available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cortex

Special issue: Original article

Switching between colors and shapes on the basis of positive and negative feedback: An fMRI and EEG study on feedback-based learning

Kiki Zanolie[a,b,1], *Santani Teng*[d,1], *Sarah E. Donohue*[e], *Anna C.K. van Duijvenvoorde*[a,b], *Guido P.H. Band*[a,b], *Serge A.R.B. Rombouts*[a,b,c] and *Eveline A. Crone*[a,b,*]

[a]Leiden University Institute for Psychological Research (LU-IPR), Leiden University, Leiden, The Netherlands
[b]Leiden Institute for Brain and Cognition (LIBC), Leiden, The Netherlands
[c]Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands
[d]University of California, Davis, CA, USA
[e]Duke University, Durham, NC, USA

ARTICLE INFO

Article history:
Received 19 March 2007
Reviewed 10 July 2007
Revised 30 August 2007
Accepted 17 November 2007
Published online 23 December 2007

Keywords:
Performance feedback
Hypothesis testing
Medial prefrontal cortex
Dorsolateral prefrontal cortex
fMRI
Intradimensional/extradimensional rule shift task

ABSTRACT

A crucial element of testing hypotheses about rules for behavior is the use of performance feedback. In this study, we used fMRI and EEG to test the role of medial prefrontal cortex (PFC) and dorsolateral (DL) PFC in hypothesis testing using a modified intradimensional/extradimensional rule shift task. Eighteen adults were asked to infer rules about color or shape on the basis of positive and negative feedback in sets of two trials. Half of the trials involved color-to-color or shape-to-shape trials (intradimensional switches; ID) and the other half involved color-to-shape or shape-to-color trials (extradimensional switches; ED). Participants performed the task in separate fMRI and EEG sessions. ED trials were associated with reduced accuracy relative to ID trials. In addition, accuracy was reduced and response latencies increased following negative relative to positive feedback. Negative feedback resulted in increased activation in medial PFC and DLPFC, but more so for ED than ID shifts. Reduced accuracy following negative feedback correlated with increased activation in DLPFC, and increased response latencies following negative feedback correlated with increased activation in medial PFC. Additionally, around 250 msec following negative performance feedback participants showed a feedback-related negative scalp potential, but this potential did not differ between ID and ED shifts. These results indicate that both medial PFC and DLPFC signal the need for performance adjustment, and both regions are sensitive to the increased demands of set shifting in hypothesis testing.

* *Corresponding author*. Department of Psychology, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands.
E-mail address: ecrone@fsw.leidenuniv.nl (E.A. Crone).

[1] Shared first-authorship.

0010-9452/$ – see front matter
doi:10.1016/j.cortex.2007.11.005

1. Introduction

In order to adjust our behavior to changing circumstances in daily life, we must monitor the outcomes of our own actions. This type of performance monitoring is especially necessary when discriminating which behavior is appropriate and should be continued, and which behavior is inappropriate and should be adjusted. Therefore, we learn on the basis of *positive and negative feedback*. For example, feedback monitoring is important when we need to adjust pre-specified rules for behavior (Miltner et al., 1997) or when we need to test hypotheses about which behavior is currently appropriate (Barcelo and Knight, 2002).

Research into event-related potentials has demonstrated a differential neural response associated with receiving positive and negative performance feedback. For example, Miltner et al. (1997) asked participants to estimate a 1-sec time interval, which was followed by positive or negative feedback based on the accuracy of the estimate. They demonstrated that negative feedback was followed by a negative brain potential, which was observed approximately 230–270 msec following the presentation of negative feedback. This potential is maximal at frontocentral locations and has similarities with the error-related negativity (ERN), observed after a response error (Falkenstein et al., 1991). Several other studies have reported this brain potential in association with the presentation of negative feedback (Hajcak et al., 2006; Holroyd et al., 2006), which is thought to reflect an evaluation process that monitors expected and unexpected events (Nieuwenhuis et al., 2004). Therefore, this potential may reflect general performance monitoring and has been referred to as the feedback-ERN.

