Brain responses to outcomes of one's own and other's performance in a gambling task

Rongjun Yu[a] and Xiaolin Zhou[a,b,c]

[a]Department of Psychology, Peking University, [b]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University and [c]Learning & Cognition Laboratory, Capital Normal University, Beijing, China

Correspondence and requests for reprints to Dr Xiaolin Zhou, Department of Psychology, Peking University, Beijing 100871, China
Tel: +8610 6275 6599; fax: +8610 6276 1081; e-mail: xzl04@pku.edu.cn

Previous studies find that the feedback-related negativity in brain potentials is sensitive to the negative outcome of one's own performance. The feedback-related negativity is suggested to reflect the learning processes using feedback about the consequences of recent actions. Learning can also take place when one observes other people's actions and the associated outcomes. We recorded brain potentials while the participant received outcome feedback for their own or for another person's performance in a gambling task. The feedback-related negativity or the feedback-related negativity-like effects were obtained in both situations, suggesting that similar neural mechanisms are involved in evaluating the outcomes of one's own and the other's actions. Thus, the neural processes in learning-by-observation resemble those in learning-by-doing. *NeuroReport* 17:1747–1751 © 2006 Lippincott Williams & Wilkins.

Keywords: feedback-related negativity, observational learning, outcome evaluation

Introduction

Humans (and other animals) use external feedback to learn how to behave. Modern psychological approaches of human behavior distinguish two principal sources of learning: personal experience and observation [1]. Experience shapes behavior as a response to a direct stimulus on the individual, while observational learning originates from observing the responses of the others. No matter which source the individual learns from, learning depends crucially on his ability to discriminate between positive outcome, indicating that the behavior is appropriate, and negative outcome, indicating that the behavior is in some way inappropriate. Fast outcome evaluation allows rapid decision-making and the adjustment of future behavior.

Recent studies utilizing event-related potentials (ERPs) to examine the evaluative processes in the brain have found an ERP component, called the feedback-related negativity (FRN), to be differentially sensitive to positive and negative feedback [2]. The FRN is distributed mainly over fronto-central regions of the scalp and it reaches maximum amplitude between 250 and 300 ms following the onset of feedback stimuli associated with unfavorable outcomes, such as incorrect responses or monetary losses. Much work has been done to investigate the FRN responses to the outcomes of one's own actions [2], and a reinforcement learning theory has been proposed to account for a wide range of experimental findings [3].

Little, however, is known about how the brain responds to the outcomes of other people's actions in observational learning, in which these outcomes are not directly relevant to the individual's own behavior or interest. Miltner *et al.* [4] asked the participant to perform a choice reaction time task or to watch a computer simulation of the same task. They found that, compared with correct responses, incorrect responses committed by the participant elicited the error-related negativity (ERN), an ERP component that may have the same underlying neural mechanisms as the FRN [3]. Importantly, an ERN-like effect was also found when the participant observed errors committed by the simulated other individual. Both effects could be accounted for by a pair of bilateral sources in the anterior cingulate cortex in the dipole source localization analysis (see also [5]). These studies, however, are limited by the fact that they examined the ERPs in a situation in which the correctness of the other's behavior was determined by comparing the observed responses with the observer's internalized, expected responses. The participant was informed of the stimuli-response mappings beforehand and there was no need or even possibility for the participant to learn new skills or strategies in observation. Moreover, the participant was asked to count the number of errors in the task, making error observation more salient than observing correct responses. It is not clear to what extent the negativity in the 'observation' condition was affected by such manipulations.

In this study, we attempt to replicate and extend the above studies by directly comparing the ERPs elicited by feedback to one's own performance and with those elicited by the observed feedback to another person's performance. Here, the positive and the negative feedbacks, involving

monetary gain or loss in a gambling task, could be provided in a naturalistic context and be equated in frequency. The participant could, in theory, learn to improve one's own performance by observing the other's behavior. Moreover, because the feedbacks are delivered some time after the responses, the FRN is dissociated from the response generation process. We hypothesize that the classic FRN would be elicited by the negative outcomes of one's own performance. Whether observing the negative outcomes of the other's performance would elicit similar brain responses is an empirical question, a finding that could have important implications for theories of FRN and learning.

