



ELSEVIER

Neuropsychologia 45 (2007) 2285–2293

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Electrophysiological correlates of common-onset visual masking

Eleni Kotsoni, Gergely Csibra, Denis Mareschal*, Mark H. Johnson

*Centre for Brain and Cognitive Development, School of Psychology, Birkbeck, University of London,
Malet Street, London WC1E 7HX, United Kingdom*

Received 14 November 2005; received in revised form 16 February 2007; accepted 18 February 2007
Available online 4 March 2007

Abstract

In common-onset visual masking (COVM) the target and the mask come into view simultaneously. Masking occurs when the mask remains on the screen for longer after deletion of the target. Enns and Di Lollo [Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, 4(9), 345–352] have argued that this type of masking can be explained by re-entrant visual processing. In the present studies we used high-density event-related brain potentials (HD-ERP) to obtain neural evidence for re-entrant processing in COVM. In two experiments the participants' task was to indicate the presence or absence of a vertical bar situated at the lower part of a ring highlighted by the mask. The only difference between the experiments was the duration of the target: 13 and 40 ms for the first and second experiment respectively. Behavioral results were consistent between experiments: COVM was stronger as a joint function of a large set size and longer trailing mask duration. Electrophysiological data from both studies revealed modulation of a posterior P2 component around 220 ms post-stimulus onset associated with masking. Further, in the critical experimental condition we revealed a significant relation between the amplitude of the P2 and behavioural response accuracy. We hypothesize that this re-activation of early visual areas reflects re-entrant feedback from higher to lower visual areas, providing converging evidence for re-entrance as an explanation for COVM.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: ERP; Visual masking; Re-entrant visual processing

1. Introduction

Several studies have shown that backward projections directly and continuously affect visual processing. For example, Hupe et al. (1998) studied backward connections from higher to lower visual areas of the macaque monkey, and reported that feedback projections served to amplify and focus the activity of units in the lower areas. Similarly, Lee, Mumford, Romero, and Lamme (1998) proposed that V1 is engaged in many levels of visual analysis through intra-cortical and feedback connections. Lamme and Roelfsema (2000) also concluded that feed-forward connections relay information from lower to higher visual cortical areas, but there are also horizontal, within-areas and, more importantly, feedback connections. In a recent report, Pascual-Leone and Walsh (2001) used transcranial magnetic stimulation (TMS) to demonstrate that the feedback projection from secondary visual areas to V1 is necessary for conscious visual

perception. Collectively these findings provide evidence for re-entrant, feedback connections, and interactions between lower and higher cortical visual areas. More importantly, they suggest that top-down connections have an impact on bottom-up processes in perception, attention, and recognition (Spratling & Johnson, 2004).

Some electrophysiological findings of dynamic shifts of voltage change at the scalp surface are also consistent with re-entrant migration of information between higher and lower levels of information processing. Curran, Tucker, Kutas, and Posner (1992) examined visual event-related brain potentials (ERPs) during a word reading task. They reported that following N1 (the first negative deflection following stimulus onset) a separate posterior positive pattern emerges (termed as the 'P1-reprise') that seemed to repeat the topography of P1. According to the authors, the scalp distribution of this effect was similar to the P1 and it seemed unlikely that it reflected stimulus offset. Another example of back-projection during a recognition task is described in De Haan, Pascalis, and Johnson (2002). These authors also used ERPs to compare the spatial and temporal characteristics of electro-cortical activation during the early stages of

* Corresponding author. Fax: +44 20 7631 6312.

E-mail address: d.mareschal@birk.ac.uk (D. Mareschal).

ce processing in adults and 6-month-old infants. They reported that in both cases there is apparent dynamic movement of voltage change consistent with migration of information along the central visual pathway. This is followed by a re-activation of overlapping visual areas. Finally, Martínez et al. (1999, 2001) used ERPs combined with functional magnetic resonance imaging (fMRI) in a series of studies in order to investigate the cortical mechanisms of visual-spatial attention. Thus, they reported that attentional modulation of activity in V1 was substantial as measured by fMRI but non-existent as measured by the early ERP components originating from V1. However, they found that a later ERP component, with a latency of 160–260 ms, was modulated by attention and was localized to V1. They concluded that V1 activity was affected by a delayed, re-entrant feedback from higher visual areas. Similar results have also been reported by others (e.g., Noesselt et al., 2002).

In the present study we use common-onset visual masking (COVM; sometimes also referred to as “object substitution masking”) as a tool for exploring re-entrant visual processing (Di Lollo, Enns, & Rensink, 2000, 2002). At a general level, visual masking refers to the reduction of visibility of one object (the target) by another object (the mask) that appears nearby in space or time. For instance, a highly visible object briefly presented in isolation can be made almost invisible when it is followed by another object occupying the same spatial location or even an adjacent, but not overlapping, location. This kind of masking is also referred to as *backward* masking since the detection of an object is impaired by events that occur subsequently (Breitmeyer, 1984). Visual backward masking is an empirically and theoretically rich phenomenon that can be a powerful methodological tool to study visual information processing and attention (Breitmeyer & Ogmen, 2000).

COVM is a form of backward masking that occurs when a brief display of the target plus a mask, that consists of only four black dots that surround but do not touch the target, is followed by the mask alone (Di Lollo et al., 2000, 2002; Enns and Di Lollo, 1997). According to Di Lollo et al., the first wave of bottom-up or bottom-up processing of the visual input is not sufficient for target identification. As a result, identification is aided by feedback or top-down neural projections during which, the circuit actively searches for a match between a descending perceptual hypothesis, and an ongoing pattern of low-level activity. The comparison between information from higher and lower areas allows a percept to be achieved, ensuring that information is consistent between both levels. However, when the target item is deleted and only the four-dot mask is left on the target location, the ongoing activity at the lower level would then consist of an image of the mask alone, and a decaying image of the target. This creates a mismatch between the ongoing (bottom-up) pattern of low-level activity and the re-entrant (top-down) perceptual hypothesis that includes both the target and the mask, leading to confusion and disruption of the target’s identification. What is consciously perceived would then depend on the number of iterations required to identify the target.

