Cognitive, Affective, & Behavioral Neuroscience (2019) 19:1458–1466
https://doi.org/10.3758/s13415-019-00730-2

The importance of agency in human reward processing

Cameron D. Hassall[1] · Greg Hajcak[2] · Olave E. Krigolson[1]

Published online: 11 June 2019
© The Psychonomic Society, Inc. 2019

Abstract
Converging evidence suggests that reinforcement learning (RL) signals exist within the human brain and that they play a role in the modification of behaviour. According to RL theory, prediction errors are used to update values associated with actions and/or predictive cues, thus facilitate decision-making. For example, the reward positivity—a feedback-sensitive component of the event-related brain potential (ERP)—is thought to index an RL prediction error. An unresolved question, however, is whether or not action is required to elicit a reward positivity. Reinforcement learning theory would predict that the reward positivity should diminish or disappear in the absence of action, but evidence for this claim is conflicting. To investigate the impact of cue, choice, and action on the amplitude of the reward positivity, we altered a two-armed bandit task by systematically removing these factors. The reward positivity was greatly reduced or absent in the altered versions of the task. This result highlights the key role of agency in producing learning signals, such as the reward positivity.

Keywords Agency · Reward positivity · Reinforcement learning

Importance of agency in human reward processing

Reinforcement learning (RL) describes how we learn to do the right thing at the right time. More formally, RL is a computational theory that describes how an agent, or decision-maker, learns to maximize its rewards by interacting with the environment (Sutton & Barto, 1998). Rewarded actions are more likely to be repeated, and punished actions are less likely to be repeated (Thorndike, 1911/2017). In particular, actions are selected via a *policy*, linking scenarios (states) to action likelihoods. Learning occurs when the policy is updated following feedback. Neuroimaging evidence suggests that RL algorithms are implemented in the human brain (O'Doherty, Cockburn, & Pauli, 2017).

One possible neural RL signal is the reward positivity, a feedback-sensitive component of the human event-related potential (ERP). The reward positivity, also called the feedback-related negativity (FRN; see Proudfit, 2015), is a positive ERP deflection that is sensitive to RL prediction errors (Holroyd & Coles, 2002; Krigolson, 2017; Sambrook & Goslin, 2015; Walsh & Anderson, 2012). Whenever feedback occurs an RL prediction error is computed reflecting the difference in value between the expected and the actual outcome. Thus, an unexpected outcome elicits a larger reward positivity than an expected outcome, and a large-magnitude outcome elicits a larger reward positivity than a small-magnitude outcome (see Walsh and Anderson, 2012, for a list of contradictory evidence). RL prediction errors are then used to update the policy, increasing or decreasing the value of selecting certain actions in a given state.[1]

The degree to which the reward positivity reflects an RL prediction error has been studied and debated (Walsh & Anderson, 2012). One aspect of this debate relates to the role of action. Usually, an agent's *own action* initiates a learning event (i.e., cue → choice → action → outcome → policy update). In other words, the reward positivity and learning should take place when individuals have *agency* (control over actions and their outcomes: Haggard, 2017). However, humans and other animals are able to learn by observing the consequences of others' actions (observational learning: Bellebaum, Kobza, Thiele, & Daum, 2010), in the absence

✉ Olave E. Krigolson
krigolson@uvic.ca

[1] Centre for Biomedical Research, University of Victoria, P.O. Box 1700 STN CSC, Victoria, British Columbia V8W 2Y2, Canada

[2] Department of Psychology and Biomedical Sciences, Florida State University, Tallahassee, FL, USA

[1] This describes *model-free*, as opposed to *model-based* learning (Daw, Niv, & Dayan, 2005). Although we focused here on the former, electroencephalography (EEG) has the potential to reveal the roles of both systems (Collins & Frank, 2018).

of action altogether (e.g., classical conditioning: Pavlov, 1927/2010; Rescorla & Wagner, 1972), and from counterfactual outcomes (Fischer & Ullsperger, 2013). Strictly speaking, observational learning and classical conditioning lie outside the scope of RL, because they lack self-initiated actions; however, these examples illustrate the diverse range of conditions under which learning takes place (and the possibility of multiple learning systems in the brain). Even within the realm of RL, self-initiated actions may differ in their sense of agency. For example, selecting one of two possible snacks involves more agency (or outcome control) than flipping a coin.

Early experiments on the importance of agency on the reward positivity have been inconclusive. For instance, Martin and Potts (2011) found that removing agency from a response (i.e., having the computer respond instead of the participant) obliterated the reward positivity. However, in prior work by Yeung, Holroyd, and Cohen (2005) participants showed reduced (but still present) reward positivities when their actions were perceived to have no impact on outcomes (also see Mühlberger, Angus, Jonas, Harmon-Jones, & Harmon-Jones, 2017). Furthermore, and in contrast to Martin and Potts (2011), Yeung et al. (2005) observed a small but significant reward positivity in the absence of action (i.e., when the computer initiated the trial; see Donkers, Nieuwenhuis, & van Boxtel, 2005, for a similar result).

