REFLEXIVE ATTENTION MODULATES PROCESSING OF VISUAL STIMULI IN HUMAN EXTRASTRIATE CORTEX

Joseph B. Hopfinger and George R. Mangun
Center for Neuroscience and Department of Psychology, University of California, Davis

Abstract—Attention can be oriented reflexively to a location in space by an abrupt change in the visual scene. In the present study, we investigated the consequences of reflexive attention on the neural processing of visual stimuli. The findings show that reflexively oriented attention produces modulations in early sensory analysis at the same extrastriate neural locus as the earliest effects of voluntarily focused attention. In addition, stimulus processing was found to be enhanced at later stages of analysis, which reflect stimulus relevance. As is the case with behavioral measures of reflexive attention, these physiological enhancement effects are rapidly engaged but short-lived. As time passes between the initial attention-capturing event and subsequent stimuli, the extrastriate effect reverses, and the enhancement of higher order processing subsides. These findings indicate that reflexive attention is able to affect perceptions of the visual world by modulating neural processing as early as extrastriate visual cortex.

Visual attention can be oriented voluntarily to a location or captured reflexively by sensory stimuli (e.g., Egly, Rafal, Henik, & Berger, in press; Jonides, 1981; Müller & Rabbitt, 1989; Yantis, 1996). Voluntarily directed visual attention results in faster and more accurate responses to stimuli at the attended location (e.g., Hawkins et al., 1990; Posner, Nissen, & Ogden, 1978). It is now well established from human and animal research that voluntary visuo-spatial attention results in a modification of cortical sensory processing (e.g., Mangun & Hillyard, 1991; Moran & Desimone, 1985; Motter, 1993). In humans, voluntary attention to location leads to a relative facilitation of visual processing in extrastriate visual cortex that begins by 70 to 80 ms after presentation of the target (e.g., Eason, 1981; Heinze et al., 1994; Luck et al., 1994; Mangun & Hillyard, 1991; Van Voorhis & Hillyard, 1977).

Reflexive attention, however, may be controlled by neural mechanisms partially or wholly separate from those involved in voluntary attention (see Rafal, 1996, for a review). In addition, psychophysical and neuropsychological studies indicate that these two forms of attentional orienting have distinct properties. For example, Briand and Klein (1987) suggested that reflexive attention performs feature integration, whereas voluntary attention alone does not. Other researchers have reported that compared with voluntary attention, reflexive attention is engaged more rapidly, is more resistant to interference, and dissipates more quickly (e.g., Cheal & Lyon, 1991; Jonides, 1981; Müller & Rabbitt, 1989; Posner & Cohen, 1984). In addition, although voluntary attention may be stable over time, the effects of reflexive attention change significantly as time passes after an attention-capturing event (i.e., a nonpredictive exogenous “cue”). These effects of reflexive orienting have been demonstrated by measuring behavioral responses to target stimuli that are or are not preceded at the same location by task-irrelevant cues. Specifically, it has been found that at short cue-to-target intervals, targets are responded to more quickly if they appear at the same location as the previous cue (e.g., Cheal & Lyon, 1991; Jonides, 1981), whereas at longer cue-to-target intervals, stimuli may actually be responded to more slowly at cued locations than at uncued locations (Posner & Cohen, 1984). To date, however, no physiological studies in animals or humans have investigated the purely reflexive effects of nonpredictive cues on visual cortical activity. Although neurophysiological studies have shown that voluntary attention results in modulations of neural activity as early as extrastriate cortex (e.g., Mangun & Hillyard, 1991; Moran & Desimone, 1985; Motter, 1993), it has remained unknown whether reflexive attention may also affect sensory stimulus processing in visual cortex.

In the present study, we investigated the effects of reflexively oriented attention on visual processing by measuring neural activity in human subjects using the event-related potential (ERP) method. This method measures the activity of neuronal assemblies in real time, thereby allowing the ongoing processing of visual information within the brain to be studied. In the present experiment, ERPs were recorded from human subjects while they performed a discrimination task in which a nonpredictive cue preceded each target stimulus.

