback on Choice-based and Prediction-based decisions in a DDM task, further work is needed to explore the possible influences of reward on decision making.

PHILOSOPHICAL ISSUES RAISED BY NEUROPSYCHOLOGICAL RESEARCH ON CHOICE-BASED AND PREDICTION-BASED DECISION MAKING

A question asked at the start of this article based on the implication of Peterson et al. (2011) and Hajcak et al.'s (2007) findings was: *What can we infer about the relationship between brain and behavior given that the changes detected at the neurophysiological level do not correspond with any observable changes in behavior at the psychological level?* The same question will now be tackled with respect to philosophical issues concerning the inferences that this and present study can make about the neurological mechanisms that support different forms of decision making.

The virtue of neuroscience is that it allows us to gain access to processes that were once inaccessible to psychologists. The rational usually follows along the lines of: If brain region X is active, then cognitive process Y will be active. For this rational to work, there also has to be an assumption that the causal arrow goes in the direction of brain to behavior. Detractors of this position can make the argument that there is a lack of functional specificity of regions in the brain which undermines any strong inferences that can be made from neuroimaging data to behavioral measures (Poldrack, 2006). As a case in point, while Peterson et al. (2011) and Hajcak et al.'s (2007) are not neuroimaging studies, nevertheless, their critical findings concern differences neurophysiologically but not behaviorally. So what can be inferred from such findings? Given that the logical of many neuropsychology studies involves detecting a change in the pattern of activation in certain brain regions and then inferring cognitive processes from observable changes in behavioral measures, it is perhaps even more problematic to make inferences about the association between brain regions and cognitive processes when the differences lie only in neurophysiological data.

Also, if, like many psychologists and neuroscientists, materialism (in which ever flavor is adopted) is the favored position, because if behavior is reducible to regions in the brain, then one is interested in discovering the etiology of human behavior by examining the processes in the brain. The rational here follows along the lines of: If my study manipulates cognitive process Y, then given what I know from work conducted in the neurosciences, brain region X should be activated. So long as neurophysiological and behavioral data converge, there are no problems in developing an explanatory account of a cognitive process based on the patterns of data at both level. The problem that is posed here is deciding what the appropriate level of explanation for prediction-based and choice-based decision making given that behavioral data imply one type of account, and neurophysiological data suggest an alternative account. As a case in point, the findings from Peterson et al. (2011) and Hajcak et al.'s (2007) studies pose this problem. The experimental manipulations in both studies were designed to pit two cognitive processes (i.e., choice and prediction) against each other. While the behavioral data from both studies implies a single mechanism that supports Choice-based and Prediction-based decisions through the generation of prediction errors, the neurophysiological data suggests there might be different underlying mechanisms that correspond to the cognitive processes.

Where as the issues discussed above concern problems in interpreting neurophysiological and behavioral data, a more general issue is that there may well be limitations in extrapolating from simple tasks to more complex task in designed to simulate real world situations (Osman, 2010b). The issue comes down to scalability. The argument concerning the practice of transforming higher-level cognitive behaviors observed in the real world to detectable lower-level neurobiological phenomena takes many forms (Bickle, 2006, 2007; Craver, 2007; Sullivan, 2009); though for simplicity this discussion will focus on two: Internal and External validity. *External validity* refers to the correspondence between results implying a causal relationship between variables in a laboratory to variables of the same kind existing outside of it (Guala, 2003). Elegant simple choice tasks used in neuropsychological research may not be sufficient tools for studying complex behaviors if they cannot adequately explain or predict complex behavior in the real world. *Internal validity* refers to the success of an experimental result that establishes a causal relationship between variables found to operate in the context of a laboratory. If there is not a general convergence of reductive practices in neuropsychological experiments in establishing causal relationships between high level behaviors and cellular/molecular processes, then mental functions are ultimately not reducible to cellular/molecular processes.

