Article

Intrinsic Emotional Relevance of Outcomes and Prediction Error

Their Influence on Early Processing of Subsequent Stimulus During Reversal Learning

Louis Nahum,[1,2] Sandra Barcellona-Lehmann,[1,3] Stéphanie Morand,[1,4] David Sander,[2,5] and Armin Schnider[1]

[1]Laboratory of Cognitive Neurorehabilitation, Department of Clinical Neurosciences, University of Geneva, Geneva University Hospitals, Switzerland, [2]Laboratory for the Study of Emotion Elicitation and Expression, Department of Psychology, University of Geneva, Switzerland, [3]Hôpital neurologique, Fondation Institution de Lavigny, Lavigny, Switzerland, [4]Department of Psychology, University of Glasgow, UK, [5]Swiss Centre for Affective Sciences and Department of Psychology, University of Geneva, Switzerland

Abstract. Infrequent events, such as unexpected absence of outcomes (prediction errors), have a detrimental effect on performance of subsequent trial in various cognitive tasks. In the present event-related potential study, we tested whether the influence of prediction error manifests itself in the early cortical processing of subsequent stimuli. Participants performed a reversal learning task in which they saw two alternating pairs of faces and indicated for each pair which one would have a declared target stimulus on its nose. The target switched to the other face after several consecutive trials with correct response, thereby inducing a prediction error, with the switch being indicated by the appearance of a disk (unexpected neutral outcome) or a spider (unexpected unpleasant outcome), depending on the condition. Results showed that after both unexpected and expected unpleasant outcomes, the amplitude of P2 decreased, while after both unexpected neutral and unpleasant outcomes, the amplitude of P1 increased on the following presentation of the pair of faces. Source localization analysis suggested that the differences mainly emanated from the cuneus and precuneus with respect to the P1 and P2 time ranges respectively. We conclude that both the intrinsic emotional relevance of outcomes and prediction error may modulate attention allocation.

Keywords: reversal learning, prediction error, orienting attention, P1, P2

Behavioral accuracy is worse and reaction times are longer after a trial with incorrect response than a trial with correct response in various experimental paradigms (Debener, Makeig, Delorme, & Engel, 2005; Hajcak & Simons, 2008; Holroyd & Krigolson, 2007). This effect might reflect the persistence of the cognitive process that led to error on the previous trial (Gehring, Goss, Coles, Meyer, & Donchin, 1993). However, studies using reversal learning also found slower reaction times after unexpected provoked errors (prediction errors) that signal a change in learning contingencies (Mullette-Gillman & Huettel, 2009; Schnider, Mohr, Morand, & Michel, 2007). Furthermore, trials with correct response also induce reaction times slowing on the subsequent trial when they are less frequent than errors (Notebaert et al., 2009). The "orienting account" hypothesis proposes that post-error slowing is due to a capture of attention by unexpected events (Notebaert et al., 2009).

From that perspective, prediction errors during reversal learning would induce a reaction time slowing because of its lower frequency of occurrence than confirmatory outcomes. The relative unexpectedness of these events would enhance arousal and capture attention. Then a longer delay in disengagement of attention from the unexpected outcome to the subsequent event would be produced in comparison to expected outcomes. Electrophysiologically, prediction errors induce larger amplitudes of two waveform components than presence of expected outcomes: the feedback-related negativity (Bellebaum & Daum, 2008; Holroyd & Krigolson, 2007; Yasuda, Sato, Miyawaki, Kumano, & Kuboki, 2004) and the P300 (Bellebaum & Daum, 2008; Hajcak, Holroyd, Moser, & Simons, 2005). These effects have been associated with a switch of attention to the unexpected change (Escera, Corral, & Yago, 2002; Escera, Yago, Corral, Corbera, & Nunez, 2003). However, the neural

DOI: 10.1027/0269-8803/a000066

effects of prediction errors on the subsequent event during reversal learning remain unknown. If the orienting account hypothesis is correct, then the nonoccurrence of anticipated outcomes which signals a need to switch behavior should influence early cortical processing of subsequent stimuli as well as behavioral performance.

The present study tested this hypothesis by using a reversal learning task (Nahum, Morand, Barcellona-Lehmann, & Schnider, 2009; Nahum, Simon, Sander, Lazeyras, & Schnider, 2011) in which participants had to predict which one of the two faces (cued stimuli) would have a declared target stimulus on the nose. The target unexpectedly switched to the alternate face after several consecutive trials with correct response, thus inducing a prediction error. The intrinsic emotional relevance of the outcomes was manipulated; in the "unexpected neutral outcome" condition, the target outcome was an unpleasant stimulus and prediction error was indicated by a neutral outcome, while in the "unexpected unpleasant outcome" condition, the target outcome was the neutral stimulus and prediction error was indicated by the unpleasant outcome.

In a previous study using this experiment and event-related potentials (ERPs) time-locked to the outcomes (Nahum et al., 2009), we found that the intrinsic emotional relevance of the outcomes influenced early cortical processing of prediction error; unpleasant stimuli signaling prediction error enhanced amplitude of the N170 and induced a specific electrocortical topography between 200 and 300 ms compared to neutral stimuli. In the current study, we used the same recordings as in Nahum et al. (2009) but focused analyses on the cued stimuli in order to study a different question: do prediction errors influence early cortical processing of subsequent stimuli?