The second aim of this study is to examine whether the participant's performance and brain responses are affected by the observation of the other's performance. Gehring and Willoughby [6] asked participants to choose one of two cards, each of which contained the numeral 5 or 25. The selected card then turned (randomly) red or green, indicating that the participant gained or lost the amount of money indicated by the number. Comparing all the loss trials with the gain trials, the authors obtained the typical FRN. Moreover, they replicated a classic finding that the outcome in the preceding trial has a great influence on the riskiness of the individual's behavior in the next trial: the proportion of trials on which the participant chose the risky card with the bigger number was greater when the previous trial was a loss than when it was a gain [7]. Correspondingly, the FRN effect was greater after loss trials than after gain trials. In this study, we adopt Gehring and Willoughby's paradigm, but ask the participant to play the game round-by-round, in alternation, with another person over the computer network. The participant could see the other person's choices of cards and the associated outcomes. The question is whether a similar pattern of sequential adjustment, as reported by Gehring and Willoughby, would be observed for the participant in the present setup.

Method

Participants

Twenty undergraduate students (10 female; mean age 21.2 ± 1.6 years) gave informed consent to participate in the experiment. They were first told that they would get paid 20 yuan (about US$2.5) for their participation, and their performance in the experiment would determine how much they would be awarded or penalized on top of this basic payment. Another female student, a collaborator of the experimenter, played the role of 'the other person'. Participants were healthy, right-handed and had normal or corrected-to-normal vision. The experiment followed the guidelines of the Department of Psychology, Peking University.

Procedures

Participants sat comfortably about 1 m in front of a computer screen in an electrically shielded room. Each participant took part in two conditions: a 'self-execution' condition in which they performed the gambling task, and an 'observation' condition in which they observed the collaborator's performance on the screen. The collaborator sat behind the participant. Each experimental block began with an 'observation' trial, followed by an 'execution' trial, which was followed by another 'observation' trial, and so on.

Each trial began with the presentation of either the collaborator's or the participant's name (4° high, 3° wide, white against a black background) above the fixation sign for 500 ms. Two gray cards (2° high, 6° wide) were then presented on the left and the right sides of the fixation. One card had the number 5 written at the center and the other card had the number 25. The participant or the collaborator pressed one of the buttons on a joystick to select one of the two numbers and their choice was then highlighted by a thickening of the white outlines of the card. After 500 ms, the chosen card turned red or green to indicate whether the participant or the collaborator gained or lost the amount of money represented by the chosen number. To emphasize the valence of the outcome, the '+' or '−' symbol was added before the number to indicate the gain/loss status of the outcome.

Unknown to the participant, the collaborator's choices and the associated feedbacks were predetermined by the computer program, such that the four types of outcomes ($+25$, $+5$, $−5$ and $−25$) had equal frequencies of appearance for the collaborator (i.e. for 'observation'). The gain/loss status of the participant's chosen number was also determined according to a prespecified pseudorandom sequence, with half the times gaining and another half losing. The assignment of the two colors as 'gain' and 'loss' was counterbalanced over participants.

Before the experiment, the participant and the collaborator were told that they would play the game in turn, and that they should pay attention to each other's performance and try to learn from it. They were informed that the value of each chosen outcome would be added to or subtracted from the total amount of bonus money awarded to them, separately, at the end of each block of trials and their goal was to earn as much money as possible. The amount each person won or lost would be cumulated through the experiment. When a block resulted in a net loss, however, this block would be taken as zero and the person concerned would receive no reward or penalty. The experiment consisted of eight blocks of 80 trials each. Each block had 40 execution trials and 40 observation trials. A practice block was administered before the formal test.

Recording and analysis

The electroencephalogram (EEG) was recorded from 64 scalp sites using tin electrodes mounted in an elastic cap (NeuroScan Inc., Herndon, Virginia, USA) according to the international 10/20 system, with the reference on the left mastoid. Eye blinks were monitored with an electrode located below the right eye. The horizontal electro-oculogram (EOG) was recorded from electrodes placed 1.5 cm lateral to the left and right external canthi. All interelectrode impedance was maintained below 5 kΩ. The EEG and EOG were amplified using a 0.05–70 Hz bandpass and continuously sampled at 500 Hz/channel for offline analysis.

