Think differently: a brain orienting response to task novelty

Francisco Barceló, José A. Periáñez and Robert T. Knight

Department of Psychology, University of the Balearic Islands, Ctra. Valldemossa, km 7.5, 07071 Palma de Mallorca; 1Department of Basic Psychology II, Complutense University, 28223 Madrid, Spain; 2Department of Psychology and the Helen Will Neuroscience Institute, University of California, Berkeley 94720, and the Veterans Administration Research Service, 150 Muir Road, Martinez, CA 94553, USA

CA Corresponding Author: f.barcelo@uib.es

Received 24 July 2002; accepted 22 August 2002

Cognitive flexibility hinges on a readiness to direct attention to novel events, and on an ability to change one’s mental set to find new solutions for old problems. Human event-related potential (ERP) studies have described a brain ‘orienting’ response to discrete novel events, marked by a frontally distributed positive potential peaking 300–400 ms post-stimulus (P3a). This brain potential has been typically related to bottom-up processing of novel non-targets under a fixed task-set (i.e., press a button to coloured targets), but had never been related to top-down attention control in dual-task paradigms. In this study, 27 subjects had their ERPs measured while they performed a version of the Wisconsin card sorting test (WCST), a dual-task paradigm where the same feedback cues signalled unpredictable shifts to a new task set (i.e., from ‘sort by colour’ to ‘sort by shape’). Feedback cues that directed a shift in the subject’s mental set to a new task-set elicited frontally distributed P3a activity, thus suggesting a role of the P3a response system in task-set shifting. Feedback cues also evoked a longer latency positive potential (350–600 ms; P3b), that was larger the more task rules were held in memory. In line with current models of prefrontal function in the executive control of attention, this P3a/P3b response system appears to reflect the co-ordinated action of prefrontal and posterior association cortices during the switching and updating of task sets in working memory. NeuroReport 13:1887–1892 © 2002 Lippincott Williams & Wilkins.

Key words: Cognitive set; Event-related potentials; Orienting response; P300; Prefrontal cortex; Task-switching; Wisconsin card sorting test

INTRODUCTION

Adaptive behaviour requires an executive control system that allows us to respond flexibly to changing environmental events or contingencies [1,2]. These may be an unexpected novel stimulus [3,4], or the need to find a new solution for a familiar problem (i.e. a command to adopt a new plan of action). The ability to alter behaviour on the basis of changing environmental contingencies requires shifting attention among learned stimulus–response associations, or task-sets [5–7]. For example, if your mobile rings while you are driving, you may not pick it up as you normally would, but instead connect the hands-free kit [1]. The brain mechanisms of this executive control have been explored using task set shifting paradigms [8]. Functional imaging and lesion studies in humans suggest a role of prefrontal cortex (PFC) in set shifting, as a key part of a distributed network also encompassing posterior association cortices [9–12]. The temporal dynamics of activation within this frontal-posterior attention network are still poorly defined.

Scalp-recorded event-related potentials (ERPs) provide a fine spatio-temporal analysis of brain activation in so-called oddball target detection tasks. In these tasks, target events evoke a distinct long latency positive potential (350–600 ms; P3b) maximal over mid-parietal scalp. The P3b potential reflects task-relevant processes such as context updating [13] or closure of the event-encoding cycle in working memory [14,15]. An earlier latency positive potential (300–400 ms; P3a) indexes attention switching to non-target novel events [16,17]. Lesion, brain imaging and intracranial studies propose anatomical and functionally distinct neural sources for the switching (P3a) and updating (P3b) mechanisms [15,16,18,19]. For instance, the PFC plays a key role in triggering the P3a potential [20,21], but has a less critical involvement in P3b elicitation [22]. In spite of evidence about an implication of PFC in both task set-shifting and the detection of novel events [8,9,11,20], to date there is no functional evidence that these two processes may be subserved by a common brain mechanism. One limitation of most ERP studies on attention is that brain activity is measured under fixed task set conditions (i.e. ‘press a button to named targets’). Further, even if the task’s rules change between successive trial blocks, ERPs are normally not recorded while the task’s rules are being changed. Thus, the critical brain potentials related to the shifting and updating of stimulus–response mappings (or task sets) in working memory still remain to be described.