According to the orienting account hypothesis, we supposed that prediction errors, indicating a switch of association between a cued stimulus and a target outcome, would influence early cortical processing of subsequent stimuli because the nonoccurrence of an outcome violates a person's expectation. As the P1 and P2 components have been associated with early visuospatial orienting of attention (Luck, Heinze, Mangun, & Hillyard, 1990; Mangun, 1995) and attention disengagement respectively (Bar-Haim, Lamy, & Glickman, 2005), we tested the hypothesis according to which these components upon presentation of the pair of faces would be influenced by prediction error in the preceding trial. Secondly, we tested whether this response would differ according to the intrinsic emotional relevance of the outcome.

Materials and Methods

Participants

The present study constitutes an analysis of the data used in Nahum et al. (2009), except for two participants who were excluded due to ERP artifacts in the period of interest. Data from 13 healthy right-handed students (3 males, 10 females) aged 24.1 years (*SD* = 2.4 years, range: 21–29 years) entered the analysis. The participants had given written informed consent to participate in the study, for which they received financial compensation. The study was approved by the Ethical Committee of the University Hospital of Geneva. All participants declared not to be spiderphobic, which was confirmed by low to moderate scores in a French version of the Fear of Spider Questionnaire (Szymanski & O'Donohue, 1995; *M* ± *SD*, 39.38 ± 20).

Experimental Task

Participants performed two conditions of a reversal learning task (see Nahum et al., 2009), in which they had to learn associations between a cued stimulus and an outcome stimulus that unpredictably reversed. Figure 1 illustrates the design of the experiment. Participants saw two alternating grayscale pairs of neutral male faces (Pair 1 and Pair 2) matched for brightness and size (taken from the Ekman & Friesen, 1975) with randomly changing right or left position. For each pair, participants had to indicate which of the two faces would have a target outcome located on its nose by pressing with the index or middle finger of the right hand the button corresponding to the side of the chosen face. The target outcome switched randomly between faces and independently for each pair of faces (1 or 2) after four to six consecutive trials with correct response (i.e., switches could occur on a different number of consecutive correct responses according to the pair of faces). Absence of the target outcome was indicated by appearance of an alternative outcome and signaled the need to switch behavior. In one condition, the declared target outcome was a black schematic spider and the alternative outcome was a black disk. In this case, the unexpected absence of the spider is referred to as "unexpected neutral outcome." In the second condition, the declared target outcome was a disk and the alternative outcome was a spider. In this case, absence of the disk is referred to as "unexpected unpleasant outcome."

Participants were informed that the target would normally reappear on the same face, but that it would occasionally switch to the other face. They were asked to make their choice on the basis of the previous outcome and to restrain from guessing.

Trials presenting the same outcome that had been correctly predicted in the previous trial were called "expected outcome trials." Trials in which the target was absent from the face, on which it had been presented in the previous trial (absence indicated by presentation of the alternate outcome), were called "prediction error" trials (Pe trials). Trials following unexpected outcomes (Pe trials) were called "unexpected outcome trials."

To assess the influence of prediction error on the cortical processing of the subsequent stimuli, we compared cued stimuli following absence of expected outcomes (unexpected outcome trials) with cued stimuli following presence of expected outcomes (expected outcome trials). The use of two alternating and independent pairs of faces allowed us to analyze cued stimuli for which participants were not