Separate EEG epochs of 700 ms (100 ms baseline) were extracted offline for the feedback stimuli. Ocular artifacts were corrected with an eye-movement correction algorithm. All trials in which EEG voltages exceeded a threshold of $\pm 60 \mu V$ during the recording epoch were excluded from further analysis. The EEG data were re-referenced offline to the linked mastoid electrodes by subtracting from each sample of data recorded at each channel one-half of the activity recorded at the right mastoid. The EEG data were
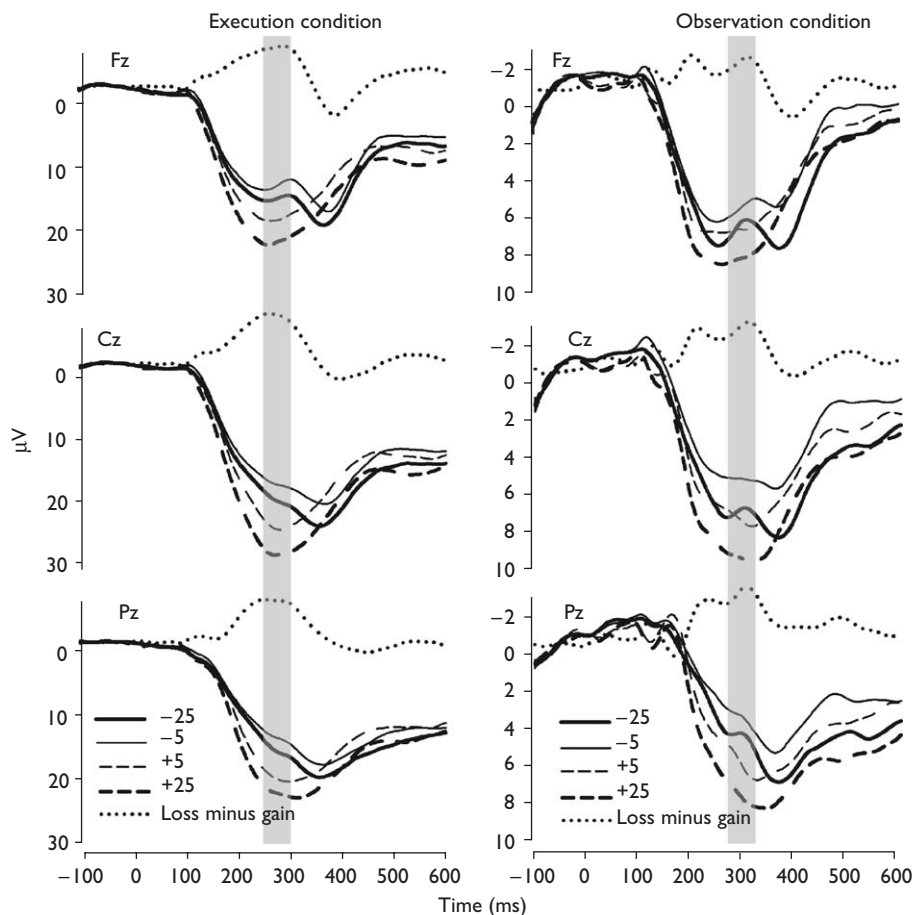
Fig. I Grand-average event-related potential waveforms and the difference waves between loss and gain trials (collapsed over reward magnitudes) from channel Fz, Cz, Pz after the presentation of feedback stimuli in the 'self-execution' condition (the left panel) and the 'observation' condition (the right panel). Feedback stimulus onset occurred at 0 ms. The gray shaded areas indicate the analysis window (253–303 ms for the execution condition and 263–313 ms for the observation condition) in which the feedback negativity was quantified.

low-pass filtered below 20 Hz. The data were baseline corrected by subtracting from each sample the average activity of that channel during the baseline period.

The FRN effect was measured separately for the execution and observation conditions as the differences between the averaged amplitudes of the gain and loss trials. For the purpose of statistical analysis, we calculated the average peak latencies of the difference waves over the five midline electrode locations (Fz, FCz, Cz, CPz and Pz) separately for the execution and observation conditions (collapsed over the 5 and 25 events) and took the average amplitudes ± 25 ms around these latencies as the FRN (Fig. 1). The data from the midline electrodes were reported, because the FRN was shown to be the greatest at these sites in previous studies. In all analyses, the Greenhouse–Geisser correction for nonsphericity was applied where appropriate.