Here we used a task-switching paradigm inspired by a classic test of prefrontal impairment, the Wisconsin card sorting test (WCST) [1,8], to study the fast brain dynamics of activation in a frontal-posterior attention network during both the feedback and card-matching stages of WCST.
performance. Recent ERP studies had shown a gradual modulation of P3b amplitudes across ‘shift’ and non-shift (‘stay’) trials time-locked to the card events, but no evidence of a frontally distributed P3a potential [23,24]. However, behavioural research on task set switching suggests that the actual shift in set may take place when a ‘shift’ feedback cue instructs the subject to flexibly adopt a new rule for the task at hand, whereas at the card-matching stage the task-set is merely implemented or rehearsed [5–7]. Thus, the main aim of this study was to examine the likely involvement of the frontal P3a and posterior P3b components of the P300 response system in switching (P3a) and updating (P3b) of task sets in working memory during the feedback stage of WCST performance, in contrast to their alleged role during the card-matching stage [24].

MATERIALS AND METHODS

Subjects: Twenty-seven right handed subjects (15 females; mean age 23 ± 4 years, range 18–34 years), took part in the study. They had normal or corrected visual acuity and no history of neurological or psychiatric disorder. The experiments were performed in accordance with the Declaration of Helsinki, and informed consent was obtained from all subjects.

Behavioural procedures: We used a computer adaptation of the WCST designed to assess attentional set shifting using ERP recordings [24]. The task protocol used the 24 choice-cards of the original 64 WCST cards that can be matched unambiguously with the four WCST key-cards based on just one stimulus dimension (i.e., either colour, shape, or number of items in the card; Fig. 1a). Unambiguous cards are required for both a sensitive scoring of WCST errors and set-shifting ability [25]. These 24 choice-cards were repeatedly used in 137 trials arranged into 18 series (see Fig. 1b).

The correct sorting rule was initially unknown to the subject, and changed randomly from one series to the next. The length of each series varied randomly between six and eight trials, so that subjects could not predict the start of a new series. Each trial began with the onset of a compound stimulus with the four WCST key-cards on top of one choice-card, all centred on a computer screen (Fig. 1a). The cards subtended a visual angle of 4° horizontally and 3.5° vertically, and remained on display until a response was given.

Subjects were instructed to match the choice-card with one of the four key-cards following one of three possible rules: number, colour, or shape of items in the choice-card. Likewise, subjects were informed that the correct sorting rule would change without notice after a variable number of card sorts, and hence, they would have to shift their sorting rule accordingly. The correct rule was to be determined on the basis of an auditory feedback cue delivered within a variable interval of 1500–2000 ms after the response (200 ms duration, 10 ms rise/fall times; 65 dB SPL; 1000 Hz for ‘stay’ cues, 500 Hz for ‘shift’ cues). Subjects used their thumbs for responding using a panel with four key-buttons aligned. The far left button designated the key-card on the far left of the display, the far right button designated the key-card on the far right, and so on. A fixed inter-trial interval of 1400 ms was employed. Subjects performed two blocks of 137 trials each, with a 5 min rest period between blocks. The average duration of each task block was 15 min. The task was practised for about 5 min, or 5–7 series, to make sure that subjects had understood the instructions, and could sort cards efficiently (see operational criteria for efficiently completed WCST series below [25]). The sequence of trials used for practice was different from that in the main task. Therefore, in our adapted WCST protocol each correct card match was followed by a ‘stay’ feedback cue prompting the subject to use the same sorting rule again. After a variable number of correct card matches, the rule changed unpredictably and the subject had to adopt a new task rule upon hearing a ‘shift’ feedback cue (e.g., from shape to number, or to colour). In summary, for our fully instructed, practised, and efficient subjects, a ‘shift’ cue was a signal to think differently and find a new answer for the same card sorting problem, whereas a ‘stay’ cue was a signal to give the same answer just used before.

Following prior fMRI studies [10,11], we defined 3D shift trials (shift3D) as those where subjects had to handle three task rules in working memory (i.e. inhibit the previous rule and consider the other two for responding; Fig. 1b). In 2D shift trials (shift2D), only two rules were handled, after having discarded one in the previous trial. In stay trials, a ‘stay’ feedback cue prompted the subject to use the same task rule again. This task design allowed us to carry out a separate analysis of attention switching (P3a) and memory
updating (P3b) processes. After the first shift feedback cue of a new series (shift3D), an ideal subject has 50% chance of choosing an incorrect task rule, and so needs to shift set again to achieve the remaining correct rule. This is a very efficient trial-and-error process in normal subjects, who can use past contextual information to optimize task set shifting [25]. Here we considered data from efficient series only, with either no errors or just one such efficient error (shift2D; Fig 1b). A detailed analysis of behavioural performance for the present subjects, as part of a larger sample, has been reported elsewhere [25].