COVM is sensitive to the attentional demands of the task (Di Lollo et al., 2000, 2002). Indeed, COVM is greater when the target is surrounded by similar distracter items, which corresponds

to conditions that increase attentional demands in visual search tasks even when no mask is used (Treisman & Souther, 1985). By contrast, COVM is significantly weaker when the participant’s attention is directed to the target by a spatial precue (Di Lollo et al., 2000, 2002). Di Lollo et al. (2000, 2002) report that for COVM to occur attention must be distributed among many potential targets and the four-dot mask must remain on the screen after deletion of the target in order for the mask-alone representation to *substitute* the target-plus-mask representation. If this object substitution hypothesis is correct, then COVM is a tool for exploring the temporal dynamics of visual perception, and more specifically, the iterative processing that occurs when an initial visual representation is discarded if it is incompatible with subsequent attention-based analysis of the visual scene (Lamme & Roelfsema, 2000).

The present studies aimed to bring together the two sets of literature discussed above by investigating the electrophysiological pattern of activity associated with re-entrant processing during common-onset visual masking in adults. Note that there is currently very little evidence for neurophysiological correlates of backward masking, possibly because authors have looked for evidence of inhibition, rather than re-entrant processing (see Enns & Di Lollo, 2000 for further discussion). Both behavioral and electrophysiological data were collected and analyzed in the present studies. From a behavioral point of view, common-onset visual masking by four dots should become stronger as a joint function of set size and trailing mask duration (Di Lollo et al., 2000, 2002). From an electrophysiological point of view, visual information initially activates the early extrastriate visual cortical areas situated at the posterior part of the cortex (P1). Then, information is projected to more anterior parts of the cortex creating an occipito-temporal negative deflection (N1). These early ERP components usually occur within 200 ms following the presentation of the stimulus (Curran et al., 1992; Fabiani, Gratton, & Coles, 2000; Lamme & Roelfsema, 2000). However, according to Di Lollo et al.’s (2000, 2002) model, competing information or representations feed back from higher to lower visual areas for confirmation. When masking is strong this renewed activation of early visual areas corresponds to a hypothetical stronger mismatch between the re-entrant visual representation and the ongoing lower-level activity produced by current sensory input. In so far as the magnitude of a component reflects the size of the population of neurons generating the signal, we therefore expect to find a larger re-entrant positivity for conditions where masking is stronger, immediately following N1 post-stimulus onset (we will refer to this component as a posterior P2). A further prediction is that there should be a relationship between the magnitude of the re-entrant positivity and behavioural response accuracy.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Participants consisted of 12 neurologically normal paid volunteers (6 males) with an average age of 27.4 years ($SD = 4.6$ years). All participants had normal or corrected-to-normal vision and four of them were left-handed.

