Better late than never? The effect of feedback delay on ERP indices of reward processing

**Anna Weinberg · Christian C. Luhmann ·
Jennifer N. Bress · Greg Hajcak**

Abstract The feedback negativity (FN), an early neural response that differentiates rewards from losses, appears to be generated in part by reward circuits in the brain. A prominent model of the FN suggests that it reflects learning processes by which environmental feedback shapes behavior. Although there is evidence that human behavior is more strongly influenced by rewards that quickly follow actions, in nonlaboratory settings, optimal behaviors are not always followed by immediate rewards. However, it is not clear how the introduction of a delay between response selection and feedback impacts the FN. Thus, the present study used a simple forced choice gambling task to elicit the FN, in which feedback about rewards and losses was presented after either 1 or 6 s. Results suggest that, at short delays (1 s), participants clearly differentiated losses from rewards, as evidenced in the magnitude of the FN. At long delays (6 s), on the other hand, the difference between losses and rewards was negligible. Results are discussed in terms of eligibility traces and the reinforcement learning model of the FN.

Keywords Reward · ERP

Feedback from the environment allows us to modify our behaviors and respond adaptively to our changing surroundings. As we learn to associate behaviors with their outcomes, actions followed by positive feedback (e.g., rewards) tend to be favored, while those followed by negative feedback (e.g., losses) are less likely to be repeated (Schultz et al., 1995; Sutton & Barto, 1998; Thorndike, 1927). Recently, the neural mechanisms underlying the evaluation of external feedback have been receiving increased attention, and evidence from event-related brain potentials (ERPs) may be helpful in clarifying processes by which rewards and losses are distinguished and evaluated. In particular, ERP research has focused on the feedback negativity (FN), a neural response peaking approximately 250–300 ms following the presentation of feedback, which is maximal at frontocentral recording sites (Holroyd & Coles, 2002; Miltner, Braun, & Coles, 1997) and which is thought to be generated in part by the anterior cingulate cortex (ACC; Gehring & Willoughby, 2002; Holroyd & Coles, 2002; Knutson, Westdorp, Kaiser, & Hommer, 2000; Luu, Tucker, Derryberry, Reed, & Poulsen, 2003; Miltner et al., 1997).

The FN has frequently been studied in the context of laboratory gambling tasks, in which participants make responses in order to win or lose money; feedback indicating monetary rewards, losses, or nonrewards (i.e., neither winning nor losing money) is presented on each trial following response selection. The FN differentiates unfavorable feedback (e.g., feedback indicating monetary loss) from favorable feedback (e.g., feedback indicating monetary reward).

Traditionally, the FN has been conceptualized as a negative deflection in the waveform that is enhanced by losses and absent for gains (e.g., Hajcak, Moser, Holroyd, & Simons, 2006; Holroyd & Coles, 2002). However, recent evidence suggests that, in addition to a negativity in response to losses, activity in the time range of the FN may also reflect an underlying positivity in response to rewards that is reduced or absent in response to losses (i.e., the reward positivity; Baker & Holroyd, 2011; Bernat, Nelson, Steele, Gehring, & Patrick, 2011; Bogdan, Santesso, Fagerness, Perlis, & Pizzagalli, 2011; Carlson, Foti, Harmon-Jones, Mujica-Parodi, & Hajcak, 2011; Foti & Hajcak, 2009; Foti, Weinberg, Dien, & Hajcak, 2011; Hewig et al., 2010; Holroyd, Krigolson, & Lee, 2011; Holroyd, Pakzad-Vaezi, & Krigolson, 2008). Thus, the trial-averaged FN observed following loss feedback might reflect the activity of two independent but overlapping processes: a

A. Weinberg (✉) · C. C. Luhmann · J. N. Bress · G. Hajcak
Department of Psychology, Stony Brook University,
Stony Brook, NY 11794-2500, USA
e-mail: anna.weinberg@gmail.com

negative deflection in the waveform that is enhanced by losses and the *absence* of a positive-going reward response following loss feedback (e.g., Bernat, Nelson, & Sommers, 2012; Bernat et al., 2011; Carlson et al., 2011; Foti et al., 2011). For this reason, it is often helpful to conceptualize the FN as a difference score, representing the degree to which rewards and losses are differentiated.