These previous findings can be summarized as follows: the reward positivity is reduced in the absence of agency and is further reduced (or absent) in the absence of action. If the reward positivity reflects an RL signal, why does it still occur in the absence of agency and in the absence of action, at least some of time? Yeung et al. (2005) offered the following explanation: the reward positivity is not only tied to learning about actions but also to learning about reward contingencies in the world (classical conditioning). Thus, expectations may be possible even in the absence of agency and action; rewards preceded only by predictive cues may still elicit a reward positivity (Donkers, Nieuwenhuis, & van Boxtel, 2005). Interestingly, predictive cues themselves may elicit a reward positivity, further suggesting that the reward positivity is not solely related to learning about choices and actions (Dunning & Hajcak, 2007; Holroyd, Krigolson, & Lee, 2011; Krigolson, Hassall, & Handy, 2013; Krigolson & Holroyd, 2007). Indeed, if the reward positivity reflects a more general prediction error signal (as opposed to an action-contingent RL signal), then it should still be present in the absence of choice and action, although perhaps diminished for other reasons. It remains to be seen whether the reward positivity would still be present in the absence of choice, action, *and* predictive cues.

Consider two casino games: roulette and slots (slot machines). Both are games of chance involving action, but roulette also has choices: players choose the bet amounts and predicted outcomes (e.g., "red" or "even"). In contrast, slot machines offer actions (insert coin, pull arm) but no choices, traditionally. Now imagine watching helplessly while someone else plays roulette with your money. You observe bets and outcomes, but the choices and actions are not your own. This scenario describes several previously used experimental tasks designed to examine the neural response to feedback in the absence of action (Donkers et al., 2005; Martin & Potts, 2011; Yeung et al., 2005). To date, however, the effects of choice and action on the reward positivity have yet to be compared within the same individuals. Additionally, the role of predictive cues in the generation of this neural signal is still somewhat unclear. The scenario above, in which bets and outcomes can only be observed, can be further modified by hiding the actual bets. Would a normal reward positivity be generated in the absence of these predictive cues (i.e., with outcomes only)?

In the present study, we sought to 1) reproduce earlier work showing a reduction in the reward positivity in the absence of choice and action, and 2) show further reduction or abolishment of the reward positivity in the absence of predictive cues. To address these hypotheses, we asked participants to play four versions of a standard decision-making task (the doors task: Proudfit, 2015). Across the tasks we manipulated agency within four experimental conditions as follows: 1) cue → choice → action → outcome, 2) cue → action → outcome, 3) cue → outcome, and 4) outcome. In line with previous work, we predicted that the reward positivity would be present in the first condition (cue → choice → action → outcome) and would be attenuated or abolished in the other three conditions.

Method

Participants

We tested 26 undergraduate students at the University of Victoria. All participants had normal or corrected-to-normal vision and no known neurological conditions. Participants were recruited via the psychology department's online recruitment system and were compensated with credit in an undergraduate psychology class and $8.40. EEG data for two participants were excluded from the results (noisy ocular channels in one case, ground electrode failure in the other). Of the remaining 24 participants, 13 were male and 2 were left-handed ($M_{age} = 21.54$, $SD_{age} = 2.72$). The study was approved by the University of Victoria Human Research Ethics Board, and all participants gave written informed consent.

Apparatus and Procedure

Participants were seated 60 cm in front of a 22-inch LCD display (75 Hz, 2-ms response rate, 1,680 x 1x050 pixels, LG W2242TQ-GF, Seoul, South Korea). Visual stimuli were presented using the Psychophysics Toolbox Extension (Brainard, 1997; Pelli, 1997) for MATLAB (Version 8.2,

Mathworks, Natick, USA). Participants were given written and verbal instructions to minimize head and eye movements throughout the experiment.

Participants completed four versions of the doors task, a computer-based guessing game (Proudfit, 2015). The order of the games was counterbalanced across participants (24 orderings in total). There were 60 trials per game, for a total of 240 trials across all games. Participants were instructed, in writing, that they would be playing four games for money, and that they would be paid their total at the end of the experiment. They were further informed that each win, indicated by the appearance of an upward green arrow, would increase their total by $0.14, and that each loss, indicated by a downward red arrow, would decrease their total by $0.07. Unknown to participants, outcomes were such that 50% of trials resulted in a win and 50% of trials resulted in a loss, in randomized order. Participants were told that detailed instructions would be provided prior to each game (see below). Finally, participants were shown the contents of a cash box (several $5 bills, $2 coins, and $1 coins) to reassure them that the money was real.

Choice condition In the choice version of the doors task—the standard version—two identical doors were presented in the center of the display. The doors were separated by 1.1° of visual angle, and each door subtended approximately 2.8° by 5.5°. The doors remained on the display until one was selected (mouse cursor moved over a door, and left mouse button clicked). Following the mouse click, and prior to visual feedback, a fixation cross (0.5° by 0.5°) appeared for 500 ms. Visual feedback—a 0.9° by 2.2° green or red arrow—then appeared for 2,000 ms. Another fixation cross appeared for 1,500 ms, following by the words "Next outcome" for 500 ms. Participants were given the following written instructions: "'In this game you will see two doors. Select one of the doors using the mouse. One door leads to a win (green arrow) and the other to a loss (red arrow). Place your hand on the mouse and press any key to begin." In line with previous uses of the doors task, the timing of stimulus presentation was not jittered (Bress, Foti, Kotov, Klein, & Hajcak, 2013; Mulligan, Flynn, & Hajcak, 2018). See Figure 1 for a sample trial with timing details.