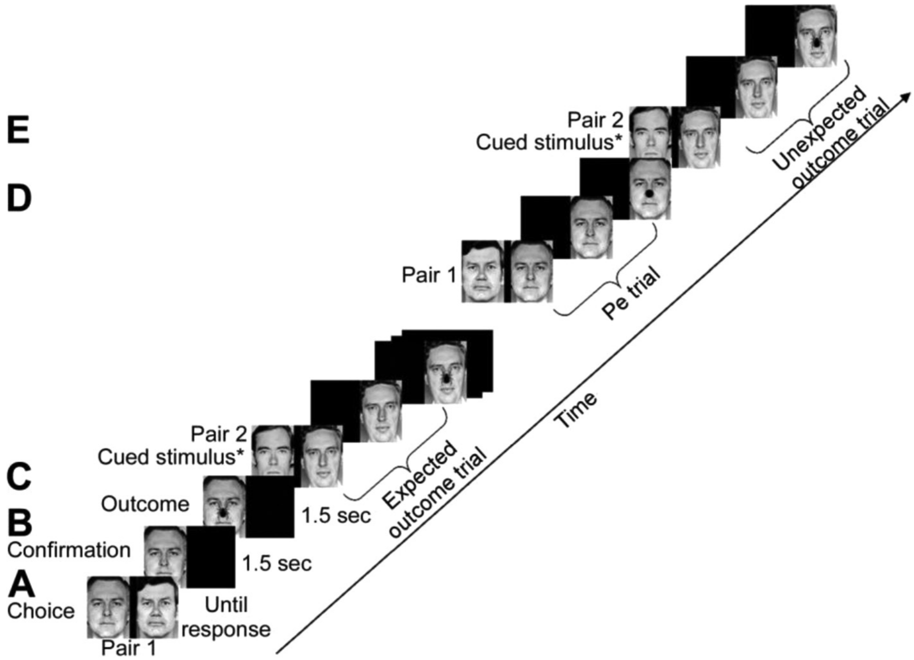

Figure 1. Sequential order of trials in the course of the reversal learning task demonstrating the two different outcome types when the target was the unpleasant outcome. Every trial consisted of the same three steps: (A) Presentation of the first pair of faces (cued stimulus, Pair 1); participants had to predict by button press which one of the two faces would have a spider on its nose. (B) After the choice, a cross appeared on the respective face for 1.5 s (confirmation). (C) Then the outcome was presented on the nose of the chosen face for 1.5 s: A spider (target outcome present) if the choice was correct or a disk (nontarget outcome), indicating absence of the target outcome. After 0.7 s, the second pair of faces was presented and participants had also to predict which one of the two faces would have a spider on its nose. (D) For each pair, after randomly four to six consecutive trials with correct response (expected outcome trials), a disk indicating absence of the spider was presented (prediction error, Pe trials). (E) After the Pe trials (unexpected outcome trials), the second pair of faces was presented (cued stimulus, Pair 2). Note that incorrect choice put back the counter of consecutive trials with correct response to zero. *Events of interest.

required to switch behavior in either condition, in order to match for cognitive control and attentional demand. The number of expected outcome trials was sufficiently variable for participants not to anticipate target switches: Questioning after the experiment revealed that 12 of 13 participants were unaware that Pe trials would always appear after four to six expected outcome trials; only one participant thought that he had learned the pattern of the trial succession but indicated a wrong estimation (6–8).

At the beginning of each block, the target to be looked for was announced (either the spider or the disk) and 20 practice trials using the respective target were performed. Two blocks had the spider, two blocks the disk as the target. The order of blocks was randomized. Participants performed a total of 1,440 trials. There were approximately 60% expected outcome trials, 20% Pe trials, and 20% unexpected outcome trials.

Prior to the experiment, the participants rated the emotional relevance of the two target outcomes that were used in the experiment. They answered six questions for each target outcome assessing current anxiety, disgust, and distress induced by these stimuli on a Likert scale from 0 (= not at all) to 9 (= extreme).

Analysis of Behavioral Data

Repeated-measures ANOVAs were computed on reaction times and accuracy with the emotional relevance of the previous outcome (neutral or unpleasant) and the predictability of the previous outcome (unexpected outcome trials, expected outcome trials) as within-participants repeated-measures factors.

EEG Acquisition and Preprocessing

The electroencephalogram (EEG) was recorded continuously using the Active-Two Biosemi EEG system (Biosemi V.O.F., Amsterdam, The Netherlands) with 128 channels covering the entire scalp. Signals were sampled at 512 Hz in a bandwidth filter of 0–134 Hz. All analyses were conducted using Cartool Software (Brain Mapping Laboratory, Geneva, Switzerland, http://brainmapping.unige.ch/Cartool.htm). Unlike our previous study (Nahum et al., 2009) in which analyses were time-locked to outcomes, epochs corresponding to −200 ms before the onset of the cued stimuli (appearance of the pair of faces) to 300 ms after stimulus onset were averaged along four conditions: expected outcome trials with the spider as the previous outcome (expected unpleasant outcome), expected outcome trials with the disk as the previous outcome (expected neutral outcome), unexpected outcome trials with the spider as the previous outcome (unexpected unpleasant outcome), and unexpected outcome trials with the disk as the previous outcome (unexpected neutral outcome). A baseline correction was applied to the 200 ms pre-stimulus period.

All trials were visually inspected and epochs with an artifact higher than ±50 μV or artifacts due to blinks, eye movements, or other sources were rejected from averaging. Dysfunctioning electrodes ($M \pm SD$, 18.5 ± 4.8) were interpolated using spherical splines (Perrin, Pernier, Bertrand, Giard, & Echallier, 1987). Before group averaging, individual data were recalculated against the average reference and bandpass filtered to 1–30 Hz. The number of trials used for analyses was 61.3 ± 11 ($M \pm SD$) for expected outcome trials and 68.6 ± 16.7 for unexpected outcome trials. Only trials with correct response were analyzed.

Analysis of ERPs

Waveform Analysis

First, traditional ERP analyses were performed (Picton et al., 2000). To test the hypothesis that prediction errors influence early cortical processing, we calculated the mean voltage (amplitude) of P1 and P2 components, time-locked to the cued stimulus (appearance of the pair of faces) in expected outcome trials and unexpected outcome trials. Based on previous studies and inspection of the grand mean ERPs, P1 amplitude was measured from 80 to 140 ms (Clark & Hillyard, 1996) over lateral occipital sites that had shown related effects in previous studies (PO3, PO4, PO7, and PO8 of the International 10–20 System; Brosch, Sander, Pourtois, & Scherer, 2008; Clark & Hillyard, 1996; Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005; Pourtois, Grandjean, Sander, & Vuilleumier, 2004) and P2 amplitude was measured from 180 to 270 ms (Eldar & Bar-Haim, 2010; Kanske, Plitschka, & Kotz, 2011) over medial parietal central and frontal electrode sites (Pz, Cz, Fz; Bar-Haim et al., 2005; Carretie, Mercado, Tapia, & Hinojosa, 2001; Kanske et al., 2011).

Amplitude of P1 was analyzed using a $4 \times 2 \times 2$ ANOVA with the repeated factors of electrode (PO3, PO4, PO7, PO8), predictability (expected outcome trials vs. unexpected outcome trials) and intrinsic emotional relevance of the previous outcome stimulus (unpleasant vs. neutral). Amplitude of P2 was analyzed using a $3 \times 2 \times 2$ ANOVA with the repeated factors of electrode (Pz, Cz, Fz), predictability (expected outcome trials, unexpected outcome trials) and intrinsic emotional relevance of the previous outcome stimulus (unpleasant vs. neutral).