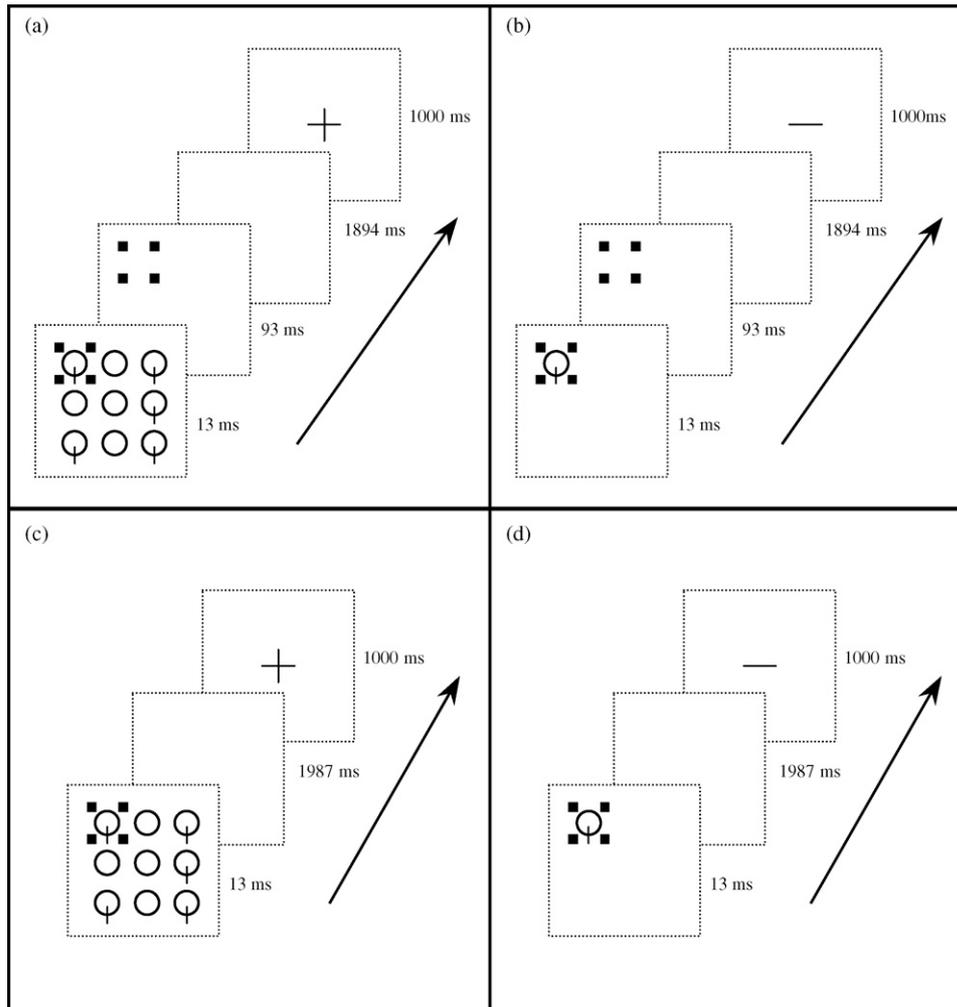


Fig. 1. Examples of stimuli sequences within different trials. Nine or one rings appeared on the screen followed (a and b) or not (c and d) by the mask (black dots). The participants' responses were acknowledged by a central feedback symbol: '+' for correct and '-' for wrong responses.

2.1.2. Stimuli

All stimuli were monochrome images displayed on a 21 in. computer monitor with 75 Hz refresh rate using in-house presentation software designed to interface with the ERP equipment. On any given trial either 1 or 9 complete rings were displayed in the cells of a notional $3.5^\circ \times 3.5^\circ$ matrix area positioned in the centre of the screen viewed from a distance of 57 cm. In the display containing nine rings, four of the rings had a vertical bar across the bottom, as shown in Fig. 1. The target stimulus was always singled out by four black dots that also served as a mask. Those dots were located approximately 0.25° from the target's contour. All rings were 1.0° , and the size of each dot was 0.2° . Overall, on half of the trials the target had a vertical bar, and on the remaining half of the trials the vertical bar was absent.

2.1.3. Procedure

At the beginning of each trial a fixation stimulus was displayed in the centre of the screen for 1 s. The display sequence that followed consisted of two frames that were presented sequentially and without interruption. The first frame, with an exposure duration of 13 ms, contained the target, the four-dot mask, and the distracters in the case of a set size of nine rings. The second frame contained only the trailing four dots for an exposure duration of either 0 or 93 ms. The only visual event that occurred between frames is the disappearance of the central target stimulus in the second frame. In between trials, the participants were given 2 s to indicate whether the target had a vertical bar or not by pressing the appropriate key on a button press box. More specifically, participants were instructed to press the "yes" key when they were sure the vertical bar was present, otherwise they were asked to press the "no" key. They were asked to

be conservative in their responses so as to try to avoid false positives. The button side was counterbalanced across participants so that half of the participants used their left hand to respond "yes" and their right hand to respond "no", and vice versa. Additionally, the participants' responses were acknowledged by a central feedback symbol ('+' for correct and '-' for wrong answers), which was displayed for 1 s and also served as the fixation marker for the next trial (see Fig. 1).

Each participant contributed nine blocks. The first block served as a practice block: no ERPs were recorded and it was not taken into account during subsequent data analysis. The practice block was followed by the actual test session, which consisted of eight blocks of 64 trials each. Although the matrix consisted of nine locations, the target never appeared in the centre of the matrix since the participants were asked to fixate at this location throughout the test session. The 64 trials within each block resulted from the factorial combination of two set sizes (one or nine rings), two trailing mask durations (0 ms or 93 ms), the presence or absence of the vertical bar at the bottom of the target, and the eight possible target locations within the matrix. Additionally, the participants were allowed to have breaks between blocks as needed. Participants typically took 2–3 breaks during the test session, which was approximately 40 min long.

2.1.4. Data acquisition and analysis of ERP data

EEG was recorded using a Geodesic Sensor Net (Tucker, 1993) consisting of 128 silver–silver chloride electrodes evenly distributed across the scalp. A ground electrode was positioned on the forehead. The EEG was recorded referenced to the vertex, with a band-pass of 0.1–100 Hz, and was sampled every 2 ms (500 Hz).

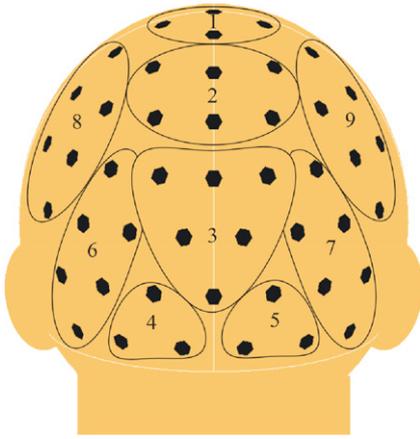


Fig. 2. Posterior view of the scalp with the electrode sites and their grouping for analysis: 1, central; 2, central parietal; 3, central posterior; 4, left occipital; 5, right occipital; 6, left posterior; 7, right posterior; 8, left parietal; 9, right parietal.

For each trial, the EEG was segmented to create an epoch from 200 ms before stimulus onset to 600 ms after stimulus onset. Data were then edited offline for artifacts. Within any given trial the activity at a sensor was excluded if it went off-scale, if the sensor was not making good contact or if its activity did not correspond to brain activity. An average of 8.77 percent of trials were excluded per participant. The entire trial was excluded if more than 10 of the sensors had been excluded or if there were eye-blinks or movement artifacts. Data were then baseline corrected for a period of 100 ms before stimulus onset. Following that, a separate average was created for each individual across trials for each of the eight conditions. Moreover, the data from the sensors with no or few (less than 10) trials in the average were interpolated using spherical spline interpolation (Herrn et al., 1989). Finally, the data were re-referenced to the average reference.