No-choice condition In the no-choice version of the doors task, participants, upon the appearance of the doors, initiated the trial by pressing the spacebar on a keyboard. After the button press, the mouse cursor on the screen moved to near the center of one of the doors, indicating the computer's choice. Door choice was random, and the cursor movement time varied between 300 ms and 500 ms. All other timing and stimuli were matched to the choice task. Participants were given the following written instructions: "In this game you will see two doors. After you press the spacebar, the computer will select one of the doors using the mouse. One door leads to a win (green arrow) and the other to a loss (red arrow). Place your hand on the spacebar and press any key to begin."

No-response condition This version was identical to the no-choice task, except that the participant was not required to initiate the trial with a button press. Rather, the computer automatically made a selection 500-700 ms after the appearance of the doors. Participants were given the following written instructions: "In this game you will see two doors. Do not press any buttons – the computer will automatically select one of the doors using the mouse. One door leads to a win (green arrow) and the other to a loss (red arrow). Press any key to begin, then remove your hands from the keyboard."

No-cue condition Here, no doors were presented, and each trial began with the appearance of a fixation cross for 500 ms. The remainder of a trial was identical to the other versions of the task. Participants were given the following written instructions: "In this game you will simply receive wins and losses. Some trials will result in a win (green arrow), and some trials will result in a loss (red arrow). Press any key to begin, then remove your hands from the keyboard."

Data Collection

Sixty-three channels of EEG data, referenced to channel AFz, were recorded using Brain Vision Recorder (Version 1.21.0004, Brain Products GmbH, Munich, Germany). Sixty-one electrodes were placed in a fitted cap according to the 10-20 system. Additionally, two electrodes were affixed to the mastoids (left and right). Conductive gel was applied to ensure that electrode impedances were below 20 kΩ before recording, and the EEG data were sampled at 500 Hz and amplified (actiCHamp, Brain Products GmbH, Munich, Germany).

Data Analysis

EEG preprocessing was done in BrainVision Analyzer (Version 2.1.2, Brain Products GmbH, Munich, Germany). EEG data were downsampled to 250 Hz and re-referenced to the average of the mastoid channels. The original reference (AFz) was recovered, and the mastoid channels were removed from the data set, leaving 62 channels in total. The data were then filtered using a phase shift-free Butterworth filter (0.1–30 Hz pass band, 60-Hz notch). Ocular artifacts were corrected by submitting all pre-feedback (fixation cross) and feedback EEG data to independent component analysis (ICA). Specifically, components associated with eye blinks were removed from the continuous EEG (Jung et al., 2000). The ICA algorithm was trained on EEG data around feedback events (−1 to 2 s) but applied to the continuous data. The continuous

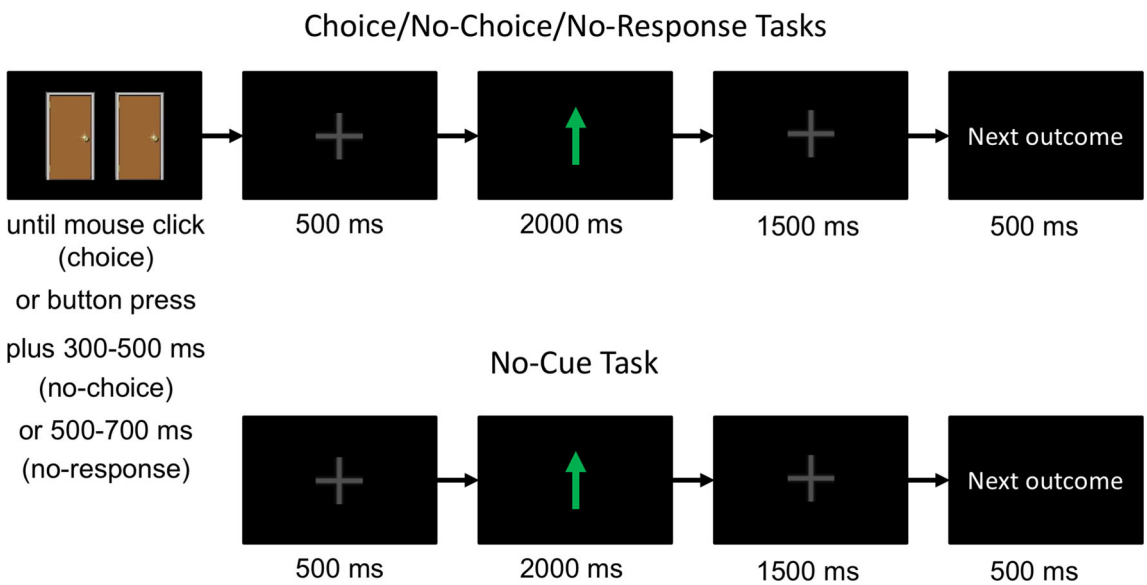

Fig. 1 Task stimuli and timing details. The choice/no-choice/no-response tasks differed in how a trial was initiated. The door stimuli were absent in the no-cue task

ICA-corrected data were then segmented into 800 ms epochs: 200 ms before 600 ms following the onset of feedback stimuli.