Source localization studies have suggested that the feedback-ERN originates in the medial prefrontal cortex (medial PFC), in or near the anterior cingulate cortex (ACC), and it has been suggested that the feedback-ERN reflects a dopaminergic learning signal (Holroyd and Coles, 2002; Miltner et al., 1997). Subsequent functional magnetic resonance imaging (fMRI) evidence, however, remains inconclusive about the role of the medial PFC in feedback-driven learning (Nieuwenhuis et al., 2005). Although some fMRI studies have shown increased activation in the caudal ACC following negative relative to positive feedback (Holroyd et al., 2004; Mars et al., 2005), others have failed to replicate this effect (Nieuwenhuis et al., 2005; van Veen et al., 2004). Thus, it is currently unclear how medial PFC, including ACC, is involved in the processing of positive and negative feedback when testing hypotheses.

Neuropsychological studies have emphasized the role of lateral PFC (lat-PFC) in feedback-driven learning, as demonstrated by perseverative behavior following feedback-induced rule switching in patients with damage to lat-PFC (Barcelo and Knight, 2002) and deviant ERN responses following damage to lat-PFC (Gehring and Knight, 2000). Kerns et al. (2004) argued that medial PFC signals response conflict and predicts activation in lat-PFC and associated performance adjustment. Therefore, in learning to adjust performance, medial PFC may be sensitive to the general information that signals that performance should be adjusted (negative *vs* positive feedback), whereas lat-PFC may be sensitive to the need to implement goal-directed and controlled behavior (Miller and Cohen, 2001). The goal of this study is to test the relative contributions of these brain regions to the processing of valence of performance feedback and the need for control.

One way to manipulate the demand for cognitive control is by the use of intradimensional (ID) versus extradimensional (ED) rule switches (O'Reilly et al., 2002). The classic ID/ED task (Dias et al., 1997; Roberts et al., 1988) is based on the Wisconsin Card Sorting Task (WCST) in that it demands a categorization of rules according to pre-specified rules. The task involves two kinds of switches: ID, in which the switch represents a category of the same dimension (e.g., switch from color to another color); and ED, in which a switch is made to a different dimension in the context of stimuli sharing the same general set of features (e.g., switch from color-to-shape). An ID switch therefore involves changing the target stimulus to another within the same dimension category, and an ED switch involves changing the target stimulus to one with a different dimension.

Dias et al. (1997) demonstrated that dorsolateral frontal lesions selectively impaired ED switches. This finding is consistent with prior studies which have suggested that lat-PFC is important for the transformation of higher levels of task rules into action, as in abstract rule switching (Cools et al., 2004). Hampshire and Owen (2006) reported that ventrolateral (VL) PFC, rather than DLPFC, played a central role in ED shifting. This result is consistent with prior reports which have suggested that VLPFC is important for the inhibition of the previously relevant response (Robbins, 2007). In contrast, they showed that the DLPFC is generally involved in solution search, a process in which the participant is actively searching the correct target response. This finding concurs with previous models of DLPFC function in which a role for the DLPFC in functions such as monitoring within working memory (e.g., Petrides, 2000) was suggested. In this study we were specifically interested in the feedback processing after having to switch to a new rule dimension (i.e., an ED condition) relative to the same rule dimension (i.e., ID condition). Based on these previous findings (Cools et al., 2004; Dias et al., 1997; Hampshire and Owen, 2006), we predict that lat-PFC will be especially sensitive to feedback-based learning following a switch to ED rules relative to ID rules.

In this study, we examined the role of medial PFC and lat-PFC in relation to feedback-based rule learning using a rule learning task that was inspired by the ID/ED switch task. The same participants were asked to participate in two sessions; fMRI data were assessed on the first occasion and EEG recordings on the second. In both sessions, a predictable switch task was performed which consisted of pairs of trials: a switch trial followed by a repetition trial. Prior to each trial pair, participants were cued to sort two nameable images on the basis of color or shape (see Fig. 1). The response was followed by a positive feedback signal (+) or a negative feedback signal (x), and participants were instructed to use this information to make the correct choice on the repetition trial. As the task required an attentional switch rather than a switch of response mappings, each trial pair presented two new stimuli in two different colors to control for the possible confound of response interference. ED conditions are more likely to require greater monitoring because it is necessary to overcome interference from the previous dimension. However, this modification made the task different from the original ID/ED switch task, because