**ERPs and data analysis:** The electroencephalogram (EEG) was recorded from 29 tin electrodes positioned at Fp1, Fp2, AF3, AF4, F7, F8, F3, Fz, F4, FC5, FC6, FC1, FC2, T7, T8, C3, Cz, C4, P7, P8, P3, Pz, P4, PO7, PO8, PO1, PO2, O1 and O2, and referenced to the left mastoid. The EEG signal was amplified (band pass, 0.01-30Hz; 12dB/octave roll/off), digitised at 250Hz/channel and stored for off-line averaging. Electrode impedances were kept below 5kΩ. The averaging window was 1600ms for the feedback epoch, and 1400ms for the card matching epoch, including a 200 ms baseline in both cases (Fig. 1a). The electrooculogram (EOG) was also recorded for eye blink correction. Epoch trials with EEG >75μV in amplitude, muscle, or any other artifacts were discarded.

ERP averages were obtained from completed WCST series only. A completed series was scored if (a) the new sorting rule was not anticipated (i.e. the first trial in a series was a shift3D trial); (b) the subject shifted set efficiently and found the new rule in the second (type A series) or third trial (type B series, Fig. 1b); and (c) the rule was not missed thereafter [25]. In the present sample of normal subjects, individual task performance ranged between 32 and 36 successfully completed WCST series [25]. ERPs were computed time-locked to both feedback cues and card onsets across both shift and stay trials. A linked-mastoid reference was obtained off-line.

Mean ERP amplitudes were measured relative to the 200ms pre-stimulus baseline. For feedback-locked averages, mean amplitude values were computed for the P3a component (375-400ms post-stimulus onset) and P3b (550-600ms post-stimulus onset). For card-locked averages, only mean P3a amplitudes were measured (550-600ms post-stimulus onset). The P3a potential was measured at Fz, and the P3b potential at Pz. Measured P3a and P3b amplitudes were normalized for testing hypotheses about their scalp distribution [26]. Mean P3a values and reaction times were subjected to a main ANOVA design with trial (shift3D, shift 2D, stay1, stay 2, stay 3, stay last) as the repeated measures factor. Mean P3b values were subjected to a main ANOVA design with stage (feedback vs card matching) and trial (shift3D, shift 2D, stay1, stay 2, stay 3, stay last) as repeated-measures factors. All post-hoc tests of simple effects were performed using the Bonferroni correction with a significance level of $p < 0.05$.

**RESULTS**

Feedback cues signalling a shift to a new rule evoked a distinct frontally distributed P3a potential that was absent after the first stay cue ($F(5,130) = 20.4, p < 0.0001$, $GG = 0.55$, for the main trial effect; Figs. 2, Fig. 3a). A sharp reduction in P3a amplitude was observed in response to the first ($F(1,26) = 10.3, p < 0.003$), and second stay cues ($F(1,26) = 45.5, p < 0.0001$; Fig. 3a), but there was no decrement in P3a amplitude from shift3D to shift2D trials ($F(1,26) = 1.5, p > 0.2$; Fig. 3a). In turn, shift2D cues evoked larger P3a potentials than stay1 cues ($F(1,26) = 10.4, p < 0.003$; Fig. 3a). Stay2 and later trials evoked similar brain responses in both type A and B series (Fig. 1b) [24]. Finally, P3a amplitudes to shift3D cues did not diminish over trial blocks ($F(1,26) < 1$), consistent with behavioural evidence that set shifting costs do not decline with practice [6].

Feedback cues also elicited a distinct pattern of P3b activation across shift and stay trials ($F(5,130) = 15.9, p < 0.0001$, $GG = 0.61$, for the main trial effect; Fig. 2, Fig. 3a). There was a reduction in P3b amplitude from shift3D to shift2D cues ($F(1,26) = 10.3, p < 0.003$), and between stay1 and stay2 cues ($F(1,26) = 59.8, p < 0.001$), but no P3b change was observed between shift2D and stay1 cues ($F(1,26) < 1$; Fig. 3a). In turn, stay1 cues evoked larger P3b potentials when they were unpredictable (in type A series), than when they were predictable events (in type B series; $F(1,26) = 9.1, p < 0.006$; see insert in Fig. 3a). Thus, unlike the P3a, the P3b response to feedback cues was sensitive both to the number of rules held in memory, and to
the subject's ability to predict the next task rule. Thus, although P3a and P3b components have never been compared in a similar task-switching paradigm, the present results reveal a significant interaction between type of P300 component (P3a vs P3b) and early task shift trials (shift3D, shift2D, stay1), suggesting their differential role in switching (P3a) and updating (P3b) of task sets in working memory (F(1,26) = 7.2, p < 0.03, for the quadratic trend; Fig. 3a).

