Failing where others have succeeded: Medial Frontal Negativity tracks failure in a social context

MAARTEN A. S. BOKSEM, a,b,c EVELIEN KOSTERMANS, c,d AND DAVID DE CREMER a
a Rotterdam School of Management, Erasmus University, Rotterdam, The Netherlands
b Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
c Department of Social Psychology, Tilburg University, Tilburg, The Netherlands
d Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

Abstract

Most of us can appreciate that it feels worse to fail when people around you are successful than when others are also failing. Indeed, comparison with other individuals is of central importance within social groups. Despite the importance of relative success or failure for human decision making and even well-being, the underlying neurobiological substrate of this social comparison process is not well understood. In the present study, ERPs were recorded while two participants received feedback on both their own, and the other participant’s performance on each trial. The results showed that medial frontal negativity, an ERP component associated with deviations from the desired outcome, is particularly enhanced when an individual’s own outcomes are worse than those of others. These results indicate that the way the brain evaluates the success of our actions is crucially dependent on the success or failure of others.

Descriptors: ERN, FRN, Feedback negativity, Social comparison, Social context

Imagine you have applied for this big and important grant, which would be greatly beneficial to your research, and indeed your career. Now as you sit at your desk, you receive an e-mail with the decision letter, stating that they are, of course, very sorry, but you will not receive the money. You have failed. Or, in terms of Reinforcement-Learning (RL) theory, you have experienced a reward-prediction error (Barto & Sutton, 1997). According to RL models, every outcome that one could possibly experience is tagged with a certain motivational value. Outcomes that have a high expected value (such as securing this big grant) are preferred over those with lower expected values (not obtaining the grant). Whenever the value of the observed outcome differs from the preferred value of the outcome, this is coded as a reward-prediction error.

The neural substrate underlying the processing of these reward-prediction errors is relatively well established and involves a network of cortical and subcortical brain areas. Schultz and colleagues have shown that, subcortically, these reward-prediction errors are encoded by midbrain dopamine (DA) neurons of non-human primates (Schultz, Dayan, & Montague, 1997). Cell recordings of these animals’ brains show that activity in DA neurons is increased when outcomes are better than expected (i.e., obtaining a grape instead of the predicted cucumber slice; a positive reward-prediction error), while decreases in activity occur when outcomes are not as good as expected (i.e., expecting a grape, but receiving cucumber; a negative reward-prediction error). The human ventral striatum also shows a strong decrease in activation when an expected reward is not obtained (Pagnoni, Zink, Montague, & Berns, 2002), resembling the prediction error signal in midbrain DA neurons as observed in primates. So, these midbrain DA neurons do not code for reward value per se, but instead signal that observed outcomes are worse than they could have been, be it not receiving a grape when you really craved one, or not securing the grant you wanted so badly.

In addition to these midbrain DA neurons, anterior cingulate cortex (ACC) has been shown to be involved in the processing of errors and reward prediction. Holroyd and Coles (2002) suggested that the prediction-error signal generated in the midbrain is conveyed to the ACC, where it produces an error signal that can be measured as a negative event-related potential (ERP) on the scalp. This prediction-error signal in the human brain is reflected by a family of negative-going ERPs that are elicited both when subjects commit errors (error negativity (Ne; Falkenstein, Hoormann, & Blanke, 1990), also referred to as error-related negativity (ERN)), as well as when subjects receive negative performance feedback (feedback-related negativity (FRN); Miltner, Braun, & Coles, 1997). It has been suggested that these ERP components are associated with common neural processes (Nieuwenhuis, Holroyd, Mol, & Coles, 2004), and, for convenience, we will refer to them as medial frontal negativity (MFN; Gehring & Willoughby, 2002).

It has been proposed that, in addition to signalling errors and negative outcomes, the MFN reflects a motivational/affective
evaluation of such outcomes (Boksem, Tops, Kostermans, & De Cremer, 2008; Gehring & Willoughby, 2002; Hajcak, Moser, Yeung, & Simons, 2005; Pailing & Segalowitz, 2004). More specifically, it has been argued that MFN amplitudes are most dependent on how concerned subjects are over making mistakes, especially in a social context. Indeed, both measures of negative affectivity (i.e., anxiety, neuroticism) and positive affectivity (i.e., agreeableness; Deneve & Cooper, 1998) have been shown to affect MFN amplitude while they also relate to concerns over social evaluation (e.g., Boksem, Ruys, & Aarts, 2011; Tops, Boksem, Wester, Lorist, & Meijman, 2006). The most salient feedback signals are arguably of a social nature, and negative social evaluation is probably one of the most potent ones, leading to strong physiological responses, such as increases in cortisol levels (Dickerson & Kemeny, 2004). Indeed, the MFN, punishment sensitivity, and cortisol responses have all been related to social evaluative threat (Boksem, Tops, Wester, Meijman, & Lorist, 2006; Cavanagh & Allen, 2008; Hajcak et al., 2005). Importantly, the ACC (the putative source of the MFN) has been shown to be involved in processing ‘error’ signals from the social environment such as potential loss of social resources: exclusion, rejection, and the experience of shame and guilt (Eisenberger, Lieberman, & Williams, 2003; Kross, Egner, Ochsner, Hirsch, & Downey, 2007).