The analysis of the electrophysiological data focused on the first peak following N1 (we will refer to this peak as the posterior P2) post-stimulus onset. This analysis was carried out over all trials, independently of response accuracy, in order to have the same number of trials per condition. To assess any significant effects around the posterior P2, the electrodes over the posterior area of the scalp were grouped into nine groupings varying from a minimum of three to a maximum of six electrodes per grouping (see Fig. 2). The time window around the posterior P2 was defined as ± 1 SD around the mean latency across all conditions of the first peak following N1. This value corresponded to the mean latency of the condition with the strongest masking effect in the performance across participants.

2. Results

2.1. Performance

The sensitivity to the presence of the vertical bar under the various masking conditions was estimated according to the signal detection theory (Green & Swets, 1966). Table 1 shows the proportion of correct responses and d' values measured in all

Table 1
Average d' values across participants for Experiment 1

Set size	Trailing mask duration (ms)	Percent correct responses		d' values
		Absent vertical bar	Present vertical bar	
0	0	64.33	93.17	2.27
0	93	74.58	68.67	1.51
9	0	64.50	84.75	1.55
9	93	82.33	33.08	0.56

presentation conditions. A 2×2 repeated measures ANOVA with set size and the duration of the trailing mask as the within subjects factors was performed on these values. This analysis revealed a main effect of set size ($F(1, 11) = 25.84, p < .001$), indicating significantly reduced sensitivity to the presence of the vertical bar for a larger set size. A main effect of the duration of the trailing mask was also revealed ($F(1, 11) = 26.58, p < .001$), showing a significantly reduced d' value for a longer trailing mask duration. However, a two-way interaction between those factors did not reach statistical significance.

2.2.2. Event-related potentials

Fig. 3 shows the ERPs from the posterior channels groups. Not surprisingly, the early responses from the visual cortex differed greatly between the conditions with one or nine stimuli, reflecting the spatial extension differences between the two conditions. After the first negative component (N1) peaking at about 180 ms, however, the difference between the ERPs to the different set sizes diminished. As our aim was to measure electrophysiological correlates of reentrant visual processes, hypothesized to be reflected by the second, much smaller positive peak over the extrastriate areas, we focused our analysis to the P2 peak.

A measure of the mean amplitude for the posterior P2 was computed by calculating the average across all of the sampling points within a target time window defined as ± 1 SD around the mean latency of the first peak following N1. The mean latency of the P2 component occurred at approximately 219 ms ($SD = 7$ ms) post-stimulus onset, for the selected channels and across all conditions, delineating the time window from 212 to 226 ms post-stimulus onset. These measures were then analyzed in a $2 \times 2 \times 2$ repeated measures ANOVA with set size (one or nine rings), duration of the trailing mask (0 or 93 ms) and vertical bar (absence or presence) as the within subject factors. A main effect of the duration of the trailing mask was found to be significant in four out of nine channel groupings ($F(1, 11) > 5.47, p < .05$) with increased amplitudes in the delayed offset condition compared to the common offset condition. These channel groups were situated in Central, Left Occipital, and Left and Right Posterior areas of the scalp.

2.3. Discussion

We recorded ERPs while participants were engaged in the COVM task developed by Di Lollo et al. (2000). Our behavioral results were consistent with the original findings. Participants were significantly worse at detecting the target among nine elements than in one element, and were significantly less sensitive to the presence of the target line with delayed than with simultaneous offset of the mask. When these two factors were added together, the target stimulus was almost not detectable at all (as reflected by the d' value close to 0).

In addition, regardless of the participants' response, we found evidence of more positive deflections for the posterior P2 when participants viewed delayed offset stimuli compared to common offset stimuli. This effect was independent of set size and may reflect increased processing of stimuli in the extrastriate cortex

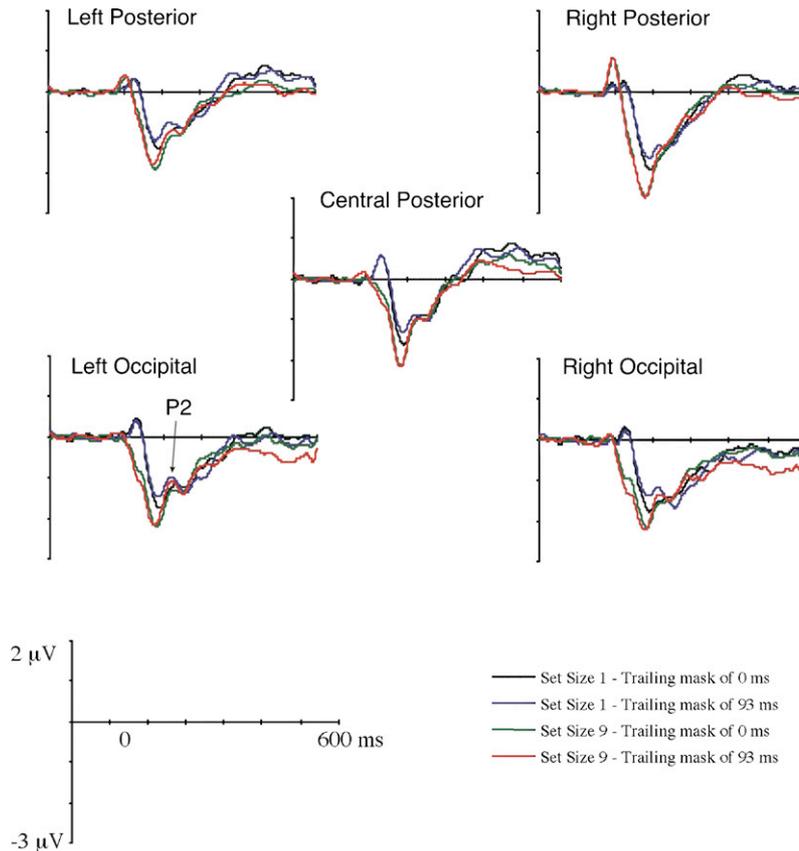


Fig. 3. Grand average waveforms representing some of the channel groupings (the numbers represent the channel groupings as specified in Fig. 2) for Experiment 1. ERPs to stimuli with or without the vertical bar are averaged together.

with trailing mask. This is in line with expectation if feedback from higher visual areas finds a visual signal, left by the trailing mask, that is incompatible with the object extracted by the first analysis of the stimulus.