The remainder of the analysis was done in MATLAB (Version 9.4, Mathworks, Natick, USA) using a combination of custom scripts and EEGLAB (Delorme & Makeig, 2004). Epochs in which the voltage changed more than 10 μV per sampling point or more than 150 μV across the entire epoch were excluded from the analysis. On average, 8% of epochs were excluded ($SD = 5\%$). ERPs were created for each participant by averaging the feedback-locked EEG data at each channel, task (choice, no-choice, no-response, no-cue), and feedback valence (win, loss). Grand average conditional waveforms (mean of all participants' win and loss ERPs) for each task were computed for each channel. The reward positivity was then analyzed using the difference wave method. For each task, a difference wave was computed for each participant by subtracting the average loss waveform from the average win waveform. A grand difference wave (mean of all participants' difference waves) was also computed for each task and channel. Based on previous work (Holroyd & Coles, 2002; Miltner, Braun, & Coles, 1997), and an examination of the grand average conditional waveforms and difference waves (Figures 2 and 3), we defined the reward positivity for each participant and task as the mean voltage from 252 ms to 288 ms post-feedback at electrode FCz.

In addition to measuring the reward positivity, we also explored the possibility that the P300 component of the ERP was impacted by task. This was done in order to replicate results from Yeung et al. (2005), who reported larger P300s for choice outcomes compared to no-choice outcomes. This is interesting because the P300 has been linked to motivation, a possible factor in our experiment (Kleih, Nijboer, Halder, & Kübler, 2010). Here, we chose to examine the conditional waveforms (win, loss), rather than the difference waveforms, as described above. This was done for two reasons. First, this was an exploratory analysis; we had no a priori P300 hypothesis about the win-minus-loss difference waves. Second, we recognized that although the reward positivity is best analyzed using the difference-wave approach (Proudfit, 2015; Krigolson, 2017), for the P300 an analysis of the conditional waveforms may be more appropriate (Polich, 2007). We defined the P300 as the mean voltage 300-412 ms post feedback at electrode Pz (time range and location of maximal response, for all conditions; see Polich, 2007). Thus, a P300 score was computed for each task (choice, no-choice, no-response, no-cue) and outcome (win, loss).

The existence of the reward positivity within each task was determined using single-sample t tests (Krigolson 2017; Krigolson & Holroyd, 2007). Additionally, we computed Cohen's d for each "existence test" as follows:

$$d = \frac{M_{\text{diff}}}{s_{\text{diff}}}$$

where M_{diff} and s_{diff} are the mean and standard deviation of the reward positivity scores (see Cumming, 2014). A one-way repeated measures ANOVA was conducted to determine the effect of task (choice, no-choice, no-response, no-cue) on the reward positivity. The P300 was subjected to a 4 (task: choice, no-choice, no-response, no-cue) x 2 (outcome: win, loss) repeated-measures ANOVA. Two different effect-size measures (partial eta squared and generalized eta squared) were computed for each ANOVA (Lakens, 2013, Olejnik & Algina, 2003). All error bars on figures and error measures for mean reward positivity scores reflect 95% confidence intervals (Loftus & Masson, 1994; Masson & Loftus, 2003).

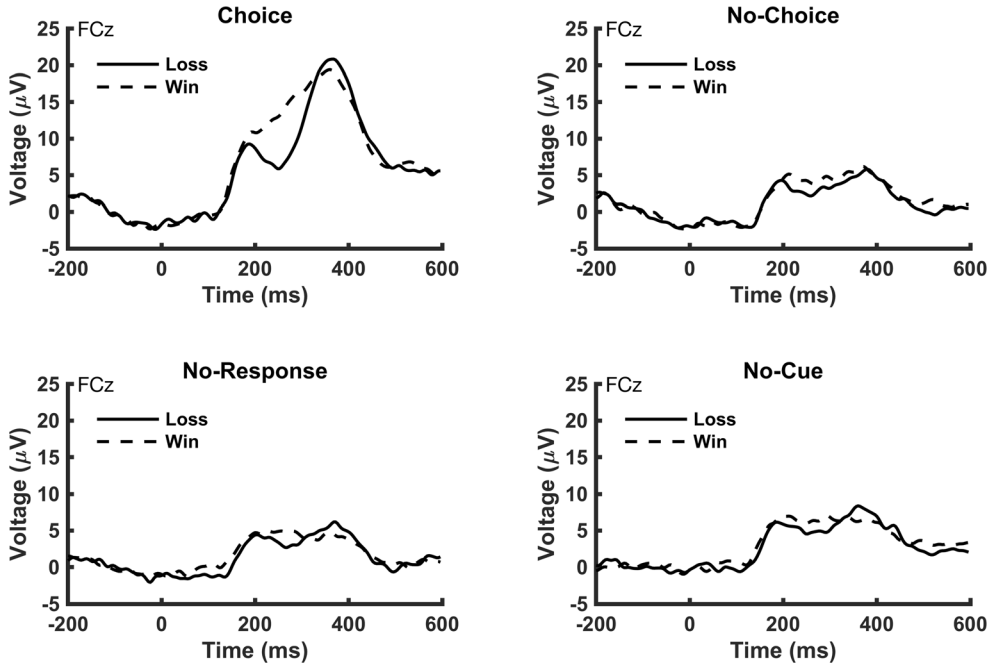

Fig. 2 Feedback-locked grand average waveforms at electrode FCz, for each task

Results

Reward Positivity

There was a significant effect of task on reward positivity, $F(3,69) = 10.67$, $p < 0.001$, $\eta_p^2 = 0.32$. A reward positivity was observed in the choice task ($t(23) = 6.16$, $p < 0.001$, Cohen's $d = 1.26$. A small reward positivity was present in both the no-choice task and no-response task according to our existence test (no-choice: $t(23) = 2.12$, $p = 0.046$, Cohen's $d = 0.43$; no-response: $t(23) = 2.11$, $p = 0.046$, Cohen's $d = 0.43$), but did not reach significance for the no-cue task ($t(23) = 1.98$,