Back to the grant you failed to secure: imagine your colleague also applied for this grant. Realizing he also must have received a decision letter, you walk up to his office to find out whether he received the grant or not. Which sound would you prefer to hear coming out of his office: the uncorking of a champagne bottle or a gentle sobbing? This, of course, may depend on how much you like your colleague, as well as on your personality, but the interesting question is why it should matter at all. After all, your outcome, your negative prediction-error in terms of absolute value, is of the same magnitude in both scenarios (you were hoping for funding, but failed to obtain it). Indeed, traditional models of economic decision making typically assume that social comparison, and therefore relative success or failure, plays no role in how we evaluate our outcomes. Nevertheless, the feeling that failing when others succeed is worse than failing when others also fail is something most of us can easily relate to. This is why these traditional models have been challenged by, for example, social psychologists, who have argued that comparison with other individuals is of central importance within social groups (Festinger, 1954). According to Festinger’s theory of social comparison, humans have a natural tendency to evaluate themselves by comparing their abilities, achievements, and possessions to those of others. At the same time, people are very much aware that others also engage in such a comparison process and that the outcome of these comparisons can have major consequences for how one will be treated by others. So, when this comparison turns out negative, as when our outcomes are not as good as those of others, this is perceived as a threat to our self-esteem. Low self-esteem in turn is something to be avoided, as it is a solid predictor of a variety of negative life outcomes such as poor health, criminal behavior, and low socio-economic status (Trzesniewski et al., 2006). Despite the importance of relative success or failure for human decision making and even well-being, the underlying neurobiological substrate of this social-comparison process is not well understood.

In the present study, two participants concurrently but independently performed a simple time-estimation task and received feedback on both their own and their co-participant’s performance on each trial. While subjects were rewarded for every correctly performed trial, these earnings were completely independent from the earnings of the other subject and, when awarded, were of the same magnitude on every trial. Indeed, the other’s performance was completely irrelevant and could have been ignored by the participant. We recorded electroencephalogram data (EEG) from both subjects while they were performing the task. If the outcomes of other individuals are indeed important for how we evaluate our own outcomes, this should be reflected in MFN amplitudes. MFN reflects the motivational value of outcomes or, in other words, the subjective importance of these outcomes. According to social-comparison theory, outcomes that compare negatively to those of others, irrespective of the objective value of the outcome, are most motivationally salient. Thus, particularly enhanced MFN amplitudes should be elicited when outcomes are worse than those of another individual.

Methods

Participants and Procedure

Thirty-six healthy participants (nine males), between 18 and 29 (M = 20.4, SD = 2.7) years of age, were recruited from the university population and were paid for their participation. Written informed consent was obtained prior to the study.

Subjects were invited to the lab in pairs. Upon arrival, they were informed that they were to play a game with the other participant present in the lab and that we would record EEG from both of them. We then proceeded to apply the electrodes to both subjects and seated them in separate dimly lit, sound-attenuated, electrically shielded rooms at 1.20 m from a 17” PC monitor, where they practiced the task for 20 trials. Because all subjects in reality played the game independently (the task was predetermined and the performance of the other player that the subject was ‘interacting’ with was simulated), having the two participants show up at the lab at the same time was only to increase credibility of the social nature of the game.