Before concluding that the effect of delayed mask on the posterior P2 amplitude reflected re-entrant visual processing, we have to consider two limitations of the present result. First, we did not find an interaction between the two main factors affecting sensitivity (d') to the target stimuli, and the interaction was also missing from the P2-effect. It is therefore possible that the modulation of the P2 amplitude reflected only the effect of different masking conditions and not object substitution *per se*, which has been shown to be dependent on the set size over which attention is initially distributed (Di Lollo et al., 2000, 2002). Second, and relatedly, the peak of the posterior P2 occurred approximately 125 ms after the offset of the delayed mask, and its modulation might have been caused by the first positive peak of the offset-related potential. Experiment 2 was designed to investigate this possibility.

3. Experiment 2

In Experiment 2, we increase the target exposure time to examine its effect on both behavior and electrophysiological response. If the posterior P2-effect found in Experiment 1 was due to the later extinction of the mask in the delayed offset

condition, we would expect the latency of this effect to move with the change in stimulus offset.

3.1. Methods

3.1.1. Participants

Participants consisted of 20 neurologically normal, paid volunteers (8 males) with an average age of 25.5 years ($SD=4.4$ years). All participants were right handed and had normal or corrected-to-normal vision.

3.1.2. Stimuli

Stimuli were identical to those used in Experiment 1.

3.1.3. Procedure

The procedure was identical to that in Experiment 1 except that the target now remained visible for 40 ms rather than the 13 ms in the first Experiment. As in the original study, the trailing mask remained visible for a full 93 ms following target extinction in the delayed offset condition. Data acquisition and analysis of ERP data was identical to that of Experiment 1.

3.2. Results

3.2.1. Performance

Similarly to Experiment 1, sensitivity to the presence of the vertical bar was analyzed in a 2×2 repeated measures ANOVA with set size (one or nine rings) and the duration of the trailing mask (0 or 93 ms) as the within subjects factors on the corresponding d' values (see Table 2). This analysis revealed a main

Table 2
Average d' values across participants for Experiment 2

Set size	Trailing mask duration (ms)	Percent correct responses		d' values
		Absent vertical bar	Present vertical bar	
0	0	69.85	94.7	2.52
93	0	81.75	85.9	2.41
0	93	75.15	83.6	1.83
93	93	88.45	38.40	1.01

Effect of set size ($F(1, 19) = 38.36, p < .001$): the d' values were significantly smaller for the larger set size. A main effect of the duration of the trailing mask was also revealed ($F(1, 19) = 19.69, p < .001$), showing a significantly reduced d' value for a longer trailing mask duration. Finally, a significant two-way interaction ($F(1, 19) = 7.96, p < .011$) demonstrated that sensitivity to the vertical bar affected by duration mask more in the larger set size than in the single stimulus condition.

2.2. Event-related potentials

Fig. 4 depicts the ERPs at posterior channel groups in all four conditions. As for the Experiment 1, the analysis focused on P2 amplitude, quantified as the average voltage in the time window defined as $\pm 1 SD$ around the mean latency of the first peak following N1. The mean latency at which the posterior P2 component occurred was approximately 220 ms post-stimulus onset

($SD = 6$ ms), for the selected channels and across all conditions. A measure of the mean latency was computed by calculating the average across all of the sampling points within the target time window from 214 to 226 ms. These measures were then analyzed in a $2 \times 2 \times 2$ repeated measures ANOVAs with set size (one or nine rings), duration of the trailing mask (0 or 93 ms) and presence or absence of the vertical bar, as the within subjects factors. This analysis revealed a main effect of the duration of the trailing mask in four out of nine channel groupings: Central and Right Posterior ($F(1, 19) = 10.58$ and 17.69 , respectively, $p < .005$) and Left and Right Occipital ($F(1, 19) = 5.61$ and 8.22 , respectively, $p < .05$) areas. We also found a significant two-way interaction between set size and the duration of the trailing mask in the Right Posterior channel grouping ($F(1, 19) = 4.45, p < .05$).

3.3. Discussion

As in Experiment 1, the behavioral results were consistent with the original Di Lollo et al. (2000, 2002) findings. In the delayed offset condition participants were significantly worse at detecting the target, and the larger set size also reduced the participants' sensitivity to the presence of the target line. Crucially, the significant interaction between these factors showed that the delayed offset of masking had a bigger impact on sensitivity in the set size 9 condition than it did with the set size 1 condition.

We also replicated the modulation of P2 amplitude by the duration of the trailing mask, found in Experiment 1. Note