To a large extent, pragmatic factors (i.e., the investigative aims of the researcher) determine which type of validity is prioritized when developing an experiment (Sullivan, 2009). But, pragmatism does not necessarily lead to any unity in the way in which phenomena (e.g., Prediction-based vs. Choice-based decision making) are examined in a cognitive psychology laboratory or an EEG laboratory. However, philosophers such as Craver (2007) would argue that the same mechanism (decision making) is being examined in at different levels in neuroscientific and cognitive science circles. There is a: (a) specialized level in the nature (e.g., neural activity) of the components of the mechanism are being examined (intralevel) – and (b) a more expansive level in which the interventions are made in order to examine the function of the components of the mechanism (interlevel). Unity is achieved when researchers refer to and try and integrate findings from both intralevel and interlevel experiments. By the same token, the behavioral differences found presently between Prediction-based and Choice-based decision making, and the differences in neural activity between the two reported in Hajcak et al.'s (2007) and Peterson et al.'s (2011), could be viewed as examples of findings from studies at intralevel and interlevel. However, the convergence of general findings at the different levels still creates a problem, because there are more still differences in the methodologies between the present study and the aforementioned EEG studies, and so this still compromises the possibility of drawing broad conclusions that the differences between prediction and choice essentially is based on volitional control.

CONCLUSION

The resent study used a DDM task to investigate the accuracy of cue-outcome knowledge when learning in dynamic environment was Prediction-based or Choice-based. In addition, the influence of reward on both was examined. To this end, the evidence suggests that Choice-based decision making leads to more accurate cue-outcome knowledge than Prediction-based learning.

However, the inclusion of reward adversely effected decision making during learning and at test. The type of DDM task included in the present study is cognitively more demanding than the typical choice tasks used in neuropsychological studies examining reward learning. The present article argues that the processing of rewards places an additional burden on cognitive resources that are already stretched when performing DDM tasks. The competition for resources leads to general decrements in decision making performance as compared to when no rewards are present. Though the general findings from this study are compatible with recent evidence from the neuropsychological domain, large differences in methodology prevent any strong conclusions being drawn with respect to supporting the claim that differences between prediction and choice are based on the level of volitional control. A number of philosophical arguments are considered with respect to generalizing evidence from neuropsychology to psychology and vice versa, in particular the inferential fallacies that are made, and the pragmatic constrains on the way studies are conducted.

ACKNOWLEDGMENTS

The support of the ESRC Research Centre for Economic Learning and Human Evolution is gratefully acknowledged. Preparation for this research project was supported by the Economic and Social Research Council, and the Engineering and Physical Sciences Research Council, EPSRC grant – EP/F069421/1 (Magda Osman). I would also like to thank Maarten Speekenbrink for preparing the experimental program and for conducting the data analysis.

REFERENCES

Behrens, T. E., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.

Bickle, J. (2006). Reducing mind to molecular pathways: explicating the reductionism implicit in current cellular and molecular neuroscience. *Synthese* 151, 411–434.

Bickle, J. (2007). Ruthless reductionism and social cognition. *J. Physiol. Paris* 101, 230–235.

Boorman, E. D., Beherns, T. E., Woolrich, M. W., and Rushworth, M. F. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62, 733–743.

Brehmer, B. (1992). Dynamic decision making: human control of complex systems. *Acta Psychol. (Amst.)* 81, 211–241.

Cohen, M. (2006). Individual differences and the neural representations of reward expectancy and reward error prediction. *Soc. Cogn. Affect. Neurosci.* 2, 20–30.

Cohen, M., Elger, C., and Weber, B. (2008). Amygdala tractography predicts functional connectivity and learning during feedback-guided decision making. *Neuroimage* 39, 1396–1407.

Cohen, M. S., Freeman, J. T., and Wolf, S. (1996). Meta-cognition in time stressed decision making: recognizing, critiquing and correcting. *Hum. Factors* 38, 206–219.

Craver, C. (2007). *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience.* Oxford: Oxford University Press.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

Feldman, J. (2004). How surprising is a simple pattern? Quantifying "Eureka!". *Cognition* 93, 199–224.

Funke, J. (2001). Dynamic systems as tools for analyzing human judgment. *Think. Reason.* 7, 69–89.

Funke, J. (2010). Complex problem solving: a case for complex cognition? *Cogn. Process.* 11, 133–142.