METHOD

Subjects viewed a NEC 4FG color computer monitor (at a distance of 75 cm) and maintained fixation on a centrally located cross throughout all trials (see Fig. 1). The background screen was dark gray. On either side of fixation, four small white dots demarcated the corners of an imaginary rectangle 1.03° wide and 1.37° tall. The center of each imaginary rectangle was located 1.5° above and 6.4° lateral to fixation. The beginning of each trial commenced with the four dots on one side of fixation (the left or right side was equally probable) being extinguished for 34 ms and then reappearing, giving the subjective impression of a blinking of one set of dots. This cue was used in order to minimize neuronal refractory effects and overlap from the cue ERP onto the target-activated ERP, while still producing an effective sensory cue. Subjects were informed that the cue would be completely nonpredictive of the location of the subsequent target.

Catch trials, during which no target appeared, accounted for 20% of the trials. The presence of catch trials reduced the likelihood of subjects forming temporal expectancies and prevented anticipatory responses. On the remaining 80% of trials, after a variable interval (ranging randomly from either 34 to 234 ms or 566 to 766 ms; rectangular distribution within each range), a vertical target bar was flashed to one side of fixation, centered between the dots on that side. The target bar was located to the left or right of fixation with equal probability, and it was as likely to be at the same hemifield location as the preceding cue as it was to be at the opposite hemifield location. The target remained on the screen for 50 ms and was either a short (1.8° by 0.69°) or tall (2.3° by 0.69°) vertical bar (for some subjects, the height of the tall bar was adjusted in order to maintain performance at around...
Electrophysiology of Reflexive Attention

Data from 8 healthy, right-handed, paid volunteers (4 female, 4 male) were analyzed. Their ages ranged from 18 to 30, and all had normal or corrected-to-normal vision. Reaction times (RTs) from the button-press responses were used as a behavioral measure of performance, and physiological measures were gathered by recording from 64 electrodes distributed over the scalp of each volunteer. These scalp-recorded signals were digitized at 256 samples per second and averaged to produce ERPs. The signals were then low-pass filtered to remove high-frequency noise, and a high-pass single-pole causal filter was used to remove the low-frequency drifts across time. Although the cue was subtle and did not overlap on the retina with the target, the scalp-recorded neural responses to the cue still overlapped with the recorded responses to the target stimulus, especially at the shortest ISIs. In order to eliminate the possibility that any differences in early ERP components might be due to overlapping neural activity produced by the cues, we employed the adjacent response (Adjar) filter method (Woldorff, 1993) to remove confounding potentials generated by the lateralized cues. This procedure iteratively estimates and subtracts the overlap from adjacent events (i.e., cue and target) until the estimates of the cue-and-target overlap do not change over successive iterations, at which point overlap is considered to have been removed from the original waveforms. (See the appendix for a further description of the effectiveness of the Adjar procedure.)

A strict criterion was used for inclusion of subjects’ data in the analysis, in order to facilitate the process of correctly subtracting overlapping components. Eye position was monitored with an infrared video camera system and by recording the electro-oculogram from electrodes placed around each eye. Trials with eye movements or blinks were rejected off-line and not included in the analysis. Subjects were then blindly ranked by an independent rater on the basis of the signal-to-noise ratio of the visually evoked ERP components (amplitude of poststimulus evoked components vs. residual activity in the prestimulus baseline). The rater had no knowledge of the effects of interest in this study. A median split was performed on the basis of this ranking, and the data from the 8 subjects (from 16 total) with the highest signal-to-noise ratios were then analyzed further. The data from these 8 subjects were individually filtered with the Adjar algorithm, in order to remove overlapping responses. The data presented here are the grand-average waveforms from the Adjar-filtered data set.

Statistical significance was evaluated using repeated measures analysis of variance (ANOVA). For the RT data, the ANOVA factors were cuing (cued- vs. uncued-location target), visual field of target (right vs. left hemifield), and subject (N = 8). For the analyses of ERP data in the 90- to 140-ms latency range (corresponding to the P1 component), the following factors were added: electrode location (medial vs. lateral) and hemisphere of electrode (right vs. left). For this time range, the occipital electrodes included in the analyses were T5, T6, O1, and OR. O1 and OR are located midway between T5 and O1, and T6 and O2, respectively, of the International 10-20 system of electrode placement (Jasper, 1958). Finally, for the analyses of the ERP data in the 250- to 500-ms latency range (corresponding to the P300 component), only midline scalp electrodes were analyzed (Cz and Poz), and thus the ANOVA factor of hemisphere was not included. For the purposes of the present report, only the main effects of cuing are presented. However, no other statistical main effects or interactions bear on the conclusions of the present study.