At the card matching stage, we observed the expected P3b response to visual targets, with a gradual P3b increment from shift to stay trials as revealed in the main trial effect (F(5,130) = 13.1, p < 0.0001; GG = 0.72; Fig. 2, Fig. 3a), but no evidence of a P3a potential [23,24]. However, the pattern of P3b responses at the card matching stage differs substantially from that observed at the feedback stage, as revealed by a significant interaction between stage and trial (F(5,130) = 44.9, p < 0.0001; GG = 0.62; Fig. 3a). These results suggest that the posterior P3b response system accomplishes rather different functions during the feedback and card matching stages of WCST performance [7,11].

The analysis of behavioural responses confirmed the well established costs in speed and accuracy related to task set shifting [6,7]. Response times were delayed during shift compared with stay trials (up to ~500 ms; F(5,26) = 15.26, p < 0.0001; GG = 0.37; Fig. 3b), indicating a gradual speed-up in responding from shift3D to shift2D trials (F(1,26) = 13.4; p < 0.02), and from shift2D to stay1 trials (F(1,26) = 5.8; p < 0.03). The analysis of errors from failed series indicated that subjects were more likely to miss the task rule in shift2D (p < 0.001) and stay1 trials (p < 0.01), compared with the last trial in the series (F(4,104) = 11.3, p < 0.001, GG = 0.62, for the main trial effect; Fig. 3b). Mean P3a amplitudes predicted these speed costs across shift and stay trials (R² = 0.58, F(1,5) = 6.8; p < 0.05). Such a linear relationship was only marginal for the P3b potential at the feedback (F(1,5) = 4.4; p < 0.1) and card matching stages (F(1,5) = 5.6; p < 0.07).

DISCUSSION

This study represents the first evidence in the literature suggesting a role of the frontally distributed P3a response in the executive control of cognitive set shifting. Past ERP studies of attention set shifting either focused on target events only [24], or failed to link the observed P3a-like activation to task set switching [27]. In our modified version of the WCST, feedback events that directed a shift in the subject's mental set to new task rules, also elicited P3a responses whose amplitude, latency and scalp topography closely resemble those elicited by non-target novel events in oddball tasks [15,16]. However, our 'shift' feedback tone cannot be defined as a novel stimulus, as it had been learned to denote a shift in the task's rules, and the same tone was used along the practice and test sessions. Instead, the 'shift' tone prompted the subject to think differently, and to flexibly adopt a new solution (i.e., a new task set) for the same card sorting problem. In consequence, the present results indicate that a common brain response system may be responsible for processing both stimulus and task novelty.

While 'shift' trials were infrequent relative to 'stay' trials (i.e. overall probabilities were 0.25 and 0.75, respectively), oddball processes like uncertainty alone cannot explain the observed modulations of P3a brain potentials. First, oddball tasks with a fixed task set and equally infrequent non-target tones elicit substantially smaller P3a potentials that decline rapidly with repetition [16,28]. Second, task uncertainty cannot account for the functional dissociation of P3a responses to feedback and card events [24]. Third, 'shift' trials from similar tasks evoke peak fMRI activation at PFC regardless of their relative frequency of occurrence [10,11]. That said, brain responses to the first stay feedback cue revealed that task uncertainty did play a role in the early trials of each new WCST series. Indeed, our feedback cues did not make it explicit which task rule was to be used next. This ambiguity disrupts WCST performance in prefrontal patients, who cannot rely on internal representations to project future actions based on past stimulus-response contingencies [1,8]. Even normal subjects who can anticipate the next set need to practice it at least once before reaching pre-shift levels of behavioural efficiency [6]. Indeed, the brain responses observed to the first stay cue may reflect residual reorienting and updating to the newly established task set (Fig. 3a) [5,7]. Activation in this P3a response system ceased completely at the second stay cue, after the new set
had been rehearsed once. Rostral anterior cingulate (BA 24/32), mid-dorsolateral (BA 9/46), and mid-ventrolateral prefrontal cortices (BA 12/47) became simultaneously active in response to feedback cues in similar set-shifting tasks [11,12]. In turn, the extant lesion, fMRI and intracranial recording data favours a lateral − rather than a mediol − prefrontal source for the P3a potential [15,19].