A prominent theory of the FN suggests that the component represents a reward prediction signal, reflecting the activity of a reinforcement learning system rooted in midbrain dopamine (DA) neurons (Holroyd & Coles, 2002). According to this theory, negative (e.g., errors or monetary loss) and positive (e.g., monetary gains) outcomes modulate mesencephalic DA activity—a system critical to reward processing (e.g., Delgado, 2007; Schultz, 2002)—which sends distinct signals to the ACC (Holroyd & Coles, 2002). Activity of this circuit is thus thought to monitor and adjust the strength of associations between actions and outcomes (Schultz et al., 1995), thereby training the ACC to select actions that maximize rewards.

Consistent with models of the FN that emphasize reward processing, emerging evidence suggests direct contributions from the basal ganglia to the magnitude of the FN (Carlson et al., 2011; Foti et al., 2011; Martin, Potts, Burton, & Montague, 2009). Furthermore, the FN appears to respond to reward in similar ways as areas of the mesencephalic DA circuit. For example, like regions of the striatum (Elliott, Newman, Longe, & Deakin, 2003), the FN differentiates rewards from losses, but it does not appear to be sensitive to the magnitude of those rewards or losses; in fact, the FN is equivalent for larger, as compared with smaller, losses (Hajcak, Moser, Holroyd, & Simons, 2006; Sato et al., 2005; Yeung & Sanfey, 2004). Likewise, expectations prior to the outcome impact the evaluation of feedback indicating rewards or losses, as reflected in activity in both the putamen (McClure, Berns, & Montague, 2003; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003) and the FN (Hajcak, Moser, Holroyd, & Simons, 2007; Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003; Walsh & Anderson, 2011), such that outcomes that violate reward predictions elicit a larger FN (Hajcak et al., 2007; Holroyd et al., 2003; Holroyd et al., 2008; Walsh & Anderson, 2011).

However, there is evidence that regions of the striatum do not merely passively assess the hedonic value of an outcome (e.g., respond to a reward received when the participant has not taken any actions). Rather, the striatum appears to respond preferentially to rewards that can be attributed to an action or sequence of actions leading to that outcome (Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; O'Doherty et al., 2004; Schultz, Tremblay, & Hollerman, 2000; Tricomi, Delgado, & Fiez, 2004). Likewise, the magnitude of the FN is enhanced when participants believe that rewards and punishments are contingent on their chosen motor response (Masaki,

Shibahara, Ogawa, Yamazaki, & Hackley, 2010; Yeung, Holroyd, & Cohen, 2005), suggesting that the FN, too, tracks action–outcome contingencies.

Yet, outside of the laboratory, rewards do not always immediately follow the actions that produce them. Instead, an action may be optimal in the sense that it reliably leads to rewards that are delivered only after some delay, perhaps even after additional subsequent actions have been taken. It may then be difficult to associate delayed rewards to the actions that produced them (Gallistel & Gibbon, 2000). Consistent with this possibility, behavioral evidence in humans suggests that behaviors are shaped more strongly by more proximal rewards (e.g., Bogacz, McClure, Li, Cohen, & Montague, 2007).

If the FN reflects reward-related modulation of behaviors, it might also be sensitive to the temporal delay between action and rewards. Although the FN has not yet been examined under conditions of delayed feedback, animal data suggest that rewards delivered after a long delay elicit reduced dopaminergic activity, as compared with rewards of the same value delivered after a short delay (e.g., Roesch, Calu, & Schoenbaum, 2007). Combined with data from studies modeling the role of DA in maintenance of action–outcome associations over a delay (Bogacz et al., 2007; Montague et al., 2004; Pan, Schmidt, Wickens, & Hyland, 2005), animal work further suggests that the FN should be enhanced for more rapidly delivered feedback and diminished for feedback presented at a delay. In order to test this hypothesis, we utilized a relatively straightforward gambling paradigm (e.g., Foti & Hajcak, 2009; Foti et al., 2011) in which participants attempted to guess which of two doors hid a monetary reward. Feedback indicating whether participants won or lost money was presented after each response, but the task was modified such that, in one condition, feedback was presented after a 1-s delay and, in the other condition, feedback was presented after a 6-s delay.