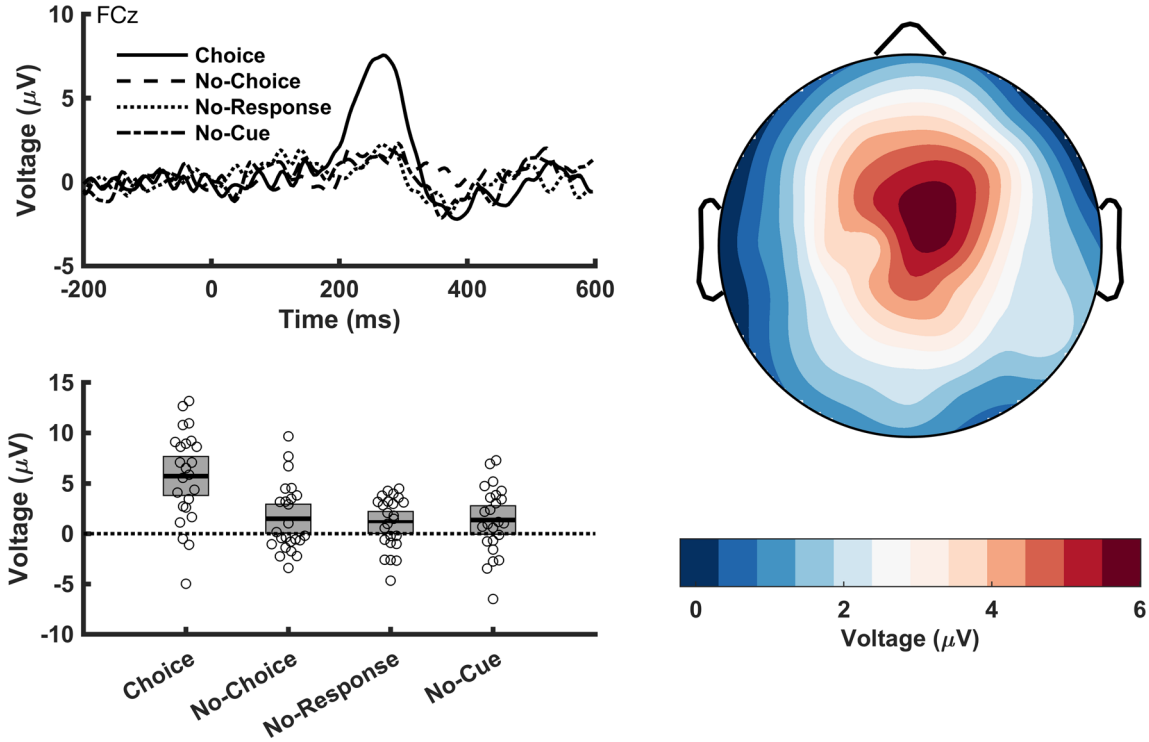

Fig. 3 Reward positivity results. Grand average difference waveforms (win minus loss) for each task (top left). Mean reward positivity scores for each task, with 95% confidence intervals (bottom left). Scalp topography of the reward positivity for the choice condition (right)

$p = 0.059$, Cohen's $d = 0.40$). However, all three of our conditions of interest (no-choice, no-response, and no-cue) had comparable effect sizes. See Figure 3 and Table 1 for exact reward positivity amplitudes.

P300

There was no interaction between task and outcome for the P300, $F(3,69) = 1.06$, $p = 0.37$, $\eta_g^2 = 0.002$, $\eta_p^2 = 0.044$. There was no main effect of feedback on the P300, $F(1,23) = 2.07$, $p = 0.16$, $\eta_g^2 = 0.006$, $\eta_p^2 = 0.082$. There was a main effect of task on P300 amplitude (enhanced for the choice task relative to other tasks), $F(3,69) = 54.40$, $p < 0.001$, $\eta_g^2 = 0.45$, $\eta_p^2 = 0.70$ (Figure 4).

Discussion

The results of the present study suggest that agency—the sense of control over our actions and their outcomes—affects the generation of a neural prediction error signal. In other words, our data support the notion that prediction error signals originating within medial-frontal cortex (Holroyd & Coles, 2002) are indicative of a volitional RL agent trying to learn the value of its actions.

In line with previous work, our analysis of the neural response to feedback in a two-armed bandit task revealed an ERP component with a timing and scalp topography consistent with the reward positivity (Holroyd & Coles, 2002; Yeung & Holroyd, 2005; Proudfit, 2015). According to the RL account of the reward positivity, this signal reflects an RL prediction error used to update action values (although other accounts exist, e.g., the conflict monitoring hypothesis: Yeung, Botvinick, & Cohen, 2004). The RL theory of the reward positivity might therefore predict a key role of choice in generating this neural signal. Consistent with this prediction, previous studies have observed that outcomes beyond our control elicit a reduced reward positivity compared to outcomes following a choice (Mühlberger et al., 2017; Yeung et al., 2005). We also observed a neural signal reminiscent of the reward positivity in the absence of choice, albeit with a much smaller effect size compared to previous work (Mühlberger et al., 2017). This signal was greatly reduced compared with our control condition in which participants made choices.