RESULTS AND DISCUSSION

As in prior RT studies using nonpredictive peripheral visual onsets (Jonides, 1981; Müller & Rabbitt, 1989; Theeuwes, 1991; Yantis, 1996) or offsets (Miller, 1989; Theeuwes, 1991), subjects in the present experiment were significantly faster in responding to targets at the cued location than at the uncued location for the short ISIs of 34 to 234 ms (Table 1). At the long ISIs (566 to 766 ms), no difference in RTs was observed between targets at cued and uncued locations. Classical inhibition of return (IOR), wherein RTs are typically slowed at cued locations at long ISIs (Posner & Cohen, 1984), was not observed, a result that is attributable to the use of a discrimination task. Such tasks are known to reduce the likelihood of obtaining RT inhibition (Egly et al., in press; Pratt, 1995; Terry, Valdes, & Neill, 1994).

When the cue-to-target ISI was short (34 to 234 ms), targets at the cued location elicited visual ERPs with significantly enhanced amplitudes compared with targets at the uncued location (Fig. 2 and Table 1). This enhancement was observed for a positive-polarity sensory ERP peak known as the P1 component. This component typically occurs 90 to 140 ms after stimulus presentation and is maximal in amplitude over lateral occipital scalp sites (e.g., Mangun, 1995; Mangun & Hillyard, 1991). At the longer cue-to-target ISIs (566 to 766 ms), however, the effect of the cues on the P1 component was

![Fig. 1. Examples of the stimulus display. The top row shows a trial with a target occurring at a cued location; the bottom row shows a trial with this target occurring at an uncued location. The cue was a 34-ms offset and then reappearance of the four dots on one side of fixation, giving the subjective impression of a blinking of the set of dots. The cue-to-target interstimulus interval was randomly varied over a short (34 to 234 ms) or long (566 to 766 ms) interval.](image-url)
reversed: Targets at cued locations elicited significantly smaller responses than targets at uncued locations. This reduction in P1 amplitude at long ISIs cannot be attributed to a simple neuronal refractory effect between cue and target because at short ISIs, when such an effect would be greatest, the pattern is opposite to that predicted by neuronal refractoriness.

By investigating topographic voltage maps during the time period of the P1 component, one can observe that the location of the maximal response at the scalp corresponding to the P1 component was highly similar for targets at cued and uncued locations (Fig. 3). This pattern is consistent with the view that the same type of neural process was evoked in both cases, with the primary difference being the strength of the response. The occipital P1 component represents the earliest stage of visual processing to be reliably modulated by voluntary spatial attention (e.g., Heinzé et al., 1994; Mangun, 1995; Mangun & Hillyard, 1991), and combined ERP and positron emission tomography studies have shown that during voluntary orienting, the P1 attention effect originates in extrastriate cortex in the ventral visual pathway (Heinzé et al., 1994; Mangun, Hopfinger, Kussmaul, Fletcher, & Heinzé, 1997). The present findings indicate that reflexive attention leads to modulations at this same stage of visual cortical processing, although, presumably, partially or wholly distinct control circuitry is involved in producing these two forms of attentional effects (Kustov & Robinson, 1996; Rafal, 1996; Robinson & Kertzman, 1995).

At short ISIs in this task, the P1 showed the same pattern as the subjects’ overt behavior—that is, facilitation at the cued location. Moreover, at longer ISIs, the P1 and RT facilitation both subsided, with the P1 actually becoming smaller in amplitude at the cued location. This pattern in the P1 is reminiscent of IOR. However, given the absence of inhibition in RT in these data, we draw no conclusions here regarding the possible contribution of this modulation of early cortical visual processing to the generation of RT IOR.

In the current experiment, it was possible to evaluate the impact of facilitated and inhibited sensory inputs on later stages of analysis. In