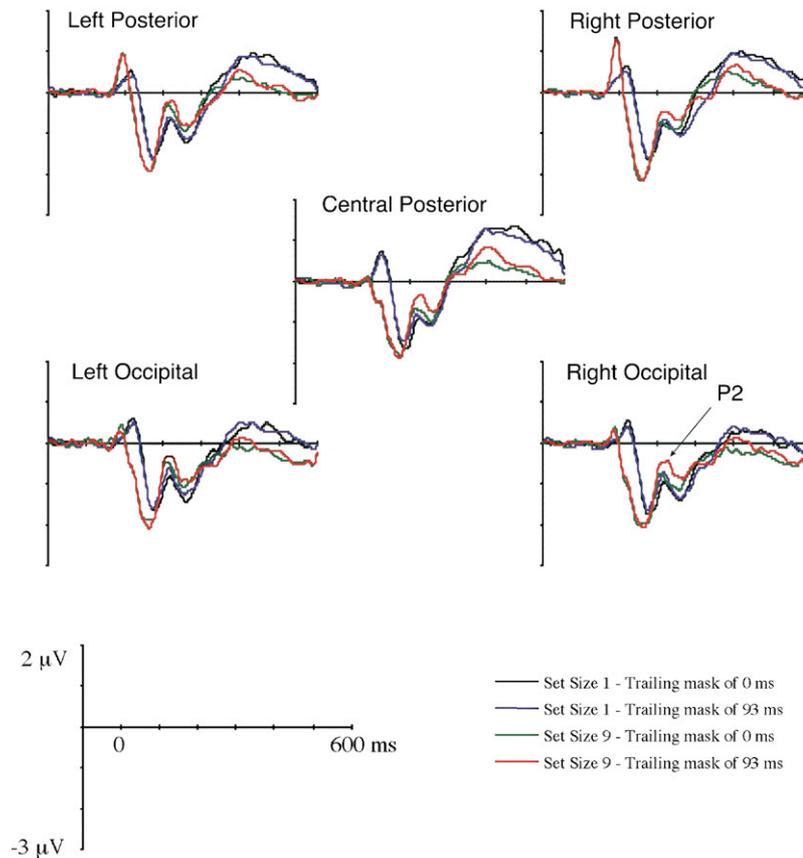


Fig. 4. Grand average waveforms representing some of the channel groupings (the numbers represent the channel groupings as specified in Fig. 2) for Experiment 2. ERPs to stimuli with or without the vertical bar are averaged together.

that the latency of this effect did not differ from that found in Experiment 1, indicating that its timing was determined by the processing of the initial visual stimulus rather than the offset of the mask. In addition to the main effect of mask duration, we also found an interaction between mask duration and set size, which was a result of bigger posterior P2 modulation by masking condition with nine possible elements on the screen. This interaction mirrors the one observed in the behavioral results, where performance decreased as a joint function of a larger set size and a longer trailing mask duration. Unlike the effect of masking duration, which was widely distributed over the visual cortex, this interactive effect appeared to be localized to the right occipito-temporal areas.

Thus, the electrophysiological effect demonstrated in this study is consistent with the hypothesis that the neural processes behind the posterior P2 component represent reactivation of primary and secondary visual cortices by feedback from higher other visual areas. The bigger the mismatch between the top-down and bottom-up signals, the higher this activation, and the lower the sensitivity to the initial stimulus. The timing of this reactivation seems to depend on the time it takes to receive feedback from the first stimulus, rather than on the timing of the second, mismatching stimulus.

4. Comparing response performance across the two experiments

In the two experiments we have found evidence of increased P2 amplitude as a function of masking conditions. In this section we assess directly whether there is a relationship between P2 amplitude and response accuracy, and whether this relation persists across the different target durations of Experiments 1 and 2.

To this end, we considered the responses and ERP amplitudes in the only condition that provided consistent evidence of masking (bar present, mask present, nine elements). An aggregate P2 amplitude score was computed by summing over the four sites in Experiments 1 and 2 that showed significant P2 effects. We then ran an analysis of covariance (ANCOVA) on the P2 aggregate measure, pooled across Experiments 1 and 2, with response accuracy as the dependent variable, Experiment (1 and 2) as a between subject factor, and aggregate P2 value (ERP) as a covariate expanded to the second quadratic order. All interaction terms were also included. This revealed a significant effect of ERP ($F(1, 26) = 6.51, p < .02$) suggesting that there is a significant linear relation between ERP (P2 amplitude) and response accuracy, as well as a highly significant quadratic relation between ERP and response accuracy ($F(1, 26) = 9.75, p < .005$). No other main effect or interaction was significant. Thus, while there is clearly a strong relation between the P2 amplitude and the participants' response accuracies, the dominant quadratic term suggests that this is not a simple linear relation.

5. General discussion

Overall, the behavioral data from both studies were consistent with the findings reported by Di Lollo et al. (2000, 2002): common-onset visual masking became stronger as a joint func-

tion of a large set size and long trailing mask duration. Indeed, an increase in set size combined with an increase in the duration of the trailing mask had an adverse effect on the sensitivity to targets. Clearly, the two factors cannot be considered in isolation: it is their interaction that produces common-onset visual masking. This is consistent with the re-entrant hypothesis, which predicts that a larger number of iterations are required in order to detect a target amongst increasing number of distracters. It also explains how a large set size (one target plus eight distracters) combined with a trailing mask of 93 ms produces significantly stronger masking than a single target item or when the mask and the target disappear simultaneously.

Set size and the duration of the trailing mask also interacted with the presence or absence of the vertical bar. This artifact was present for both studies. When a vertical bar was present, the pattern of results indicated that an augmentation in set size combined with longer trailing mask duration had an adverse effect on the accuracy of target reports. However, for the conditions where the vertical bar was absent it seemed as if the identification process was not affected by set size or longer trailing mask duration. On the contrary, performance seemed to be higher for a larger set size or longer trailing mask duration. This was due to the fact that in this case a negative response was always the correct one to produce. Indeed, either the participants were able to detect the absence of the vertical bar and produce a negative response, or a masking effect led them to a negative response, which in this case added to their overall performance. Similarly, Di Lollo et al. (2000, 2002) reported that on trials that did not contain a vertical bar participants' performance was at ceiling except at a trailing mask duration of zero, where results were comparable with those obtained on trials where the vertical bar was present.