Source Localization Analysis

The intracranial current distribution for each subject's ERP at each time point was estimated by applying a distributed linear inverse solution based on a Local Auto-Regressive Average (LAURA) model using a 3D realistic head model with a solution space of 3,005 nodes (Grave de Peralta Menendez, Murray, Michel, Martuzzi, & Gonzalez Andino, 2004; Menendez, Andino, Lantz, Michel, & Landis, 2001). Current distribution was calculated within the gray matter of the average brain provided by the Montreal Neurological Institute (Montreal, Canada).

To compare statistically the sources between conditions on the two periods of interest (P1: 80–140 ms; P2: 180–270 ms), we computed the contrasts of local electrical current densities with time-point wise paired t-tests similar to statistical parametric mapping (SPM) used in fMRI studies. Only periods for which $p < .05$ for at least 40 consecutive milliseconds were retained (James, Britz, Vuilleumier, Hauert, & Michel, 2008).

Results

Behavioral Results

As expected, paired t-tests demonstrated that the unpleasant outcome was perceived as negative; it induced more anxiety ($M \pm SD$ = 1.9 ± 2.7), disgust (2.5 ± 2.7), and distress (1 ± 1.7) than the neutral outcome (all $p < .05$).

A repeated-measures ANOVA on the mean error rate, with the intrinsic emotional relevance (unpleasant; neutral) and predictability of the previous outcome stimulus (unexpected outcome trials; expected outcome trials) as within-subjects factors, revealed a significant main effect of the intrinsic emotional relevance of the previous outcome stimulus ($F(1, 12) = 7.3$; $p = .02$). Figure 2A shows that participants committed significantly more errors after unexpected ($M \pm SD$ = 6.1 ± 1.4) and expected (4.4 ± 0.8) unpleasant outcomes than after unexpected (3.1 ± 0.7) and expected (2.6 ± 0.4) neutral outcomes.

In contrast to error rates, a repeated-measures ANOVA on the mean reaction times revealed a significant main effect of the trial type ($F(1, 12) = 16.4$; $p = .001$). Figure 2B indicates that reaction times were slower after unexpected neutral (1,112 ms ± 113) and unpleasant outcomes ($M \pm SD$ = 1,224 ms ± 74) than after expected neutral (888 ms ± 80) and unpleasant outcomes (885 ms ± 70).

ERPs Results

Waveform Analysis

Early processing was assessed by means of P1 and P2 in response to cued stimuli for the predictability (expected outcome vs. unexpected outcome trials) and for the intrinsic emotional relevance of the previous outcome stimulus (unpleasant vs. neutral). Figure 3A shows ERPs at posterior lateral electrodes (PO3 and PO4) for expected outcome trials and unexpected outcome trials when the previous outcome was neutral and when it was unpleasant. Analyses revealed

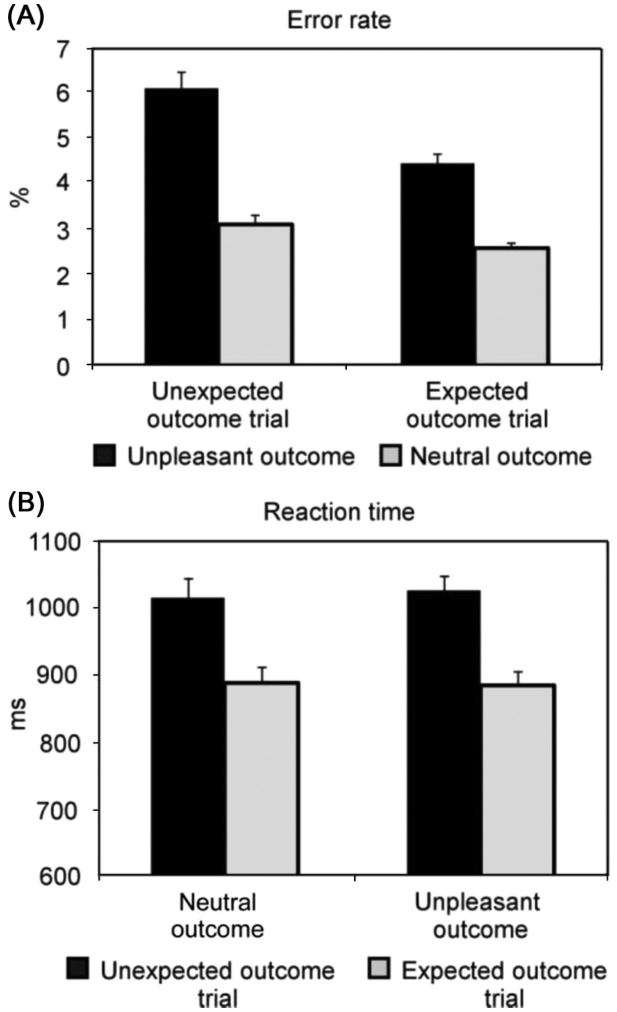

Figure 2. Performance in the reversal learning task. (A) Mean error rate and (B) reaction time to cued stimuli preceded by unexpected (unexpected outcome trials) and expected (expected outcome trials) neutral and unpleasant outcomes. Bars indicate the standard error of mean.

a significant main effect of outcome predictability, indicating that P1 amplitude was larger after absence of expected outcomes than after presence of expected outcomes ($F(1, 12) = 20.4$; $p = .0007$). Neither the main effect of the intrinsic emotional relevance of the previous outcome, nor the interaction effects (Electrode × Predictability × Intrinsic Emotional Relevance) were statistically significant. The topography associated to P1 was quite similar across conditions (Figure 3A).