On the basis of the evidence discussed above, we predicted that the difference between reward and loss trials in the time range of the FN would be substantial at short delays (i.e., 1 s after response selection) but that, when a longer delay between action and outcome was introduced (i.e., 6 s after response selection), the difference would be diminished. In addition, we examined whether the P300, a positive-going deflection in the waveform occurring soon after the FN, would be impacted by delay. Although the P300 is inconsistently related to outcome valence (e.g., Foti & Hajcak, 2009, 2012; Foti et al., 2011; Goyer, Woldorff, & Huettel, 2008; Sato et al., 2005), it appears to index distinct aspects of feedback evaluation. Thus, the P300 was also evaluated in the present study in order to examine general orienting and attention to feedback, which might be reduced overall at the long delay.

Method

Participants

Twenty Stony Brook University undergraduates (11 female) participated in the study for course credit. The mean age of participants was 19.35 years (SD = 2.50); 55 % were Caucasian, 10 % were Hispanic, 25 % were Asian, 5 % were African-American, and 5 % described themselves as "other." All participants were screened for a history of neurological disorders.

Procedure

Subsequent to verbal instructions indicating that they would be engaging in multiple tasks while electroencephalograph (EEG) recordings were made, participants were seated, and EEG sensors were attached. The EEG was recorded while tasks were administered on a Pentium class computer, using Presentation software (Neurobehavioral Systems, Inc.) to control the presentation and timing of all stimuli.

Gambling task

The gambling task consisted of six blocks of 16 trials. Half the blocks were associated with a short feedback delay (1 s), and half were associated with a long feedback delay (6 s). The order of the blocks was counterbalanced across participants, such that half the participants received an ABABAB order and half a BABABA order. On each trial within the blocks, participants were shown a graphic displaying two doors (occupying 61° of the visual field vertically and 81° horizontally) and were told to choose which door they wanted to open. Participants were told to press the left mouse button to choose the left door or the right mouse button to choose the right door. Following each choice, a feedback stimulus appeared on the screen informing the participants whether they had won or lost money on that trial. A green arrow pointing up indicated a reward of $0.50, while a red arrow pointing down indicated a loss of $0.25 (each occupying 31° of the visual field vertically and 11° horizontally). These monetary values were selected in order to ensure that the subjective values of gains and losses were equivalent (Tversky & Kahneman, 1992). Positive feedback was given on 8 trials in each block (i.e., 50 % of trials) in both the short and long feedback delay conditions. Twenty-four trials were therefore associated with reward feedback in the short-delay condition, and 24 with loss feedback in the short-delay condition; likewise, 24 trials each were associated with reward and loss feedback in the long-delay condition. Within each block, feedback was presented in a random order for each participant. The order and timing of all stimuli were as follows: (1) The graphic of two doors was presented until a response was made; (2) a fixation mark was presented for either 1 s (short-delay blocks) or 6 s (long-delay blocks); (3) a feedback arrow was presented for 2 s; (4) a fixation mark was presented for 1.5 s; and (5) "click for the next round" was presented until a response was made. After three blocks, feedback indicating the amount of money participants had won up to that point was presented.

Psychophysiological recording, data reduction, and analysis

Continuous EEG recordings were collected using an elastic cap and the ActiveTwo BioSemi system (BioSemi, Amsterdam, Netherlands). Thirty-four electrodes were used, based on the 10/20 system, as well as two electrodes on the right and left mastoids. An electrooculogram generated from eye movements and eyeblinks was recorded using four facial electrodes. Horizontal eye movements were measured via two electrodes located approximately 1 cm outside the outer edge of the right and left eyes. Vertical eye movements and blinks were measured via two electrodes placed approximately 1 cm above and below the right eye. The EEG signal was preamplified at the electrode to improve the signal-to-noise ratio and was amplified with a gain of 1× by a BioSemi ActiveTwo system (BioSemi, Amsterdam). The data were digitized at 24-bit resolution with an LSB value of 31.25 nV and a sampling rate of 512 Hz, using a low-pass fifth-order sinc filter with a −3 dB cutoff point at 104 Hz. Each active electrode was measured online with respect to a common mode sense (CMS) active electrode, located between PO3 and POz, producing a monopolar (nondifferential) channel. CMS forms a feedback loop with a paired driven right leg electrode, located between POz and PO4, reducing the potential of the participants and increasing the common mode rejection rate. Offline, all data were referenced to the average of the left and right mastoids and were band-pass filtered with low and high cutoffs of 0.1 and 30 Hz, respectively; eye blink and ocular corrections were conducted per Gratton, Coles, and Donchin (1983).