Analyses on sensitivity of the presence of the vertical bar (expressed in d' values) from the first study revealed a main effect of set size and a main effect of the duration of the trailing mask. Both of these effects indicated that sensitivity to the vertical bar was affected by larger set size and longer trailing mask duration. Similar analyses for the second study (with a larger n) indicated that the participants' sensitivity to the target (d') significantly decreased as a joint function of larger set size and longer trailing mask duration.

Taken collectively, these data demonstrate the robustness of common-onset visual masking. Indeed, although the task was facilitated by longer target exposure, the participants' performance, as well as their sensitivity, were still very much affected by the combination of the two key factors: set size and trailing mask duration. In general, most behavioral effects reported in Experiment 1 got even larger in Experiment 2.

Analysis of the electrophysiological data from Experiment 1 indicated a main effect of the duration of the trailing mask for posterior P2 amplitudes over the left posterior temporal and occipital areas as well as in the right temporal areas of the brain. This effect was evident as less negative amplitude around 212–226 ms post-stimulus onset in those conditions with a trailing mask. Electrophysiological data from Experiment 2 revealed a significant interaction between set size and the duration of the trailing mask in the amplitude of the posterior P2. This inter-

tion was evident as a less negative amplitude around 220 ms post-stimulus onset in those conditions where stronger masking was observed. Most interestingly, the posterior P2 component was located as posteriorly as in Experiment 1 and its mean latency was very similar, even though the duration of the target in Experiment 2 was much longer. Visual information usually arrives at the medial part of the occipital lobes and then travels towards more anterior and lateral parts of the brain. Thus, we suggest that the P2 component reflects the effects of a feedback mechanism on visual processing.

The posterior P2 was reduced (less positive-going) when little or no behavioral masking occurred. In our view, this activity corresponds to a mismatch between a descending code representing the perceptual hypothesis (the target and the surrounding dots) and an ongoing pattern of low-level activity (a blank screen or the task alone). The magnitude of a component is generally thought to reflect the size of the population of neurons generating the signal (Nunez, 1981). According to this interpretation, a larger P2 means that more cells are active. Suppose a population code used to represent feature information (as is the case in much of the visual cortex; see Nicolelis, 2001). If cells are activated both by the feed-forward perceptual input and the top-down feedback information, then the set of cells activated will be at a minimum when the top-down information activates exactly those cells being activated by the bottom-up code. Any mismatch will result in a greater number of cells being activated (e.g., some different cells might be activated by new perceptual input signal while the original cells will be activated by the top-down returning signal, thus the total number of cells activated will be greater than in the match condition). Taken collectively, the results of both studies indicate that the P2 component indexes a re-entrant mechanism. The importance of the P2 for the behavioural masking effect was further reinforced by our analysis of the relation between the amplitude of this component and response accuracy in the critical condition across the two experiments. This analysis revealed both linear and quadratic relations between P2 amplitude and response accuracy, indicating a tight causal coupling between the two measures.

Our electrophysiological results are also consistent with previous studies reporting recurrent processing or feedback projections towards primary visual areas (Hupe et al., 1998; Lamme & Roelfsema, 2000; Lee et al., 1998; Martínez et al., 1999, 2001; Mason, Chun, & Allison, 2001; Pascual-Leone & Walsh, 2001; Rodorff et al., 2002). Although these studies differ in terms of timing, polarity, laterality, or amplitude data, they all provide clear evidence of re-entrant, feedback or back-projections and interactions between cortical areas. Similarly, our results demonstrated the existence of dynamic changes in cortical activity consistent with a feedback re-entrant pathway during target detection.

Visual cortical neurons are not just simple detectors limited to one aspect of the visual scene. Primary visual areas remain active long after their initial participation in the feed-forward, bottom-up pathway. Feedback connections originating from higher visual areas allow lower ones to contribute to different levels of analysis at later points in time. It is at those longer latencies that information feeding back from higher cor-

tical areas exerts its influence and can be incorporated into a perceptual awareness (Lamme & Roelfsema, 2000; Lee et al., 1998).

Although the COVM phenomenon has been replicated several times there is considerable debate over whether it requires re-entrant processing as suggested by DiLollo and colleagues, or whether a feed-forward account involving competition between two neural object representations can explain the data equally well (e.g., Di Lollo et al., 2002; Francis & Hermens, 2002; Keyser & Perrett, 2002; Neill, Hutchison, & Graves, 2002). It is unclear to us whether this debate can ever be fully resolved. In part, this is because such arguments depend on how competition is implemented. For example, one recent neural computational account of attentional competition between two simultaneously presented visual stimuli relies crucially on local re-entrant processing (Spratling & Johnson, 2004). In such cases, the distinction between feed-forward competition and re-entrant processing is unclear.

That said, our data does shed light on the debate surrounding whether COVM is explained by a feed-forward competitive process. Broadly speaking, two types of feed-forward competitive accounts have been put forward to explain masking. One relies on the temporal contiguity of two rapidly presented stimuli (Keyser & Perrett, 2002; Neill et al., 2002). According to this view, the processing of the later arriving stimulus overcomes the processing of the initial stimulus whose encoding has not been fully completed prior to the arrival of information from the next stimulus. The second account relies on the spatial contiguity of two objects in space. When two objects are in close spatial proximity a competition for visual attention ensues, with one stimulus winning out at the expense of the other (e.g., Desimone & Duncan, 1995). Our data can rule out the first (temporal contiguity) account because we found no latency effects on the P2 across the two experiments. Any account that relied on the encoding of the target plus the mask followed by the mask only as two distinct chunked stimuli separated in time would predict latency difference as a function of the target duration (a factor that varied across experiments). We found no evidence of this suggesting that P2 latencies are determined by internal processing constraints (such as the time for required for re-entrant information to feedback) as opposed to stimulus temporal presentation characteristics. Our data remains ambiguous with regards to the second (spatial proximity) account of competition. However, we note that some authors have questioned whether such accounts can actually be tested empirically using ERPs (Driver et al., 2004).