Gonzalez, C. (2005). Decision support for real-time, dynamic decision-making tasks. *Organ. Behav. Hum. Decis. Process* 96, 142–154.

Guala, F. (2003). Experimental localism and external validity. *Philos. Sci. Suppl.* 70, 1195–1205.

Hajcak, G., Moser, J., Holyroyd, C., and Simons, R. (2007). It's worse than you thought: the feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology* 44, 905–912.

Hammond, K., and Summers, D. (1972). Cognitive control. *Psychol. Rev.* 79, 58–67.

Harvey, N. (2011). "Learning judgment and decision making from feedback: an exploration- exploitation tradeoff?" in *Judgment and Decision Making as a Skill: Learning, Development, and Evolution*, eds M. K. Dhami, A. Schlottmann, and M. Waldmann (Cambridge: Cambridge University Press).

Hollerman, J. R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Holroyd, C. B., and Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709.

Jocham, G., Neumann, J., Klein, T., Danielmeier, C., and Ullsperger, M. (2009). Adaptive coding of action values in the human rostral cingulate zone. *J. Neurosci.* 29, 7496–7489.

Joslyn, S., and Hunt, E. (1998). Evaluating individual differences in response to time-pressure situations. *J. Exp. Psychol. Appl.* 4, 16–43.

Kluger, A. N., and DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol. Bull.* 119, 254–284.

Kluger, A. N., and DeNisi, A. (1998). Feedback interventions: toward the understanding of a double-edged sword. *Curr. Dir. Psychol. Sci.* 7, 67–72.

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.

O'Doherty, J. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776.

Osman, M. (2008a). Observation can be as effective as action in problem solving. *Cogn. Sci.* 32, 162–183.

Osman, M. (2008b). Evidence for positive transfer and negative transfer/anti-learning of problem solving skills. *J. Exp. Psychol. Gen.* 137, 97–115.

Osman, M. (2010a). Controlling uncertainty: a review of human behavior in complex dynamic environments. *Psychol. Bull.* 136, 65–86.

Osman, M. (2010b). *Controlling Uncertainty: Learning and Decision Making in Complex Worlds.* Oxford: Wiley-Blackwell Publishers.

Osman, M., and Speekenbrink, M. (in press). Prediction and control in dynamic decision making environments. *Front. Neuropsychol.*

Osman, M., Wilkinson, L., Beigi, M., Parvez, C., and Jahanshahi, M. (2008). The striatum and learning to control a complex system? *Neuropsychologia* I46, 2355–2363.

Peterson, D., Lotz, D., Halgren, E., Sejnowski, T., and Poizner, H. (2011). Choice modulates the neural dynamics of prediction error processing during rewarded learning. *Neuroimage* 54, 1385–1394.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci. (Regul. Ed.)* 10, 59–63.

Rutledge, B., Lazzaro, S., Lau, B., Myers, C., Gluck, M., and Glimcher, P. (2009). Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *J. Neurosci.* 29, 15104–15114.

Sailer, U., Robinson, S., Fischmeister, F., Moser, E., Kryspin-Exner, I., and Bauer, H. (2007). Imaging the changing role of feedback during learning in decision making. *Neuroimage* 37, 1474–1486.

Salmoni, A. W., Schmidt, R. A., and Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. *Psychol. Bull.* 95, 355–386.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* 57, 87–115.

Schultz, W., Dayan, P., and Montague, R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Schultz, W., Preuschoff, K., Camerer, C., Hsu, M., Fiorillo, C., Tobler, P., and Bossaerts, P. (2008). Explicit neural signals reflect reward uncertainty. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 3801–3811.

Sullivan, J. (2009). The multiplicity of experimental protocols: a challenge to reductionist and non-reductionist models of the unity of neuroscience. *Synthese* 167, 511–539.

Wörgötter, F., and Porr, B. (2005). Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. *Neural Comput.* 17, 245–319.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 October 2011; paper pending published: 25 November 2011; accepted: 26 February 2012; published online: xx March 2012.

Citation: Osman M (2012) The role of reward in dynamic decision making. Front. Neurosci. ***6****:35. doi: 10.3389/fnins.2012.00035*

This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.