A semiautomatic procedure was employed to detect and reject artifacts. The criteria applied were a voltage step of more than 50.0 µV between sample points, a voltage difference of 300.0 µV within a trial, and a maximum voltage difference of less than 0.50 µV within 100-ms intervals. Visual inspection of the data was then conducted to detect and reject any remaining artifacts.

The EEG was segmented for each trial beginning 200 ms before each response onset and continuing for 1,000 ms (i.e., for 800 ms following feedback); a 200-ms window from −200 to 0 ms prior to feedback onset served as the baseline. The FN appears maximal around 300 ms at central sites; therefore, the FN was scored as the average activity at FCz, between 250 and 350 ms. Following the FN, the P300 is maximal between 350 and 600 ms, and it was scored as the average activity at Pz in this time window.

The FN and the P300 were statistically evaluated using SPSS (Version 17.0) Repeated-Measures General Linear Model software. In the case of both the FN and the P300, a 2 (outcome: reward, loss) × 2 (delay: short, long) ANOVA was conducted. Greenhouse–Geisser corrections were applied to p values associated with multiple-df, repeated measures comparisons when necessitated by violation of the assumption of sphericity; p-values were adjusted with the Bonferroni correction for multiple post hoc comparisons. Finally, post hoc paired-samples t-tests were conducted following significant interactions determined by the ANOVAs.

Results

The FN

Figure 1 presents the grand average stimulus-locked ERPs at FCz and Pz for reward and loss feedback, as well as the differences between reward and loss feedback, for long-delay (left) and short-delay (right) trials. Figure 2 presents topographic maps depicting voltage differences (in μV) for losses minus rewards in the long (left) and short (right) conditions, in the time range of the FN (i.e., 250–350 ms). Average ERP values from the two conditions are presented in Table 1. The magnitude of the ERPs in the time window of the FN did not vary, overall, as a function of feedback

delay, $F(1, 19) = 0.95$, $p > .05$, $\eta_p^2 = .05$, or as a function of feedback type, $F(1, 19) = 4.15$, $p > .05$, $\eta_p^2 = .18$. However, as is suggested by Figs. 1 and 2, the effect of feedback type varied significantly as a function of delay, $F(1, 19) = 5.41$, $p < .05$, $\eta_p^2 = .22$, such that losses were characterized by a relative negativity, as compared with gains, in the short-delay condition, $t(19) = 2.94$, $p < .01$, but not in the long-delay condition, $t(19) = 0.37$, $p > .05$.

The P300

P300 amplitudes following short and long delays are presented in Table 1. The magnitude of the P300 did not vary significantly as a function of feedback delay, $F(1, 19) = 3.49$, $p > .05$, $\eta_p^2 = .16$, or as a function of feedback type, $F(1, 19) < 1$, $\eta_p^2 = .01$; furthermore, the interaction between feedback type and delay did not reach significance, $F(1, 19) = 1.36$, $p > .05$, $\eta_p^2 = .07$.