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{Interstimulus interval} & \textbf{Target} & \textbf{Difference} & \textbf{Significance} \\
\hline
 & \textbf{Cued} & \textbf{Uncued} & \textbf{(Cued – uncued)} & \\
\hline
\textbf{Manual response time to targets (ms)} & & & & \\
\textbf{Short} & 517 & 533 & –16 & \(p < .01\) \\
\textbf{Long} & 546 & 544 & +2 & n.s. \\
\hline
\textbf{ERP amplitude (µV) at occipital scalp sites (OL, OR, T5, T6)} & & & & \\
\textbf{90 to 140 ms posttarget: P1 component} & & & & \\
\textbf{Short} & 0.23 & –0.19 & +0.42 & \(p < .01\) \\
\textbf{Long} & 0.88 & 1.21 & –0.33 & \(p < .01\) \\
\hline
\textbf{ERP amplitude (µV) at central-parietal scalp sites (Pz, Cz)} & & & & \\
\textbf{250 to 500 ms posttarget: P300 component} & & & & \\
\textbf{Short} & 4.16 & 3.14 & +1.02 & \(p < .001\) \\
\textbf{Long} & 4.95 & 4.80 & +0.15 & n.s. \\
\hline
\end{tabular}
\caption{Mean values and summary statistics for the behavioral and electrophysiological measures.}
\end{table}
order to track the fates of cued and uncued targets, we evaluated longer latency ERP components known to reflect higher order aspects of target processing. One such component of the ERP that has been used in conjunction with RTs to examine human information processing is the P300 component (latency of 250 to 500 ms, maximal over central and central-parietal scalp sites; e.g., Duncan-Johnson & Donchin, 1982). The amplitude of the P300 is not directly tied to RTs, but instead indexes aspects of information processing such as expectancy and perceived stimulus relevance. The P300 is typically larger to attended than unattended targets, and larger to infrequent than frequent stimuli (Donchin, 1981). In the present study, the P300 was enlarged for cued-location targets, but only at the short cue-to-target ISIs; the P300 to cued and uncued targets did not differ at long ISIs (Fig. 4). This pattern indicates another key function of reflexive orienting. At short ISIs, reflexive attention not only facilitates processing in sensory cortex (i.e., the P1 modulation), but also leads cued-location stimuli to be treated as more relevant at higher stages of stimulus evaluation. Additionally, because there was no P300 difference observed at long ISIs, it is possible to conclude that subjects in the present task were not invoking voluntary orienting toward the task-irrelevant cue. If they had done so, a difference in P300 amplitude would have been expected at the long ISIs as well. The present findings suggest that another result of reflexive orienting is that very shortly after a sensory stimulus, the stimulated location is briefly tagged as being more relevant than other locations in the environment.

**CONCLUSION**

These findings provide evidence that when attention is reflexively attracted by a sensory event, cortical visual processing is subsequently altered. Furthermore, these early modulations occur at the same stage of visual processing as the earliest effects of voluntary attention. Although these two types of attention proceed with different time courses, both voluntary focusing of attention and reflexive shifts of attention are able to modulate early sensory processing of visual stimuli. In addition, reflexive orienting transiently affects the perceived importance of objects in the environment, as at short ISIs, stimuli at locations that have just captured attention are processed as if they are more relevant than events elsewhere. These findings suggest that perceptions of the visual world can be affected not only by voluntary, top-down processes, but also by bottom-up, reflexive mechanisms that influence perceptual processing as early as the extrastriate cortex in humans.
Fig. 4. Scalp topographic voltage maps, referenced to the right mastoid, in the latency range of the P300 component (250 to 500 ms poststimulus), in a rear view of the head. Event-related potentials (ERPs) elicited by left-visual-field targets are presented in the two panels on the left; ERPs elicited by right-visual-field targets are presented in the two right panels. The top panels show results for short interstimulus intervals (ISIs), and the bottom panels show results for long ISIs. The small black dots on each map indicate the locations of the electrodes.

Acknowledgments—This work was supported by the National Science Foundation, the National Institutes of Health, the National Institute of Mental Health, the Human Frontiers Science Program, and the James S. McDonnell Foundation. We thank M. Woldorf and C. Kussmaul for their assistance and advice, and we are grateful to M. Posner, S. Hillyard, S. Yantis, S. Luck, M. Girelli, and J. Schaal for their helpful comments on earlier versions of this manuscript. These data were presented in preliminary form at the November 1995 meeting of the Society for Neuroscience in San Diego (Hopfinger & Mangun, 1995).

REFERENCES


Eason, R.G. (1981). Scalp topographic voltage maps, referenced to the right mastoid, in the latency range of the P300 component (250 to 500 ms poststimulus), in a rear view of the head. Event-related potentials (ERPs) elicited by left-visual-field targets are presented in the two panels on the left; ERPs elicited by right-visual-field targets are presented in the two right panels. The top panels show results for short interstimulus intervals (ISIs), and the bottom panels show results for long ISIs. The small black dots on each map indicate the locations of the electrodes.