Results
Behavior results
Participants gained on average 27.1 yuan (SD = 9.8) for extra money at the end of the experiment. The best score was a gain of 43 yuan and the worst was a gain of 13 yuan. At debriefing, all participants reported that they did pay attention to the outcomes of the other person's performance.

To examine whether the other's performance affected the participant's choices, we compared the participant's riskiness of choices (i.e. choosing 25 rather than 5) following the four possible outcomes in the observation condition. A 2 (valence: gain, loss) × 2 (magnitude: 25, 5) repeated-measures analysis of variance (ANOVA) found no main effect of the gain/loss status of the previous trial (F < 1) and no main effect of reward magnitude (F < 1). The interaction between valence and magnitude missed the significance [$F(1,19) = 4.25$, $P = 0.053$]. The mean probability of a risky choice after each of the four observed outcomes was as follows: gain 25 (48.8%, SD = 17.1%), gain 5 (50.2%, SD = 12.4%), loss 5 (52.9%, SD = 11.6%), loss 25 (48.4%, SD = 12.3%).

Event-related potential results
The average peak latency of the FRN effect over the five midline electrodes was 278 ms (SD = 42) for the execution condition and 288 ms (SD = 64) for the observation condition. A 2 (condition: execution, observation) × 5 (location: Fz, FCz, Cz, CPz Pz) repeated measures ANOVA on the peak latencies found no significant main effect of condition (F < 1), suggesting the negativities in the two conditions had the same peak latencies.

An ANOVA on the average amplitudes, with valence (gain/loss), reward magnitude (small/large) and electrode location (Fz, FCz, Cz, CPz, Pz) as three within-participant factors, revealed a main effect of valence in the execution condition [$F(1,19) = 33.8$, $P < 0.01$], indicating that the FRN was greater for loss trials than for gain trials. No interaction was found between valence and magnitude ($F < 1$), suggesting that the reward magnitude had no significant impact upon the FRN effect. The interaction between valence and location was significant [$F(4,76) = 6.1$, $P < 0.05$]. Paired t-tests showed that the FRN effect was greater at the fronto-central locations FCz and Cz than at other locations ($P < 0.05$).

Similarly, for the observation condition, the main effect of valence was significant [$F(1,19) = 24.2$, $P < 0.01$], and this effect did not interact with reward magnitude [$F(1,19) = 1.3$, $P > 0.1$] . The interaction between valence and location, however, was significant [$F(4,76) = 10.9$, $P < 0.01$]. Paired t-tests showed that the effect was the greatest at the centro-posterior locations CPz and Pz.

Clearly, the magnitudes of the FRN effects differed significantly between the two conditions (-8.7 vs. $-3.6 \mu V$ on average, see Fig. 1). More importantly, we found no evidence that observing the other person's performance would influence the FRN to one's own performance: a 4 (observed outcomes: $+25$, $+5$, -5, $+25$) $\times 2$ (outcome in the execution condition: gain, loss) $\times 5$ (location: Fz, FCz, Cz, CPz, Pz) ANOVA revealed no significant interaction between the observed outcome and the outcome in the execution condition ($F < 1$), indicating that observing the other's performance did not affect the FRN responses to one's own performance, consistent with the behavioral data.

Discussion

The finding of an FRN effect in the execution condition replicated many previous studies [2]. Importantly, an FRN-like effect, although smaller in magnitude and more posteriorly located than the FRN effect in the execution condition, was also found for the observation condition. The similarity between the two FRN effects suggests that similar neural mechanisms are involved in evaluating the outcomes of one's own actions and the outcomes of the other's actions (also see [4,5]).

According to the reinforcement learning theory, the FRN reflects the arrival in the anterior cingulate cortex of a reward signal conveyed by the mesencephalic dopamine system [3]. These signals are used to guide action selection mediated by the anterior cingulate cortex, through the reinforcement of actions associated with reward and the punishment of actions associated with penalty. The present study, together with some other studies [8–10], however, challenged the assumption that the FRN is only elicited by negative outcomes that are contingent upon recent actions: the FRN can be elicited in simple monetary gambling tasks in which the participant makes no active choices and no overt actions [8–10], or in which the participant merely observes the other performing the task.