Discussion

In the present study, monetary losses and rewards were robustly differentiated when responses were followed by rapid feedback (i.e., 1-s delay), such that losses elicited a reliably larger negativity than did gains. When feedback was delayed (6-s delay), however, the responses to losses and
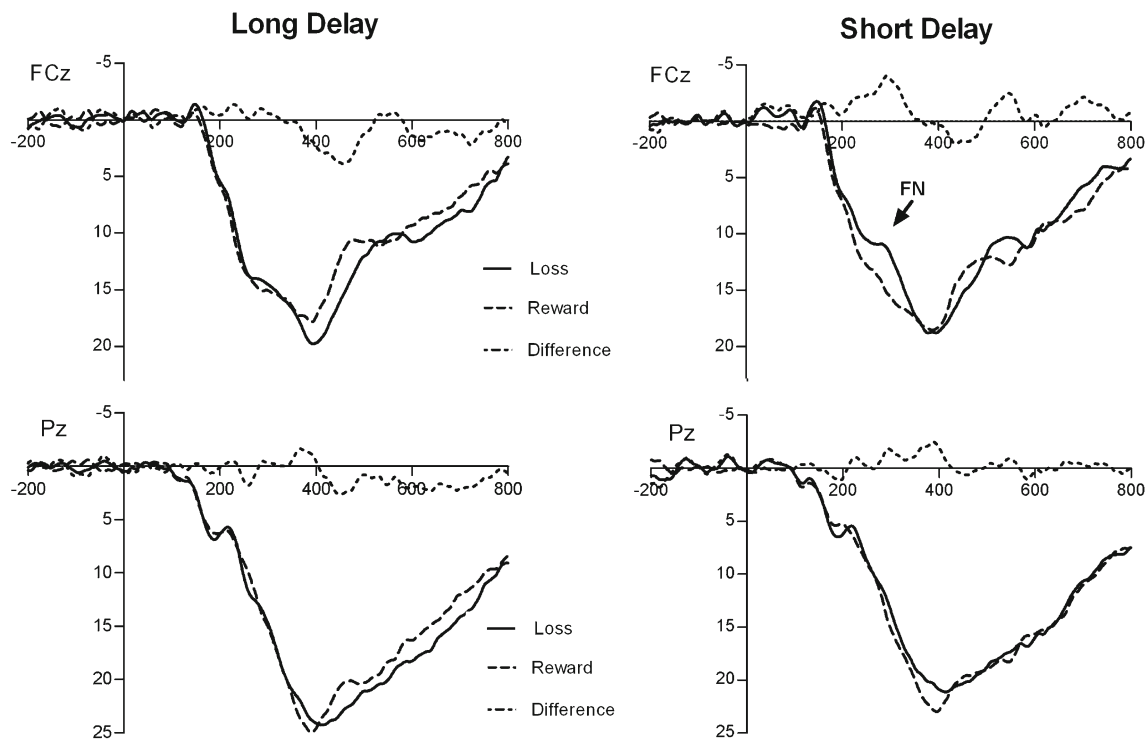

Fig. 1 Response-locked ERP waveforms at FCz (top) and Pz (bottom) comparing reward and loss trial waveforms, as well as the difference between them, in the long (left) and short (right) feedback delay conditions. For each panel, response onset occurred at 0 ms, and negative is plotted up

Fig. 2 Topographic maps depicting differences (in μV) between response to rewards and losses in the long-delay (left) and short-delay (right) conditions in the time range of the FN (250–350 ms)

Loss Minus Reward

Long Delay   Short Delay

250-350 ms

-3.1 μV   0μV   1.5μV
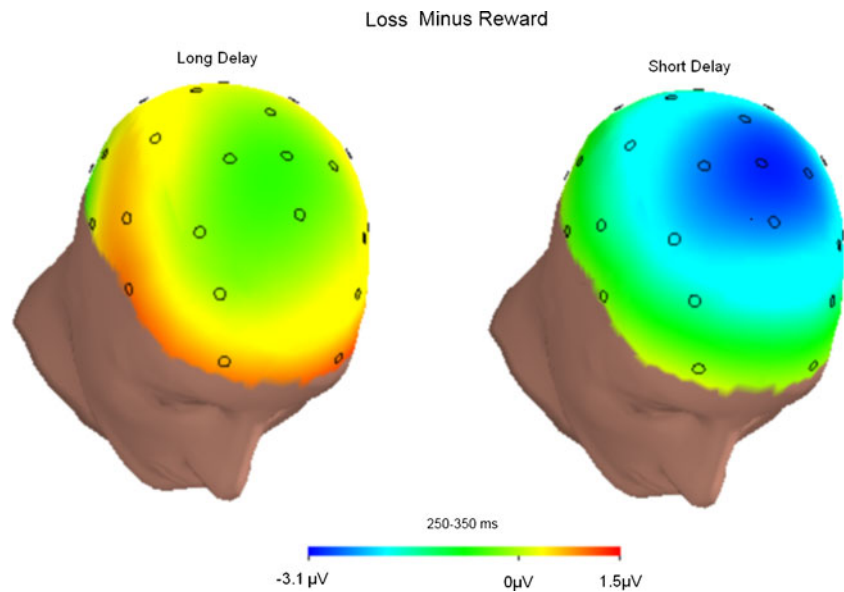
rewards were nearly identical, suggesting that, at greater delays, the FN is diminished. This suggests that the introduction of a delay between responses and feedback may interfere with the ability to maintain associations between actions and the resulting rewards. Furthermore, this effect appeared specific to the FN; delay did not appear to impact the magnitude of the P300, although with a relatively small sample size ($n = 20$), the study may have been underpowered to detect small but significant effects.