Figure 3B shows ERPs at medial electrodes (Cz and Fz) for the unpleasant stimulus and the neutral stimulus when the previous outcome was expected and unexpected. Analyses revealed a significant main effect of the intrinsic emotional relevance of the previous outcomes indicating that P2 amplitude was smaller after unpleasant outcomes than after neutral outcomes ($F(1, 12) = 6.9$; $p = .02$). Neither the main effect of the predictability, nor the interaction effects (Electrode × Predictability × Intrinsic Emotional Relevance) were statistically significant. The topography associated to P2 was quite similar across conditions (Figure 3B).

Source Localization Analysis

In order to estimate the localization of the neural generators of the differences observed at 80–140 ms and 180–270 ms in the waveform analysis, paired *t*-tests on the current source density (as determined using LAURA) were performed. Figure 4 indicates that the cuneus was more strongly activated by the stimulus following both unexpected neutral (Figure 4A) and unpleasant outcomes (Figure 4B) than after expected outcomes between 80 and 140 ms. In addition, the medial frontal and cingulate gyri were more strongly activated after unexpected than after expected unpleasant outcomes (Figure 4B). The left precuneus was more strongly activated by the stimulus following unpleasant expected outcomes than neutral expected outcomes between 180 and 270 ms (Figure 4C). Figure 4 shows that the right precuneus was more strongly activated by the stimulus following unexpected unpleasant outcomes than neutral unexpected outcomes (Figure 4D, area in purple), while the right middle frontal gyrus was more strongly activated by the stimulus following neutral unexpected outcomes than unpleasant unexpected outcomes between 180 and 270 ms (Figure 4D, area in red).

Discussion

The major aim of this ERP study was to test whether prediction errors, which signal a change with respect to an anticipated outcome, do influence early cortical processing of subsequent stimuli. Secondly, we tested whether such a response differs according to the intrinsic emotional relevance of the previous unexpected outcome (unpleasant vs. neutral).

Consistent with previous studies, unexpected outcomes had a negative impact on performance, inducing a reaction time slowing on the subsequent trial (Castellar, Kühn, Fias, & Notebaert, 2010; Notebaert et al., 2009; Schnider et al., 2007). ERP results indicated that this effect had a neural correlate: both unexpected unpleasant and neutral outcomes induced a larger amplitude of the early visual P1 component on the subsequent trial than expected unpleasant and neutral stimuli. Source estimations indicated stronger activation of the cuneus after prediction error trials than after confirmatory trials.

It is unlikely that the early neural effects reflect differences of visual features, perceptual load or task difficulty between the conditions. Indeed, short-term memory load and stimuli were equivalent between unexpected and expected outcome trials. Neither, spatial location or motor preparation can account for this effect since side of apparition of the correct face and button to press were randomized. In agreement with the orienting account hypothesis (Notebaert et al.,