Table 1 Reward positivity scores for each condition, in microvolts

	M	*SD*	95% CI	Cohen's *d*
Choice	5.72	4.55	[3.80, 7.65]	1.26
No-choice	1.49	3.43	[0.04, 2.93]	0.43
No-response	1.11	2.59	[0.02, 2.21]	0.43
No-cue	1.36	3.36	[−0.06, 2.78]	0.40

To manipulate agency, we not only removed choice but also action. This was done in part to replicate previous work (Yeung et al., 2005) but also because of evidence that our sense of agency may work retrospectively. Actions that lead to unintended outcomes can be reframed as intentional after the fact (Johansson, Hall, Sikstrom, & Olsson, 2005). More importantly, an action in the absence of a choice can still be reinforced, thus engaging RL systems within the brain. We therefore predicted that the removal of choice and action from our task would result in a further reduction of the reward positivity. Unlike previous literature, the removal of action did not result in further reduction of the reward positivity (Yeung et al., 2005). Similar effect sizes were seen in in both our no-choice and no-response conditions, subordinating the contribution of action to the reward positivity. Finally, because RL systems are sensitive to cues, we introduced a no-cue condition designed to push the RL theory of the reward positivity to its limits. As others have shown, presenting a predictive cue sets up an expectation that impacts the reward positivity (Donkers et al., 2005; Krigolson et al., 2013). If the doors in our task served as such cues, then their removal should have resulted a reduced or absent reward positivity. Once again, however, we observed an effect size in our no-cue condition that was similar to our no-choice and no-response conditions. Thus, a major factor in generating the reward positivity (at least in this study) appears to be choice.

The observation that our sense of agency may work retrospectively is especially relevant to studies that contrast a choice condition (e.g., picking a card) to a no-choice condition in which participants respond only to initiate a random event (e.g., spinning a roulette wheel). Problem gamblers, for example, will mistakenly view random outcomes as under their control (illusion of control: Langer, 1975).[2] It is therefore possible that participants in previous "choice versus no-choice" experiments might have experienced some sense of agency when initiating random outcomes, accounting for the moderate-sized reward positivities seen in those studies (Mühlberger et al., 2017; Yeung et al., 2005). The current experimental design, however, left little doubt as to when participants were not in control; in two of our experimental conditions (no-choice and no-response) participants were told that the computer would select a door. This instruction was emphasized by the animation, on each trial, of a mouse cursor moving toward one of the doors. We speculate that these design details may have emphasized non-agency (the sense that participants were not in control) within our no-choice and no-

[2] Interestingly, problem gamblers also display an enhanced reward positivity, relative to controls, following a random outcome (Oberg, Christie, and Tata, 2011).