The results of this study have real-world implications; oftentimes, optimal actions are not followed by immediate rewards. The challenge, then, is to accurately associate action with a distal or delayed outcome, yet behavioral data suggest that humans often fare poorly at this (e.g., Bogacz et al., 2007). The results of the present study are consistent with behavioral findings: The neural response to rewards seen at shorter delays is not present at longer delays. This would suggest that people are simply unable to associate actions with outcomes that occur after a few seconds. And yet people often do put off immediate rewards for the sake of longer-term goals and, presumably, are able to attribute

their longer-term success to their earlier actions. This raises the possibility that the system by which these longer-term goals are processed is separate from the system that processes immediate rewards.

The possibility of distinct systems is consistent with neuroimaging studies of anticipatory reward processing, which have suggested that the choice to receive an immediate monetary reward and the choice to receive a delayed reward are processed in distinct areas of the brain. Namely, while decisions about immediate rewards are processed by the midbrain DA system—a system thought to be critically involved in the generation of the FN—decisions about delayed rewards are associated with activity in prefrontal and parietal regions associated with higher-level cognitive control (McClure, Ericson, Laibson, Loewenstein, & Cohen, 2007; McClure, Laibson, Loewenstein, & Cohen, 2004). It is possible that the ability to connect longer-term rewards to their associated actions requires higher-level reasoning, which might require cortical areas not necessary for the more automatic processing of immediate rewards. Future studies might combine ERP measures with imaging or source localization techniques to clarify

Table 1 Means and standard deviations for the FN and P3 in the short-delay (left) and long-delay (right) conditions

	Short delay						Long delay					
	Reward		Loss		Difference (loss minus reward)		Reward		Loss		Difference (loss minus reward)	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
FN (FCz) (250–350 ms)	15.26*	7.86	12.68*	9.60	−2.58	3.92	15.06	7.11	14.74	7.66	−0.32	3.78
P300 (Pz) (350–600 ms)	19.47	6.33	19.00	7.57	0.47	4.84	20.72	7.30	21.65	7.52	0.93	3.19

* p<.05

the process by which people learn associations between delayed rewards and the actions that caused them.

The present study used a relatively unambiguous reward paradigm in which, on each trial, regardless of their response selection, participants had an equal chance of gaining or losing money. This has the advantage of isolating the impact of delay on the processing of rewards versus losses. However, although reward processing is thought to be integral to learning (Schultz et al., 1995; Sutton & Barto, 1998), the design also limits the extent to which we can generalize from the present results to learning processes. Thus, future studies might directly examine the impact of a delay between response and feedback in paradigms requiring more complex learning calculations.

Nonetheless, given models of the FN and reinforcement learning, the results of the present study have implications for learning paradigms. For example, Holroyd and Coles (2002) have postulated a role for eligibility traces—or simple memories of actions that have been taken in the recent past—in their computational model of the FN. Eligibility traces represent the degree to which previously taken actions are eligible for consideration when attempting to attribute rewards to actions. When an action is taken, an eligibility trace for that action is initiated and immediately begins to decay with time (Doya, 2008). Holroyd and Coles suggested that the FN represents (1) a negative reward prediction error signal that coincides with (2) an eligibility trace presumably represented in the ACC. Indeed, the simulations they reported simply multiply the prediction error signals expected on a given trial by the sum of all current eligibility traces. Thus, if there were no significant eligibility traces—for example, if previous actions were taken only in the distant past—then their model would predict no FN even in the presence of significant prediction error.

Interestingly, and unlike traditional approaches, the eligibility traces employed by Holroyd and Coles (2002) do not decay systematically over time. Instead, their traces are essentially static, allowing a selected action to remain eligible indefinitely (i.e., until the end of the trial in their simulations). These authors argued that the eligibility traces embodied in their model represent, not intrinsic neuronal dynamics (cf. Bogacz et al., 2007), but something more akin to working memory, possibly subserved by persistent activity in cortical circuits (e.g., the dorsolateral prefrontal cortex). The present results are clearly inconsistent with this proposal. We observed robust differentiation of losses from rewards when feedback was delayed 1 s, but not when it was delayed 6 s. Furthermore, unlike studies in which the action–feedback association straddles other events of interest—for example, other cues (as in Pan et al., 2005), other actions (as in Bogacz et al., 2007), or both (as in Walsh & Anderson, 2011)—the delays in the present study were empty (participants spent them staring at a blank screen). Thus, there were no strong sources of interference that might have disrupted the eligibility traces associated with prior actions.