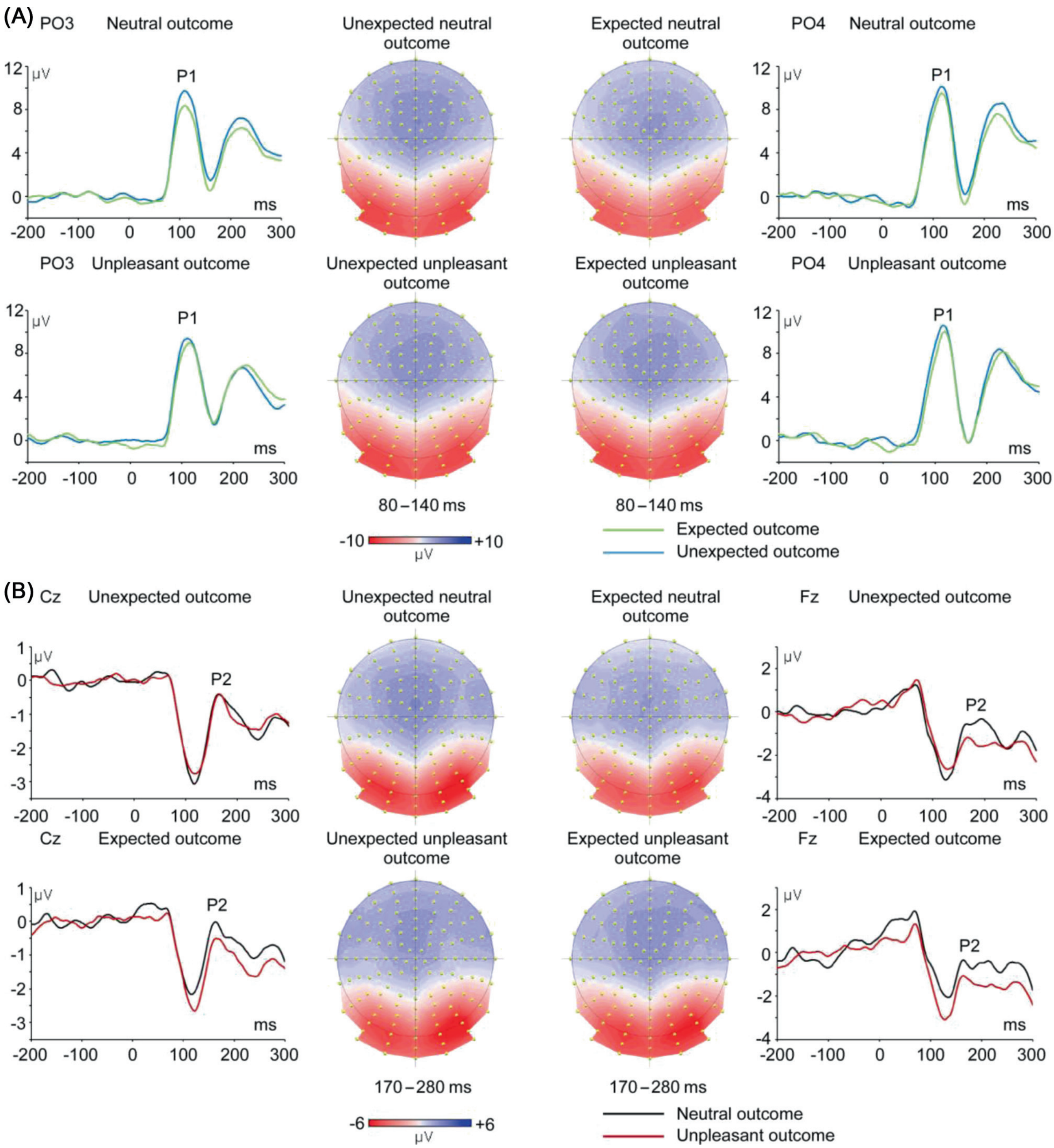

Figure 3. (A) ERP wave forms at electrodes PO3 and PO4 for cued stimuli following expected (green lines) and unexpected (blue lines) neutral or unpleasant outcomes. Maps show the scalp topographies of the P1 component (80–140 ms) for each trial type. (B) ERP wave forms at electrodes Cz and Fz for cued stimuli following neutral (black lines) and unpleasant (red lines) unexpected or expected outcomes. Maps show scalp topographies of the P2 component (180–270 ms) for each trial type.

2009), we rather interpret our findings as evidence for the fact that the prediction errors capture attention; the early modulation of the amplitude of the electrocortical signal might be the marker of attentional capture by prediction errors that causes the observed decrease of performance on the subsequent trial. Result of the source localization is compatible with such an interpretation, as the cuneus has been associated with attentional control processes (Adler et al., 2001; Hopfinger, Buonocore, & Mangun, 2000; Sander et al., 2005).

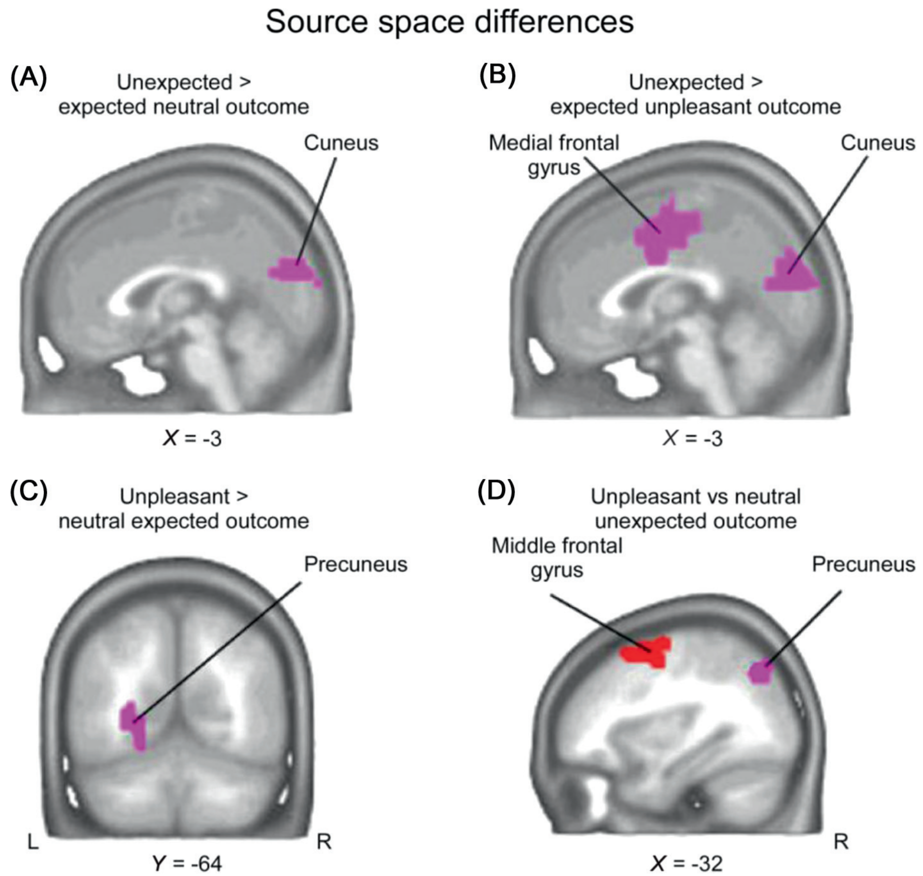

Figure 4. Areas of brain activation. Areas having significantly different current densities as determined from group-averaged source estimations during the time periods 80–140 ms (AB) and 180–270 ms (CD) for the following contrasts: (A) cued stimuli following unexpected neutral outcome > cued stimuli following expected neutral outcome; (B) cued stimuli following unexpected unpleasant outcome > cued stimuli following expected unpleasant outcome; (C) cued stimuli following expected unpleasant outcome > cued stimuli following neutral expected outcome; (D) cued stimuli following unexpected unpleasant outcome > cued stimuli following neutral unexpected outcome (area in purple); cued stimuli following unexpected neutral outcome > cued stimuli following unpleasant unexpected outcome (area in red). Solution points with significant statistical differences are depicted in red or purple on sagittal and coronal slices of the brain template of the Montreal Neurological Institute. Coordinates *x* and *y* are given in Talairach space.