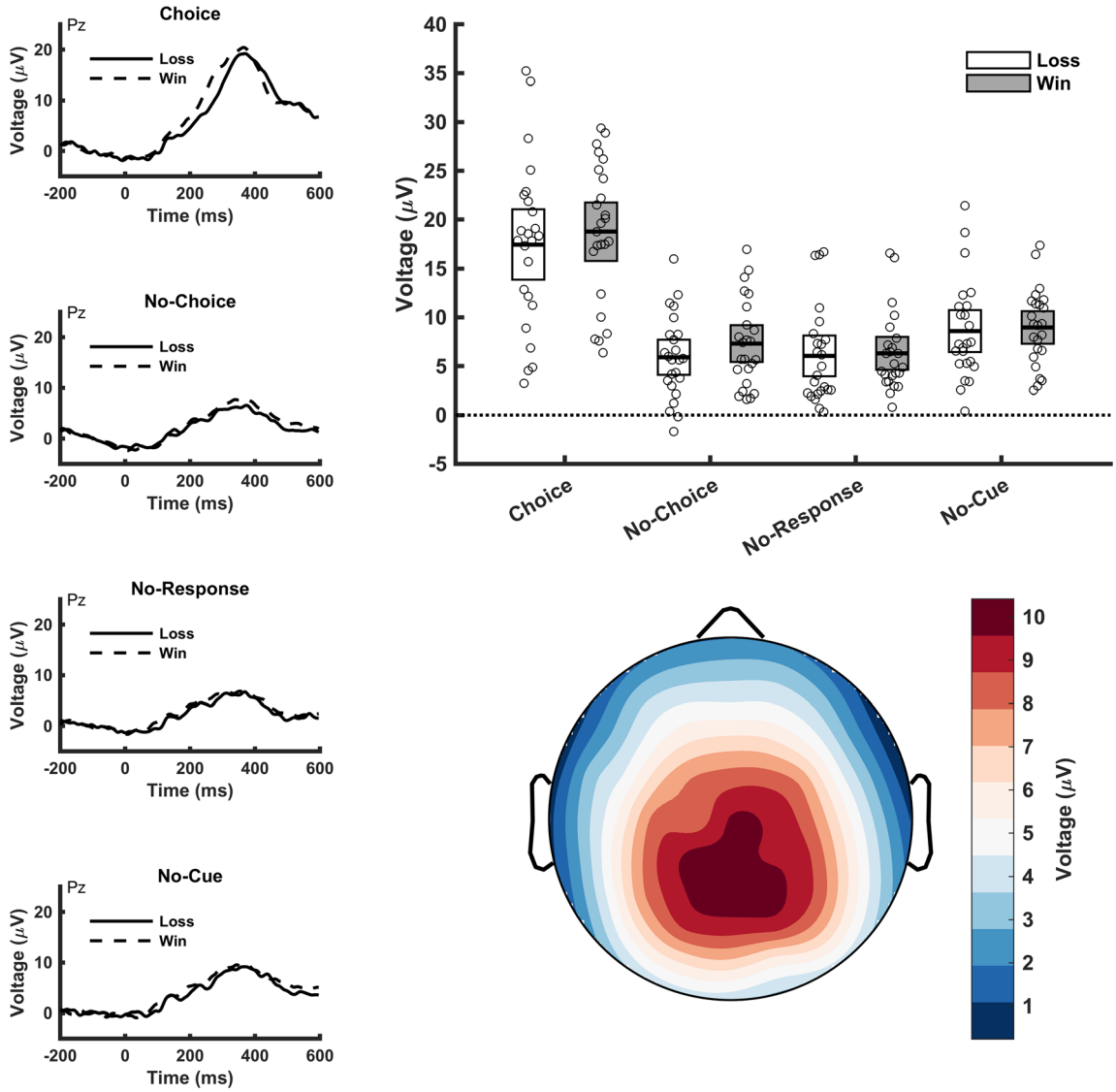

Fig. 4 P300 results. Feedback-locked grand average waveforms at electrode Pz, for each task (left). Mean P300 scores for each task, with 95% confidence intervals (top right). Scalp topography of the P300 for the mean of all conditions (bottom right). There was an overall P300 increase in the choice task relative to the other tasks

response conditions, accounting for the extremely small no-choice and no-response reward positivities that we observed.

Although we have highlighted agency, other factors are likely involved in our observed attenuation of the reward positivity. One such candidate, motivation, was investigated by Yeung et al. (2005). Their participants reported, via survey, that outcomes in the absence of choice were less interesting compared to outcomes following a choice. Furthermore, Yeung et al. (2005) noted that the degree to which participants' interest differed between tasks was predictive of the degree to which their reward positivity changed. In other words, participants who found the choice task more interesting had a larger reward positivity in the choice task, and participants who found the no-choice task more interesting had a larger reward positivity in the no-choice task. Could our reward positivity results be affected similarly? Previous research suggests that the P300 is affected by factors related to motivation. For example, larger P300s are elicited when participants are told their results will be compared with their peers' results (Carrillo-de-la-Peña & Cadaveira, 2000) and when money is at stake (Begleiter, Porjesz, Chou, & Aunon, 1983; Schmitt, Ferdinand, & Kray, 2015). Additionally, P300 magnitude correlates with reward magnitude (Goldstein et al., 2006; Meadows, Gable, Lohse, & Miller, 2016; Yeung & Sanfey, 2004) and self-reported motivation (Kleih et al., 2010). Like Yeung et al. (2005), we observed an enhanced P300 in the choice condition (the default doors task) compared with our other conditions. A motivation account of our results might suggest that this was because our participants were less motivated in the absence of choice. Although motivational

effects on the reward positivity are still an open area of research, we cannot rule out the possibility that they may have played a role here.

Although somewhat surprising given previous research, our data highlight the importance of agency in generating the reward positivity, a component of the human ERP thought to reflect an RL prediction error (Holroyd & Coles, 2002). These data provide further support for the existence of an RL system within the human brain tasked with learning the values of actions.

Acknowledgements This research was supported by the University of Victoria Neuroeducation Network (first author) and the Natural Sciences and Engineering Research Council of Canada (third author).

References

Begleiter, H., Porjesz, B., Chou, C. L., & Aunon, J. I. (1983). P3 and Stimulus Incentive Value. *Psychophysiology*, *20*(1), 95–101.

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It Was Not MY Fault: Event-Related Brain Potentials in Active and Observational Learning from Feedback. *Cerebral Cortex*, *20*(12), 2874–2883. https://doi.org/10.1093/cercor/bhq038

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Bress, J. N., Foti, D., Kotov, R., Klein, D. N., & Hajcak, G. (2013). Blunted neural response to rewards prospectively predicts depression in adolescent girls. *Psychophysiology, 50*(1), 74–81.

Carrillo-de-la-Peña, M. T., & Cadaveira, F. (2000). The effect of motivational instructions on P300 amplitude. *Neurophysiologie Clinique/Clinical Neurophysiology, 30*(4), 232–239.

Collins, A. G. E., & Frank, M. J. (2018). Within- and across-trial dynamics of human EEG reveal cooperative interplay between reinforcement learning and working memory. *Proceedings of the National Academy of Sciences, 115*(10), 2502–2507.