It is worth noting again that in the present study, rewards and losses were presented randomly and were not consistently associated with specific actions; thus, participants were unable to form strong associations between actions and outcomes. Therefore, although the present results appear inconsistent with the model suggested by Holroyd and Coles (2002), future studies will be needed in order to verify that the same pattern of results emerges in the context of a reinforcement learning paradigm. Additionally, future studies might examine the effects of a randomized design, in which the duration of feedback delay on each trial is unpredictable, in order to examine what role expectations about the timing of feedback might play in maintaining representations of action–outcome contingencies.

In sum, the present results are the first demonstration that action updating is subject to finite time constraints. However, because the present study evaluated the FN only at two time delays, the mechanisms and time course of this effect are unclear. Future studies might also profitably investigate multiple time delays to more thoroughly evaluate the time course of this decay.

References

Baker, T., & Holroyd, C. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology, 87,* 25–34.

Bernat, E., Nelson, L., & Baskin-Sommers, A. (2012). *Time-frequency theta and delta measures index separable components of feedback processing in a gambling task.* Manuscript under review.

Bernat, E., Nelson, L., Steele, V., Gehring, W., & Patrick, C. (2011). Externalizing psychopathology and gain–loss feedback in a simulated gambling task: Dissociable components of brain response revealed by time-frequency analysis. *Journal of Abnormal Psychology, 120,* 352–364.

Bogacz, R., McClure, S. M., Li, J., Cohen, J. D., & Montague, P. R. (2007). Short-term memory traces for action bias in human reinforcement learning. *Brain Research, 1153,* 111–121. doi:10.1016/j.brainres.2007.03.057

Bogdan, R., Santesso, D., Fagerness, J., Perlis, R., & Pizzagalli, D. (2011). Corticotropin-releasing hormone receptor type 1 (CRHR1) genetic variation and stress interact to influence reward learning. *Journal of Neuroscience, 31,* 13246–13254.

Carlson, J., Foti, D., Harmon-Jones, E., Mujica-Parodi, L., & Hajcak, G. (2011). The neural processing of rewards in the human striatum: Correlation of fMRI and event-related potentials. *NeuroImage, 57,* 1608–1616. doi:10.1016/j.neuroimage.2011.05.037

Delgado, M. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences, 1104,* 70–88. doi:10.1196/annals.1390.002

Delgado, M., Nystrom, L., Fissell, C., Noll, D., & Fiez, J. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology, 84,* 3072–3077.

Doya, K. (2008). Modulators of decision making. *Nature Neuroscience, 11,* 410–416. doi:10.1038/nn207

Elliott, R., Newman, J. L., Longe, O. A., & Deakin, J. F. (2003). Differential response patterns in the striatum and orbitofrontal

cortex to financial reward in humans: A parametric functional magnetic resonance imaging study. *Journal of Neuroscience, 23,* 303–307.

Foti, D., & Hajcak, G. (2009). Depression and reduced sensitivity to non-rewards versus rewards: Evidence from event-related potentials. *Biological Psychology, 81,* 1–8. doi:10.1016/j.biopsycho.2008.12.004

Foti, D., & Hajcak, G. (2012). Genetic variation in dopamine moderates neural response during reward anticipation and delivery: Evidence from event-related potentials. *Psychophysiology, 49,* 617–626.

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from non-rewards: Temporospatial principal components analysis and source localization of the feedback negativity. *Human Brain Mapping, 32,* 2207–2216. doi:10.1002/hbm.21182

Gallistel, C., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107,* 289–344. doi:10.1037/033-295X.107.2.289

Gehring, W., & Willoughby, A. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295,* 2279–2282. doi:10.1126/science.1066893

Goyer, J. P., Woldorff, M. G., & Huettel, S. A. (2008). Rapid electrophysiological brain responses are influenced by both valence and magnitude of monetary rewards. *Journal of Cognitive Neuroscience, 20,* 2058–2069. doi:10.1162/jocn.2008.20134

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55,* 468–484.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology, 71,* 148–154. doi:10.1016/j.biopsycho.2005.04.001

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology, 44,* 905–912. doi:10.1111/j.1469-8986.2007.00567

Hewig, J., Kretschmer, N., Trippe, R., Hecht, H., Coles, M., Holroyd, C., . . . Miltner, W. (2010). Hypersensitivity to reward in problem gamblers. *Biological Psychiatry, 67,* 781–783.