Task

Subjects performed a simple time-estimation reaction-time task. At the start of each trial, a blue circle was presented, which changed color to green after 2–2.5 s. The participants were required to press a response button exactly 1 s after the circle had turned green. Subjects were rewarded with 3 euro-cents (ct) every time they responded correctly. Responses were considered correct when they were within a certain critical time interval. Two seconds after the circle changed color, subjects were given feedback on their performance: a smiley face accompanied by ‘+3 ct’ when they responded within the critical time interval, or a sad face and ‘+0 ct’ when they responded too fast or too slow. Unbeknownst to the subjects, we covertly adjusted the critical interval based on the subjects’ performance. That is, when subjects responded within the critical interval, we decreased the length of the interval by 5 ms; when they responded too fast or too slow, the interval length was increased by 5 ms. This allowed us to manipulate the percentage of positive and negative feedback that the subjects received, such that every subject received 50% positive feedback. It is important to note that, although the percentage of positive and negative feedback was manipulated, this feedback was actually still contingent upon their performance. What differed between subjects was the time interval within which responses were considered correct.
Subjects performed 520 trials of the time-estimation task, lasting for approximately 40 min. During the experiment, the names of both players were presented on screen at all times. The names presented were the actual names of the two participants (see Figure 1). Importantly, subjects received feedback on the performance of both participants, and were informed that their co-player also received feedback on the performance of both participants. It was stressed, however, that the earnings of both subjects were completely independent, so that the outcome of the other player had no effect on the subject’s own outcome. Subjects were informed that, at the end of the experiment, they would be paid the amount of money that they had earned during the experiment. We encouraged them to perform as best as they could so they would earn a large amount of money. In reality, because all subjects received positive feedback 50% of the time, they all earned about 8 euros.

Electrophysiological Recording and Analyses

EEG recordings were made using 128 active Ag–AgCl electrodes (Biosemi ActiveTwo, Amsterdam, Netherlands) mounted in an elastic cap. Horizontal electro-oculograms (EOGs) were recorded from two electrodes placed at the outer canthi of both eyes. Vertical EOGs were recorded from electrodes on the infra-orbital and supraorbital regions of the right eye placed in line with the pupil. The EEG and EOG signals were sampled at a rate of 256 Hz, digitally low-pass filtered with a 52 Hz cut-off (3 dB) and offline re-referenced to an averaged mastoid reference.

All ERP analyses were performed using the Brain Vision Analyser software (Brain Products GmbH, Gilching, Germany). The data were down-sampled to 100 Hz and further filtered with a 0.53–40 Hz notch-pass filter with a slope of 48 dB/oct. Out-of-range artefacts were rejected and eye-movement artefacts were corrected, using the Gratton, Coles, and Donchin (1983) method.

ERPs from each individual subject and condition were averaged separately, and a baseline voltage averaged over the 200-ms interval preceding feedback was subtracted from these averages.
Visual inspection of grand-averaged waveforms and their scalp distributions (Figures 2 and 3) indicated an MFN that reached its maximum between 220 and 320 ms after presentation of the feedback on midline frontal electrode sites, centred around FCz and Cz. To minimize the effects of overlap between ERP components, most notably the P3, we created difference waves by subtracting ERPs elicited by ‘correct’ feedback from ERPs associated with ‘incorrect’ feedback (Holroyd & Krigolson, 2007). Mean amplitudes of these difference waves recorded from FCz, Cz, and Pz in the two time intervals were used for statistical analyses. In addition, P3 amplitudes reached their maximum between 320 and 420 ms on Pz and were also submitted to analyses to make sure differences in MFN amplitudes were unrelated to P3 amplitudes.

Results

**MFN**

Figures 2 and 3 show that, in accordance with previous studies reporting MFN, negative feedback elicited a negative-going waveform that reached its maximum over frontocentral scalp positions, peaking at FCz. Amplitude data from this electrode site was submitted to a general linear model (GLM) with own outcome and other’s outcome as factors, showing that, as expected, feedback indicating own incorrect performance elicited a larger negativity in the MFN latency range compared to feedback indicating correct performance, $F(1,35) = 82.46, p < .001$. In addition, we found a main effect of other’s outcome: MFN was slightly more negative when the other player experienced a loss, compared to a gain, $F(1,35) = 14.10, p < .005$. Most importantly, these factors were shown to interact, $F(1,35) = 23.11, p < .001$. Follow-up contrasts showed that, after negative feedback, MFN amplitudes were significantly larger when the other player experienced a gain, compared to when the other player experienced a loss, $t(35) = -6.91, p < .001$. After positive feedback, MFN amplitudes did not differ according to the outcome for the other player, $t(35) = 1.29, n.s.$ Similar but smaller effects were observed on Cz and Pz.