Cumming, G. (2014). The New Statistics: Why and How. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8*(12), 1704–1711. https://doi.org/10.1038/nn1560

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/J.JNEUMETH.2003.10.009

Donkers, F. C. L., Nieuwenhuis, S., & van Boxtel, G. J. M. (2005). Mediofrontal negativities in the absence of responding. *Cognitive Brain Research*, *25*(3), 777–787. https://doi.org/10.1016/j.cogbrainres.2005.09.007

Dunning, J. P., & Hajcak, G. (2007). Error-related negativities elicited by monetary loss and cues that predict loss. *Neuroreport*, *18*(17), 1875–1878. https://doi.org/10.1097/WNR.0b013e3282f0d50b

Fischer, A. G., & Ullsperger, M. (2013). Real and Fictive Outcomes Are Processed Differently but Converge on a Common Adaptive Mechanism. *Neuron, 79*(6), 1243–1255.

Goldstein, R. Z., Cottone, L. A., Jia, Z., Maloney, T., Volkow, N. D., & Squires, N. K. (2006). The effect of graded monetary reward on cognitive event-related potentials and behavior in young healthy adults. *International Journal of Psychophysiology, 62*(2), 272–279.

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196–207. https://doi.org/10.1038/nrn.2017.14

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037//0033-295X.109.4.679

Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *Neuroreport*, *22*(5), 249–252. https://doi.org/10.1097/WNR.0b013e328345441d

Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to Detect Mismatches Between Intention and Outcome in a Simple Decision Task. *Science*, *310*(5745), 116–119. https://doi.org/10.1126/science.1111709

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*(2), 163–178.

Kleih, S. C., Nijboer, F., Halder, S., & Kübler, A. (2010). Motivation modulates the P300 amplitude during brain–computer interface use. *Clinical Neurophysiology*, *121*(7), 1023–1031.

Krigolson, O. E. (2017). Event-related brain potentials and the study of reward processing: Methodological considerations. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2017.11.007

Krigolson, O. E., Hassall, C. D., & Handy, T. C. (2013). How We Learn to Make Decisions: Rapid Propagation of Reinforcement Learning Prediction Errors in Humans. *Journal of Cognitive Neuroscience*, *26*(3), 635–644. https://doi.org/10.1162/jocn_a_00509

Krigolson, O. E., & Holroyd, C. B. (2007). Predictive information and error processing: The role of medial-frontal cortex during motor control. *Psychophysiology*, *44*(4), 586–595. https://doi.org/10.1111/j.1469-8986.2007.00523.x

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00863

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, *32*(2), 311–328.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 476–490.

Martin, L. E., & Potts, G. F. (2011). Medial Frontal Event Related Potentials and Reward Prediction: Do Responses Matter? *Brain and Cognition*, *77*(1), 128–134. https://doi.org/10.1016/j.bandc.2011.04.001

Masson, M. E., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *57*(3), 203.

Meadows, C. C., Gable, P. A., Lohse, K. R., & Miller, M. W. (2016). The effects of reward magnitude on reward processing: An averaged and single trial event-related potential study. *Biological Psychology, 118*, 154–160.

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798.

Mühlberger, C., Angus, D. J., Jonas, E., Harmon-Jones, C., & Harmon-Jones, E. (2017). Perceived control increases the reward positivity and stimulus preceding negativity. *Psychophysiology*, *54*(2), 310–322. https://doi.org/10.1111/psyp.12786

Mulligan, E. M., Flynn, H., & Hajcak, G. (2018). Neural response to reward and psychosocial risk factors independently predict antenatal depressive symptoms. *Biological Psychology*. In press, corrected proof.

O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, Reward, and Decision Making. *Annual Review of Psychology*, *68*(1), 73–100. https://doi.org/10.1146/annurev-psych-010416-044216

Oberg, S. A. K., Christie, G. J., & Tata, M. S. (2011). Problem gamblers exhibit reward hypersensitivity in medial frontal cortex during gambling. *Neuropsychologia*, *49*(13), 3768–3775. https://doi.org/10.1016/j.neuropsychologia.2011.09.037

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434–447. https://doi.org/10.1037/1082-989X.8.4.434

Pavlov, P. I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, *17*(3), 136–141. (Original work published in 1927) https://doi.org/10.5214/ans.0972-7531.1017309

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, *52*(4), 449–459.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, *2*, 64–99.

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. https://doi.org/10.1037/bul0000006

Schmitt, H., Ferdinand, N. K., & Kray, J. (2015). The influence of monetary incentives on context processing in younger and older adults: an event-related potential study. *Cognitive, Affective, & Behavioral Neuroscience, 15*(2), 416–434.

Sutton, R. S., & Barto, A. G. (1998). Reinforcement Learning: An Introduction. MIT Press, Cambridge.

Thorndike, E. (2017). Animal Intelligence: Experimental Studies. New York: Routledge. (Original work published in 1911)

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews*, *36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The Neural Basis of Error Detection: Conflict Monitoring and the Error-Related Negativity. *Psychological Review, 111*(4), 931–959. https://doi.org/10.1037/0033-295X.111.4.931

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP Correlates of Feedback and Reward Processing in the Presence and Absence of Response Choice. *Cerebral Cortex*, *15*(5), 535–544. https://doi.org/10.1093/cercor/bhh153

Yeung, N., & Sanfey, A. G. (2004). Independent Coding of Reward Magnitude and Valence in the Human Brain. *Journal of Neuroscience, 24*(28), 6258–6264. https://doi.org/10.1523/JNEUROSCI.4537-03.2004

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.