Holroyd, C., & Coles, M. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109,* 679–709. doi:10.1037//0033-295X.109.4.67

Holroyd, C., Krigolson, O., & Lee, S. (2011). Reward positivity elicited by predictive cues. *NeuroReport, 22,* 249–252.

Holroyd, C., Nieuwenhuis, S., Yeung, N., & Cohen, J. (2003). Errors in reward prediction are reflected in the event-related brain potential. *NeuroReport, 14,* 2481–2484. doi:10.1097/01.wnr.0000099601.41403.a

Holroyd, C., Pakzad-Vaezi, K., & Krigolson, O. (2008). The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology, 45,* 688–697. doi:10.1111/j.1469-8986.2008.00668.x

Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage, 12,* 20–27. doi:10.1006/nimg.2000.0593

Luu, P., Tucker, D., Derryberry, D., Reed, M., & Poulsen, C. (2003). Electrophysiological responses to errors and feedback in the process of action regulation. *Psychological Science, 14,* 47–53. doi:10.1111/1467-9280.01417

Martin, L., Potts, G., Burton, P., & Montague, P. (2009). Electrophysiological and hemodynamic responses to reward prediction violation. *NeuroReport, 20,* 1140–1143. doi:10.1097/WNR.0b013e32832f0dca

Masaki, H., Shibahara, Y., Ogawa, K., Yamazaki, K., & Hackley, S. (2010). Action–outcome contingency and feedback-related negativity. *Psychophysiology,* Suppl. 1, S100.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron, 38,* 339–346. doi:10.1016/S0896-6273(03)00154-5

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *Journal of Neuroscience, 27,* 5796–5804.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306,* 503–507.

Miltner, W., Braun, C., & Coles, M. (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,* 788–798. doi:10.1162/jocn.1997.9.6.788

Montague, P., McClure, S., Baldwin, P., Phillips, P., Budygin, E., Stuber, G., . . . Wightman, R. (2004). Dynamic gain control of dopamine delivery in freely moving animals. *Journal of Neuroscience, 24,* 1754–1759. doi:10.1523/JNEUROSCI.4279-03.2004

O'Doherty, J., Dayan, P., Friston, K., Critchley, H., & Dolan, R. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 38,* 329–337. doi:10.1016/S0896-6273(03)00169-7

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304,* 452–454. doi:10.1126/science.1094285

Pan, W. X., Schmidt, R., Wickens, J. R., & Hyland, B. I. (2005). Dopamine cells respond to predicted events during classical conditioning: Evidence for eligibility traces in the reward-learning network. *Journal of Neuroscience, 25,* 6235–6242. doi:10.1523/JNEUROSCI.1478-05.2005

Roesch, M., Calu, D., & Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience, 10,* 1615–1624. doi:10.1038/nn2013

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., . . . Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *NeuroReport, 16,* 407–411.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36,* 241–263. doi:10.1016/S0896-6273(02)00967-4

Schultz, W., Romo, R., Ljungberg, T., Mirenowicz, J., Hollerman, J., & Dickinson, A. (1995). Reward-related signals carried by dopamine neurons. In J. Houk, J. Davis, & D. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 233–248). Cambridge: MIT Press.

Schultz, W., Tremblay, L., & Hollerman, J. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex, 10,* 272–283. doi:10.1093/cercor/10.3.272

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction* (Vol. 116). Cambridge: Cambridge University Press.

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology, 39,* 212–222.

Tricomi, E., Delgado, M., & Fiez, J. (2004). Modulation of caudate activity by action contingency. *Neuron, 41,* 281–292. doi:10.1016/S0896-6273(03)00848-1

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5,* 297–323. doi:10.1007/BF00122574

Walsh, M., & Anderson, J. (2011). Learning from delayed feedback: Neural responses in temporal credit assignment. *Cognitive, Affective, & Behavioral Neuroscience, 11,* 1–13. doi:10.3758/s13415-011-0027-0

Yeung, N., Holroyd, C., & Cohen, J. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex, 15,* 535–544. doi:10.1093/cercor/bhh153

Yeung, N., & Sanfey, A. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience, 24,* 6258–6264. doi:10.1523/JNEUROSCI.4537-03.2004