To further investigate this interaction effect, we created difference waves, by subtracting the ERPs associated with positive feedback from the ERPs associated with negative feedback. As can be observed in Figure 4, the amplitude of this difference wave...
dependently on the outcome for the other player. A paired samples t-test showed that, on FCz, the amplitude of the difference waves elicited when the other player received positive feedback was significantly larger compared to when the other received negative feedback, t(35) = −8.24, p < .001. Similar results were obtained when we analyzed data from Cz and Pz, t(35) < −4.81, p < .001.

P3
In the P3 latency range, feedback indicating correct performance resulted in a larger positivity (see Figures 2 and 3) compared to feedback indicating incorrect performance on Pz, F(1,35) = 36.97, p < .001, and Cz, F(1,35) = 7.34, p < .05, but not on FCz, F(1,35) = 0.13, n.s. Figures 2 and 3 show that, in accordance with previous studies reporting P3 and different from findings on MFN, this increased positivity reached its maximum over parieto-occipital scalp positions. In addition, negative outcomes experienced by the other player were associated with larger P3 amplitudes on Pz, F(1,35) = 3.58, p < .05; this effect was even more prominent at more frontal electrodes, F(1,35) > 6.56, p < .05. However, the interaction between own and other’s outcome was only significant at Pz, F(1,35) = 18.90, p < .001.

Again, the amplitude of the difference wave depended on the outcome for the other player (Figure 4). A paired samples t-test showed that on Pz the amplitude of the difference waves elicited when the other player received positive feedback was significantly larger compared to when the other received negative feedback, t(35) = −4.35, p < .001. This effect was not present on more frontal sites, t(35) > −1.58, n.s.

Submitting the ERP difference wave data to a GLM with Time (220–320 vs. 320–420 ms) as a within-subject factor showed a significant interaction between Time and ‘Outcome Other’ on FCz, F(1,35) = 41.03, p < .001 and also on Cz, F(1,35) = 32.92, p < .001, while this interaction only approached significance on Pz, F(1,35) = 2.71, p = .11. Together with the difference in topographical distribution between these time intervals, this suggests that amplitudes in these two time intervals may reflect activity of partly overlapping but independent effects on frontal sites while, on Pz, it is more difficult to distinguish activity associated with MFN and P3.

Discussion
In the present study, two participants concurrently but independently performed a simple time-estimation task and received feedback on both their own and the other participant’s performance on each trial. The results showed that MFN amplitudes elicited by feedback indicating failure were larger compared to those elicited by feedback indicating success, mirroring a reward-prediction error signal. Importantly, we found that this neural response associated with failure was significantly enhanced when the other participant experienced a positive outcome, compared to when he or she also failed. It should be stressed here that the outcomes of both subjects were completely independent, so objectively the outcome of the co-player was inconsequential for the outcome of the participant. Nevertheless, the results showed that the neural prediction error signal is modulated by the outcome for others: MFN is particularly enhanced when own outcomes are worse than those of others.

Of course, the idea that we evaluate rewards more by comparison than by their intrinsic value has been around for a long time (e.g., Hume, 1778). Relative reward processing has previously been demonstrated in the striatum (Cromwell, Hassani, & Schultz, 2005) in primates, where neuronal responses to a given reward depend on possible alternative reward outcomes. Similarly, context dependent reward activity in the human striatum has been demonstrated using fMRI (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Nieuwenhuis et al., 2005). In these studies, the neural response to winning or losing nothing (i.e., receiving $0) depended on the distribution from which the outcome was drawn (a set of possible gains or a set of possible losses). A similar pattern of outcome-context dependency has been observed in the ACC, which receives extensive projections from the striatum: using a similar experimental design, Holroyd and colleagues have shown that the amplitude of the MFN—the ERP correlate of ACC error-processing—evoked by a particular outcome also depends on the relative value compared to the possible range of outcomes, rather than on the intrinsic value of the evoking outcome (Holroyd, Larsen, & Cohen, 2004).

In the present study, however, there were only two possible outcomes: success (earning 3 cents) or failure (earning nothing). Therefore, in our study reward expectation on every trial should be the same and independent from the rewards received by the other participant. As such, the differential neural activation in response to the other’s success or failure shows an immediate impact of contextual social information on neural processes involved in evaluation of own outcomes. In this sense, our study shows that the brain also takes the range of possible outcomes for others into account when predicting and evaluating own rewards. In other words, the other’s outcome is perceived as an important reference point by which own outcomes are evaluated. The idea that people evaluate the value of alternative outcomes relative to a subjective reference point is a central idea of prospect theory (Kahneman & Tversky, 1979). Indeed, Loewenstein and colleagues (1989) already suggested that, in an interpersonal context, the outcomes of another person may emerge as an alternative (or additional) potentially salient reference point guiding subsequent decision making.