The reason for the attentional bias toward unexpected events is probably linked to the importance of detecting environmental changes, which are potentially rewarding or threatening events. Appraisal theories of emotion suggest that a "predictability appraisal" determines the extent to which an event is expected (Ellsworth & Scherer, 2003; Sander, Grandjean, & Scherer, 2005). "Predictability appraisal" refers to the notion that organisms expect the appearance of particular events as based on past observations of regularities and probabilities for specific events. We thus propose that unexpected events, irrespective of their intrinsic emotional relevance, are thought to affect subsequent reaction times and early cortical processing because they are potentially highly relevant events that require further cognitive processing, which in turn may delay disengagement of attention. This interpretation is in line with studies indicating that prediction errors induced different electrocortical responses compared to presence of anticipated outcomes, independently of a rewarding context (Schnider, 2008; Schnider et al., 2007) or of the intrinsic emotional relevance of the outcome (Nahum et al., 2009).

The intrinsic emotional relevance of the unexpected outcome also had an influence on the subsequent stimulus. Indeed, participants committed more errors, and the amplitude of the P2 component was smaller after an unpleasant than a neutral outcome, irrespective of its predictability. Source estimations indicated stronger activation of the precuneus after unpleasant than neutral outcomes. In addition, unexpected unpleasant outcomes induced stronger activation of the medial frontal and cingulate gyri, but lower activation of the middle frontal gyrus than unexpected neutral outcomes.

Altogether the results are in concordance with those of previous studies showing that emotional stimuli modulate attention allocation (Brosch, Grandjean, Sander, & Scherer, 2009; Brosch et al., 2008; Pourtois et al., 2004, 2005). More specifically, they are compatible with those of previous studies showing that threat-related stimuli can bias attention allocation (Vuilleumier, 2005) and that P2 is associated with emotional evaluation (Carretie, Martin-Loeches, Hinojosa, & Mercado, 2001; Carretie, Mercado, et al., 2001), attentional orienting to emotional stimuli (Kanske et al., 2011), and attention disengagement (Bar-Haim et al., 2005). The precuneus, the cingulate anterior cortex, and the medial frontal gyrus have also been associated with emotional processing (Brázdil et al., 2009; Vuilleumier & Driver, 2007). In addition, the precuneus is involved in the allocation of attentional resources (Barber & Carter, 2005; Pourtois, 2011). According to appraisal theories of emotion, a "pleasantness appraisal" evaluates on the basis of genetically fixed schemata or overlearned associations, the extent to which a stimulus is likely to result in pleasure or pain (Ellsworth & Scherer, 2003; Sander et al., 2005; Scherer, 2001). The outcome of this appraisal determines the orienting of attention and resource processing; unpleasant and pleasant stimuli increase the speed of the orientating response, the amount of allocated resource, and the disengagement of attention compared to neutral stimuli.

In conclusion, this study indicates that both prediction errors indicating a switch of association between a stimulus and a target outcome and the intrinsic emotional relevance of outcomes modulate performance and early cortical processing of subsequent stimuli during reversal learning. The results support the predictions of appraisal theories of emotion that the intrinsic emotional relevance of outcomes and prediction errors modulates attention allocation.

Acknowledgments

The study was supported by Swiss National Science Foundation Grant No. 320030-132447 to Armin Schnider. The Cartool software was developed by Denis Brunet, with the support of the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne.

References

Adler, C. M., Sax, K. W., Holland, S. K., Schmithorst, V., Rosenberg, L., & Strakowski, S. M. (2001). Changes in neuronal activation with increasing attention demand in healthy volunteers: An fMRI study. *Synapse, 42*, 266–272.

Bar-Haim, Y., Lamy, D., & Glickman, S. (2005). Attentional bias in anxiety: A behavioral and ERP study. *Brain and Cognition, 59*, 11–22.

Barber, A. D., & Carter, C. S. (2005). Cognitive control involved in overcoming prepotent response tendencies and switching between tasks. *Cerebral Cortex, 15*, 899–912.

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience, 27*, 1823–1835.

Brázdil, M., Roman, R., Urbánek, T., Chládek, J., Špok, D., Mareček, R., ... Rektor, I. (2009). Neural correlates of affective picture processing – a depth ERP study. *NeuroImage, 47*, 376–383.

Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2009). Cross-modal emotional attention: Emotional voices modulate early stages of visual processing. *Journal of Cognitive Neuroscience, 21*, 1670–1679.

Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008). Beyond fear: Rapid spatial orienting toward positive emotional stimuli. *Psychological Science, 19*, 362–370.

Carretie, L., Martin-Loeches, M., Hinojosa, J. A., & Mercado, F. (2001). Emotion and attention interaction studied through event-related potentials. *Journal of Cognitive Neuroscience, 13*, 1109–1128.