Finally, our findings are also consistent with a recent fMRI study showing increased activation during object substitution match (Weidner, Shah, & Fink, 2006). The relative increase in BOLD signal obtained during masking trials is entirely consistent with the increase of the ERP amplitude we found, thus providing converging evidence for our interpretation. While the fMRI results revealed increased activation in V1, the middle occipital gyrus, the transvers occipital gyrus, and the left intraparietal sulcus, these effects operate at a much longer time scale than those reported here. Thus, the scalp topology that we report is not inconsistent with the fMRI data and provides crucial com-

plementary information concerning the co-timing of neural and visual events in the COVM.

In summary, these studies represent the first step in establishing an ERP marker related to common-onset visual masking. A combination of ERPs and functional magnetic resonance methodologies within a single study would be an interesting follow-up in order to establish more specifically the dynamics as well as the sources of the P2 component during common-onset visual masking. Good temporal resolution paired with excellent spatial resolution should provide us with a much more detailed account of this phenomenon. This will assist us in understanding how cortical areas intercommunicate or how top-down processes interact or influence bottom-up information processing.

Acknowledgements

We would like to thank all the participants who took part in this study. This work was supported by the European Commission grants HPRN-CT-1999-00065 and 516542-NEST, and MRC Program Grant G9715587.

References

- Breitmeyer, B. G. (1984). *Visual Masking: An Integrative Approach*. Oxford, UK: Oxford University Press.
- Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, *62*, 1572–1595.
- Curran, T., Tucker, D. M., Kutas, M., & Posner, M. I. (1992). Topography of the N400: Brain electrical activity reflecting semantic expectancy. *Electroencephalography and Clinical Neurophysiology*, *88*, 188–209.
- De Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, *14*, 199–209.
- Desimone, R., & Duncan, J. (1995). *Annual Review of Neuroscience*, *18*, 193–222.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of re-entrant visual pathways. *Journal of Experimental Psychology: General*, *129*(4), 481–507.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2002). Object substitution without reentry? *Journal of Experimental Psychology: General*, *131*, 594–596.
- Driver, J., Eimer, M., Macaluso, E., & van Velzen, J. (2004). Neurobiology of human spatial attention: Modulation, generation, and integration. In N. Kanawisher & J. Duncan (Eds.), *Functional neuroimaging of visual cognition: Attention and performance XX* (pp. 267–300). Oxford, UK: Oxford University Press.
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, *8*, 135–139.
- Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345–352.
- Fabiani, M., Gratton, G., & Coles, M. G. H. (2000). Event-related brain potentials: Methods, theory, and applications. In J. T. Cacioppo, L. G. Tassinary, & G. C. Bernston (Eds.), *Handbook of psychophysiology*. Cambridge (UK): Cambridge University Press.
- Francis, G., & Hermens, F. (2002). Comment on “Competition for consciousness among visual events: The psychophysics of reentrant visual processes (Di Lollo, Enns, and Rensink, 2000)”. *Journal of Experimental Psychology: General*, *131*, 590–593.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley.
- Hupe, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and ground by V1, V2, and V3 neurons. *Nature*, *394*, 784–787.
- Keysers, C., & Perrett, D. I. (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences*, *6*, 120–125.
- Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience*, *23*, 571–579.
- Lee, T. S., Mumford, D., Romero, R., & Lamme, V. A. F. (1998). The role of the primary visual cortex in higher-level vision. *Vision Research*, *38*, 2429–2454.
- Martínez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., et al. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature: Neuroscience*, *2*(4), 364–369.
- Martínez, A., DiRusso, F., Anllo-Vento, L., Sereno, M. I., Buxton, R. B., & Hillyard, S. A. (2001). Putting spatial attention on the map: Timing and localization of stimulus selection processes in striate and extrastriate visual areas. *Vision Research*, *41*, 1437–1457.
- Neill, W. T., Hutchison, K. A., & Graves, D. F. (2002). Masking by object substitution: Dissociation of masking and cuing effects. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 682–694.
- Nicolelis, M. A. L. (2001). *Advances in neural population coding*. Elsevier.
- Noesselt, T., Hillyard, S. A., Woldorff, M. G., Schoenfeld, A., Hagner, T., Jancke, L., et al. (2002). Delayed striate cortical activation during spatial attention. *Neuron*, *35*, 575–587.
- Nunez, P. L. (1981). *Electric fields of the brain: The neurophysics of EEG*. London: Oxford University Press.
- Olson, I. R., Chun, M. M., & Allison, T. (2001). Contextual guidance of attention: Human intracranial event-related potential evidence for feedback modulation in anatomically early, temporally late stages of visual processing. *Brain*, *124*, 1417–1425.
- Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science*, *292*, 510–512.
- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current-density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*, 184–187.
- Spratling, M., & Johnson, M. H. (2004). A feedback model of visual attention. *Journal of Cognitive Neuroscience*, *16*, 219–237.
- Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal Of Experimental Psychology: General*, *114*, 285–310.
- Tucker, D. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *EEG and Clinical Neurophysiology*, *87*, 154–163.
- Weidner, R., Shah, N. J., & Fink, G. R. (2006). The neural basis of perceptual hypothesis generation and testing. *Journal of Cognitive Neuroscience*, *18*, 258–266.
- Woldorff, M. G., Liotti, M., Seabolt, M., Busse, L., Lancaster, J. L., & Fox, P. T. (2002). Cognitive. *Brain Research*, *15*, 1–15.