Acknowledgments—This work was supported by the National Science Foundation, the National Institutes of Health, the National Institute of Mental Health, the Human Frontiers Science Program, and the James S. McDonnell Foundation. We thank M. Woldorf and C. Kussmaul for their assistance and advice, and we are grateful to M. Posner, S. Hillyard, S. Yantis, S. Luck, M. Girelli, and J. Schaal for their helpful comments on earlier versions of this manuscript. These data were presented in preliminary form at the November 1995 meeting of the Society for Neuroscience in San Diego (Hopfinger & Mangun, 1995).

REFERENCES


Electrophysiology of Reflexive Attention


(Received 7/7/97; Accepted 2/11/98)

**APPENDIX**

Comparison of the cue and target ERPs in Figure A1 demonstrates that the cue was indeed a subtle stimulus, which produced a relatively small overall ERP compared with that produced by the target stimulus. Nevertheless, inspection of Figure A1a shows that the stimulus-evoked ERPs before Adjar filtering likely contained overlapping activity from adjacent stimuli. The Adjar procedure was determined to be effective at removing this overlapping activity on the basis of three observations. First, the iterative estimation and subtraction procedure converged after 10 iterations, indicating that a stable solution had been obtained (e.g., Woldorff, 1993). Second, the prestimulus baseline of the Adjar-corrected target ERPs is flat (see Fig. A1b, right panel), indicating that the Adjar procedure did indeed remove systematic overlap from preceding cue stimuli. Finally, the effects of interest were not created by the Adjar procedure: Comparison of the target waveforms during the likely period of greatest overlap shows that the P1 attention effect is present in both the before-Adjar and the after-Adjar waveforms (right panels, Figs. A1a and A1b, respectively).

In addition, although not shown here, the estimate of the overlap of the cue responses onto the target responses (convolution of the cue ERPs with the cue-to-target ISI distribution) showed that the difference between left and right cues had a different polarity of effects for the contralateral and ipsilateral occipital electrodes. However, the corrected target ERP attention effects did not display this pattern; in contrast, the polarity of the effect was in the same direction for the Adjar-corrected waveforms over contralateral and ipsilateral occipital scalp (as is typically seen for the P1 attention effect during voluntary attention). Thus, these effects could not have been produced solely by the Adjar process of subtracting the residual overlap of the cue responses.

Figure A1c shows the ERP from the cue on catch trials. The cue ERP on catch trials cannot be assumed to be a perfect model for a cue that is followed shortly in time by another event. There were only one half the number of catch-trial cues as short-ISI cues, and therefore, the catch-trial cues contain more background noise. In addition, it is likely there are real differences between a cue that is not followed by another stimulus (catch trial) and one that is. In fact, that is the reason why we did not simply subtract ERPs for catch trials from the target ERPs, but rather used the actual data to obtain the cue response. Nevertheless, the catch-trial ERP does give a rough estimate of the ERP to cue stimuli in general, without the overlapping processing from target stimuli. Comparison of Figures A1b (left column) and A1c shows that the Adjar-corrected cue ERP does in fact capture the form of the ERP for the cue alone (catch-trial cue), providing further validation that the procedure was not fabricating responses, but was actually finding true responses amid the overlapping activity.

Although the Adjar correction was applied to both ISIs in the data presented in the Results section, only the short-ISI data are shown in Figure A1. This is because the most severe overlap of cue responses onto target responses would be expected to occur when the two events are closest to each other in time. Indeed, the overlap of cue responses onto target responses was minimal at the long ISI, because the waveform of the cue ERP had largely returned to baseline by the time the long-ISI target was presented (see Fig. A1c).
Fig. A1. Illustration of the effectiveness of the Adjar procedure. Shown in (a) are event-related potentials (ERPs) to cues and targets with short interstimulus intervals (ISIs), before being filtered with the Adjar procedure. For each ERP plot, the vertical axis indicates the onset of the stimulus. The gray boxes indicate the likely region where overlapping components could be most severely contaminating the ERP. The time ranges over which adjacent stimuli were presented are indicated above the gray boxes. Shown in (b) are ERPs to cues and targets with short ISIs, after being filtered with the Adjar procedure. The ERP to the cue stimulus on catch trials is shown in (c). All ERPs shown here are averaged over occipital scalp sites contralateral to the visual field of stimulus presentation (sites OL and OR).