Carretie, L., Mercado, F., Tapia, M., & Hinojosa, J. A. (2001). Emotion, attention, and the 'negativity bias', studied through event-related potentials. *International Journal Psychophysiology, 41*, 75–85.

Castellar, E. N., Kühn, S., Fias, W., & Notebaert, W. (2010). Outcome expectancy and not accuracy determines posterror slowing: ERP support. *Cognitive, Affective, & Behavioral Neuroscience, 10*, 270–278.

Clark, V. P., & Hillyard, S. A. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience, 8*, 387–402.

Debener, S., Makeig, S., Delorme, A., & Engel, A. K. (2005). What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. *Brain Research. Cognitive Brain Research, 22*, 309–321.

Ekman, P., & Friesen, W. (1975). *Unmasking the human face: A guide to recognizing emotions from facial expressions*. Englewood Cliffs, NJ: Prentice-Hall.

Eldar, S., & Bar-Haim, Y. (2010). Neural plasticity in response to attention training in anxiety. *Psychological Medicine, 40*, 667–677.

Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. In R. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 572–595). New York, NY: Oxford University Press.

Escera, C., Corral, M. J., & Yago, E. (2002). An electrophysiological and behavioral investigation of involuntary attention towards auditory frequency, duration and intensity changes. *Brain Research. Cognitive Brain Research, 14*, 325–332.

Escera, C., Yago, E., Corral, M. J., Corbera, S., & Nunez, M. I. (2003). Attention capture by auditory significant stimuli: Semantic analysis follows attention switching. *European Journal of Neuroscience, 18*, 2408–2412.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science, 4*, 385–390.

Grave de Peralta Menendez, R., Murray, M. M., Michel, C. M., Martuzzi, R., & Gonzalez Andino, S. L. (2004). Electrical neuroimaging based on biophysical constraints. *NeuroImage, 21*, 527–539.

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology, 42*, 161–170.

Hajcak, G., & Simons, R. F. (2008). Oops! I did it again: An ERP and behavioral study of double-errors. *Brain and Cognition, 68*, 15–21.

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology, 44*, 913–917.

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience, 3*, 284–291.

James, C. E., Britz, J., Vuilleumier, P., Hauert, C. A., & Michel, C. M. (2008). Early neuronal responses in right limbic structures mediate harmony incongruity processing in musical experts. *NeuroImage, 42*, 1597–1608.

Kanske, P., Plitschka, J., & Kotz, S. A. (2011). Attentional orienting towards emotion: P2 and N400 ERP effects. *Neuropsychologia, 49*, 3121–3129.

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalography and Clinical Neurophysiology, 75*, 528–542.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32*, 4–18.

Menendez, R., Andino, S., Lantz, G., Michel, C., & Landis, T. (2001). Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topography, 14*, 131–137.

Mullette-Gillman, O. A., & Huettel, S. A. (2009). Neural substrates of contingency learning and executive control: Dissociating physical, valuative, and behavioral changes. *Frontiers in Human Neuroscience, 3*, 23.

Nahum, L., Morand, S., Barcellona-Lehmann, S., & Schnider, A. (2009). Instinctive modulation of cognitive behavior: A human evoked potential study. *Human Brain Mapping, 30*, 2120–2131.

Nahum, L., Simon, S., Sander, D., Lazeyras, F., & Schnider, A. (2011). Neural response to the behaviorally relevant absence of anticipated outcomes and the presentation of potentially harmful stimuli: A human fMRI study. *Cortex, 47*, 191–201.

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition, 111*, 275–279.

Perrin, F., Pernier, J., Bertrand, O., Giard, M. H., & Echallier, J. F. (1987). Mapping of scalp potentials by surface spline interpolation. *Electroencephalography and Clinical Neurophysiology, 66*, 75–81.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R. Jr., . . . Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37*, 127–152.

Pourtois, G. (2011). Early error detection predicted by reduced pre-response control process: An ERP topographic mapping study. *Brain Topography, 23*, 403–422.

Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: Time course and topographic evoked-potentials mapping. *Human Brain Mapping, 26*, 65–79.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting toward fearful faces. *Cerebral Cortex, 14*, 619–633.

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., . . . Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *NeuroImage, 28*, 848–858.

Sander, D., Grandjean, D., & Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Networks, 18*, 317–352.

Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92–120). New York, NY: Oxford University Press.

Schnider, A. (2008). *The confabulating mind: How the brain creates reality.* New York, NY: Oxford University Press.

Schnider, A., Mohr, C., Morand, S., & Michel, C. M. (2007). Early cortical response to behaviorally relevant absence of anticipated outcomes: A human event-related potential study. *NeuroImage, 35*, 1348–1355.

Szymanski, J., & O'Donohue, W. (1995). Fear of Spiders Questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry, 26*, 31–34.

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences, 9*, 585–594.

Vuilleumier, P., & Driver, J. (2007). Modulation of visual processing by attention and emotion: Windows on causal interactions between human brain regions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*, 837–855.

Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., & Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *NeuroReport, 15*, 2561–2565.

Accepted for publication: December 26, 2011

Louis Nahum

Service de Neurorééducation
Hôpitaux Universitaires de Genève
26, av. de Beau-Séjour
1211 Geneva 14
Switzerland
Tel. +41 22 382-3527
Fax +41 22 372-3705
E-mail louis.nahum@hcuge.ch