Loewenstein and colleagues originally considered these ‘social reference points’ in the context of modelling the apparent preference of people to reduce inequality in outcome distributions, which had been disregarded by traditional models of economic decision making (which assumed only a preference for the maximizing of absolute utility; see also Fehr & Schmidt, 1999). Indeed, there is a large body of behavioral evidence consistent with the predictions of these ‘social utility’ theories, both from primate (Brosnan & De Waal, 2003) and human (Fehr & Schmidt, 1999; Loewenstein et al., 1989) studies.

More recently, these models have also been supported by neuroimaging findings. In one paradigm, for example, two participants concurrently play the same game and receive monetary reward for correct responses. When both participants respond correctly, activity in the ventral striatum increases when participants receive a larger monetary reward than their co-players, while activity in the same region decreases if participants receive a smaller monetary reward than their co-players (Fliessbach et al., 2007). In other words, the subjects are evaluated by the same criterion (both participants provide the correct answer), but there is inequality in the payment they receive, activity in the ventral striatum is more closely related to the subject’s payment relative to the other’s payment than to the absolute payment. In addition, the ACC and also the MFN have been shown to respond to unfairness in outcome distributions, such as when subjects receive an unfair offer in an Ultimatum Game (Boksem & De Cremer, 2010; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003).
These studies show that the neural circuitry involved in reward processing and outcome evaluation also code for the (in)equity in the distribution of payoffs between individuals, providing a neural substrate for the phenomenon of inequity aversion predicted by models of social utility and fairness (Fehr & Schmidt, 1999): when two people are equally successful or put in an equal amount of effort, but one person receives a higher payoff than another, this is registered in the striatum and ACC. The present study adds to these findings by showing that it is not only the discrepancy between the invested effort or the result of performance and the rewards received that impacts on these neural processes. Even when there is no asymmetry in payoff, the social context in which failure occurs has an effect on how the prediction error is processed: when failure occurs in the context of another’s success, it is coded as a larger prediction error compared to when this failure occurs in the context of another’s failure.

The experience of envy seems closely related to these findings. Recently, Takahashi and colleagues (2009) showed that, when outcomes for others are superior to outcomes for oneself, people reported intense feelings of envy. Interestingly, the intensity of envy has been shown to be positively correlated to ACC activation: stronger ACC activation was observed when subjects felt more envious. These findings may suggest that the reward prediction error experienced when others are better off than oneself is associated with the emotion we call envy. This emotion is usually elicited by a perceived threat to our self-esteem. Since we are usually motivated to maintain high self-esteem (Tesser & Cornell, 1991), we feel discomfort when our self-esteem is threatened by others who outperform us in a self-relevant domain (see also Takahashi et al., 2009). Indeed, activity in the ACC is enhanced when our positive self-concept conflicts with external information (Amadio et al., 2004). Therefore, the association between ACC activity and envy suggests that ACC activation in this social context is generated when the result of social comparison conflicts with the motivation to maintain high self-esteem. Results by Zink and colleagues (2008) provide some support for this idea. These authors found that only outcomes with hierarchical results by Zink and colleagues (2008) provide some support for this idea. These authors found that only outcomes with hierarchical social context is generated when the result of social comparison is associated with the emotion we call envy. This emotion is usually elicited by a perceived threat to our self-esteem. Since we are usually motivated to maintain high self-esteem (Tesser & Cornell, 1991), we feel discomfort when our self-esteem is threatened by others who outperform us in a self-relevant domain (see also Takahashi et al., 2009). Indeed, activity in the ACC is enhanced when our positive self-concept conflicts with external information (Amadio et al., 2004). Therefore, the association between ACC activity and envy suggests that ACC activation in this social context is generated when the result of social comparison conflicts with the motivation to maintain high self-esteem. Results by Zink and colleagues (2008) provide some support for this idea. These authors found that only outcomes with hierarchical social context is generated when the result of social comparison conflicts with the motivation to maintain high self-esteem. Results by Zink and colleagues (2008) provide some support for this idea. These authors found that only outcomes with hierarchical social context is generated when the result of social comparison conflicts with the motivation to maintain high self-esteem. Results by Zink and colleagues (2008) provide some support for this idea. These authors found that only outcomes with hierarchical

REFERENCES


MFN and failure in a social context


*(Received July 19, 2010; Accepted October 29, 2010)*