The present results may help us resolve apparent inconsistencies in past brain imaging and clinical WCST research. First, an outdated model of prefrontal function, and the inherently limited temporal resolution of most metabolic brain imaging studies, had led us to expect maximal prefrontal ERP activation during the card matching stage rather than the feedback stage of WCST performance [9,10,24]. In contrast, the observed P3a activation suggests that the shift in set actually takes place during the feedback period, and well before the next target card is on display. This new account is consistent with more recent fMRI [11,12] and behavioural evidence [7], suggesting that the internal representation of task rules must be activated (i.e. updated) in anticipation of the behaviour they govern (i.e. card sorting) [2]. Second, our ERP results provide support for the view that PFC acts in concert with posterior association cortices for the executive control of cognitive set shifting [1]. Figure 2 reveals instant widely distributed neural activation across both frontal and posterior brain generators in response to 'shift' feedback cues. This argues against the strict localizationist view conveyed by some studies that present isolated foci of prefrontal activation during WCST performance (see Fig. 5 in [10]). Third, an ERP index of set shifting may help us integrate apparent inconsistencies in the anatomy reported by different metabolic studies. The stereotaxic coordinates of prefrontal regions with significant fMRI/PET activation during WCST performance show a good deal of variability across studies [9–11]. One possibility is that different anatomical fMRI activation elicited by the same task reflects disparate cognitive processes. Another possibility is that these different fMRI patterns may be showing 'tip-of-iceberg' activation from different parts of the same neural network that gives rise to the P3a response [29]. The present ERP results support the latter alternative. Finally, the finding of feedback-locked frontally distributed P3a activity helps to complete the picture offered by our previous ERP studies [23,24], suggesting that efficient WCST performance demands the activation of a widespread network of brain areas, with a key role played by prefrontal cortex. Further research will be necessary to map specific types of WCST deficits to specific anomalies in the frontal and posterior aspects of the P300 components reported here [4,8,25,30,31].

As in previous studies, a steady build-up in P3b activation to card onset was apparent as the new task set became established and gradually rehearsed [23,24]. This was paralleled by a steady improvement in response speed and efficiency, indicating a growing degree of automaticity in task performance [1,5]. Unlike P3b responses to feedback cues, P3b activity during card matching was modulated neither by the number of task-sets in memory, nor by their predictability [24]. This suggests a differential role of the posterior association cortices responsible for P3b elicitation during the updating (feedback) and rehearsal (card matching) of task rules in working memory [11,12]. It could be argued that long-term memory networks at posterior association cortex need to be differentially engaged both for the rapid retrieval of new task rules during set-shifting and updating, as well as for the gradual rehearsal and consolidation of practised task sets, leading to proficient task performance [32]. Such a proposal could help us to integrate apparently contradictory accounts of the functional role of the P3b response in terms of either context updating [13] or perceptual closure processes [14,18]. Future ERP research with task-switching paradigms should explore further the double dissociation of P3b responses reported here, as well as its likely implication for current models of P300 function.

CONCLUSION

Earlier studies had unsuccessfully searched for an ERP index of prefrontal activation time-locked to the card matching stage of WCST performance [23,33]. Here we show frontally-distributed P3a activity time-locked to discrete feedback events that directed a shift in the subject’s mental set to new task rules. This evidence suggests that the PFC becomes maximally engaged in controlling set-shifting operations, and well before the target WCST card is on display. Besides, this suggests a role of the P3a response in the executive control of attention, and reveals that the brain P3a orienting response may be a more general mechanism of cognitive flexibility involved in the processing of both stimulus and task novelty. Traditionally, the P3a and P3b potentials have been portrayed as separate brain indexes of cognition based on evidence from so-called oddball paradigms with a fixed-task set [16,28]. In turn, in our task-set switching paradigm, these topographically and functionally distinct P300 components may be best regarded as a tightly integrated P3a/P3b response system involved in cognitive set-shifting. The frontal P3a aspect of this response system appears to work in concert with the posterior P3b aspect to activate the task rules stored in long-term memory and bring them on-line in working memory in preparation for the next card sort [1,2].

REFERENCES

Acknowledgements: Supported by grants from Fundación Complutense del Amo y Comunidad de Madrid (to F.B.) and a NINDS grant 21135 (to R.T.K.). We thank Pedro Montoya and an anonymous reviewer for their valuable comments on a previous version of the manuscript.