One possibility is that the anterior cingulate cortex uses reward signals not only to reinforce representations of one's own actions (i.e. instrumental conditioning), but also to learn about contingencies in the external environment (i.e. observational learning). From this view, observing the actions of others can be considered to be a covert learning of stimuli–response mappings, such that the predictive validity of the mappings tends to improve with experience and observation. Observational learning allows us to learn without actually doing something or suffering from the negative consequences. Thus, the anterior cingulate cortex functions as a general monitoring system that evaluates either one's own actions or the observed actions of others along the good–bad dimension. This general function allows the individual to learn from their own experience and from observation of other people's performance and outcomes. Clearly, on this account, learning-by-observation and learning-by-doing would have similar neural mechanisms.

Nevertheless, it is clear from this study that the two types of learning differ considerably in whether a previous choice and the associated outcome have impact upon the subsequent action and brain responses. Gehring and Willoughby [6] found that the participant was more likely to make a risky choice after loss in the previous trial than after gain. Accordingly, the FRN was greater after loss trials than after gain trials, indicating that the mesencephalic dopamine system is actively adjusting the nature or the level of signals sent to the anterior cingulate cortex in response to contextual changes. In this study, however, we did not find such sequential adjusting between observation of the other's performance and one's own behavior and brain responses. The discrepancies may suggest that the participant maintains clear distinctions between the execution of one's own actions and the observation of another person's actions, even though the brain responds in similar ways to the outcomes of these actions.

One distinction between the self-executed actions and the observed actions lies perhaps in the involvement of emotion and motivation. The reward and penalty for one's own performance would elicit stronger emotional and motivational reactions than the observation of someone else receiving reward or penalty. It has been suggested that the FRN reflects an evaluation of the motivational impact of outcomes and, as such, is associated with feedback signals in general [6,9]. Although the participant may be empathetic with the loss of the other person in the game, the observation of this person's performance does not elicit strong emotional responses in the observer, especially as the observed performance and reward have no direct relationship with the observer's own performance and reward. Consequently, observing someone else losing money not only elicits a smaller FRN effect but also induces no change in the observer's gambling behavior and brain responses. In fact, the present findings can also be interpreted as the brain (FRN) responses to the emotional or motivational consequences of the feedback, given the tie-up between action and its emotional consequence [11].

Conclusion

Observing someone else losing money in a gambling task elicits an FRN-like effect in brain potentials, mirroring the brain responses to the outcomes of one's own performance. These results suggest that similar neural mechanisms are involved in evaluating the outcomes of one's own and the other's actions. Thus, the neural processes in learning-by-observation resemble those in learning-by-doing.

References

1. Armantier O. Does observation influence learning? *Game Econ Behav* 2004; **46**:221–239.

2. Nieuwenhuis S, Holroyd CB, Mol N, Coles MGH. Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neurosci Biobehav Rev* 2004; **28**: 441–448.

3. Holroyd CB, Coles MGH. The neutral basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 2002; **109**:679–709.

4. Miltner WHR, Brauer J, Hecht H, Tippe R, Coles MGH. Parallel brain activity for self-generated and observed errors. In: Ullsperger M, Falkenstein M, editors. Errors, conflicts, and the brain: current opinions on performance monitoring. Leipzig: Max Planck Institute for Cognitive Neuroscience; 2004. pp. 124–129.

5. van Schie HT, Mars RB, Coles MGH, Bekkering H. Modulation of activity in medial frontal and motor cortices during error observation. *Nat Neurosci* 2004; **7**:549–554.

6. Gehring WJ, Willoughby AR. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 2002; **295**:2279–2282.

7. Kahneman D, Tversky A. Prospect theory. *Econometrica* 1979; **47**:263–292.

8. Yeung N, Holroyd CB, Cohen JD. ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cereb Cortex* 2005; **15**:535–544.

9. Donkers FCL, Nieuwenhuis S, Van Boxtel GJM. Mediofrontal negativities to averted gains and losses in the absence of responding. *Cogn Brain Res* 2005; **25**:777–787.

10. Yu R, Zhou X. Brain potentials associated with outcome expectation and outcome evaluation. *Neuroreport* 2006; **17**:1649–1653.

11. Yeung N. Relating cognitive and affective theories of the error-related negativity. In: Ullsperger M, Falkenstein M, editors. Errors, conflicts, and the brain: current opinions on performance monitoring. Leipzig: Max Planck Institute for Cognitive Neuroscience; 2004. pp. 63–70.