

Bridging the Gap between Monkey Neurophysiology and Human Perception: An Ambiguity Resolution Theory of Visual Selective Attention

Steven J. Luck, Massimo Girelli, Michele T. McDermott, and Michelle A. Ford

University of Iowa

When the visual system must process multiple objects simultaneously, as in the visual search paradigm, the neural coding of individual objects can become ambiguous due to the visual system's extensive use of coarse coding and distributed representations. Here we propose that the primary role of visual selective attention within the ventral object recognition pathway is to resolve these ambiguities. We begin by reviewing previous studies of the effects of attention on neural responses in monkeys, which provide the basis for this hypothesis, and then describe a new set of experiments showing that similar attentional mechanisms operate in the human brain. In these new experiments, event-related potentials (ERPs) were recorded from normal human observers while they performed tasks analogous to those used previously in monkeys. The central finding was that an attention-related ERP wave called the "N2pc component" was present under the same conditions that led to attentional modulations of neural responses in monkey visual cortex. These human electrophysiological results provide a bridge between cognitive-level theories of visual attention and the behavior of individual neurons in visual cortex. © 1997 Academic Press

Virtually all theories of selective attention assume that the primary function of selecting some stimuli and rejecting others is to allow the brain to process the selected stimuli more efficiently than would be possible if the brain were to process all of the stimuli simultaneously. Theories often differ, however, in explaining why processing is inefficient when many stimuli are processed simultaneously. The most common type of explanation assumes that the brain contains a fixed-size pool of some sort of "resource" that is divided among the stimuli that are being processed at a given moment, such that fewer resources are available for any given stimulus when many stimuli are processed simultaneously.

This study was supported by Grant 95-38 from the McDonnell–Pew Program in Cognitive Neuroscience. We thank Leonardo Chelazzi, Robert Desimone, and Steven A. Hillyard for many stimulating discussions that influenced the ideas expressed in this article. We also thank Kyle Cave and Anne Treisman for providing comments on an earlier version of the manuscript.

Address correspondence and reprint requests to Steven J. Luck, Department of Psychology, University of Iowa, 11 Seashore Hall E, Iowa City, IA 52242-1407. E-mail: steven-luck@uiowa.edu.

Although resource theories may be very elaborate, they rarely specify exactly why a given cognitive process cannot operate on multiple stimuli simultaneously without suffering from interference. In other words, most theories do not provide a detailed description of the computational problem that selective attention solves. A notable exception to this generalization is Treisman's feature integration theory (Treisman, 1988; Treisman & Gelade, 1980).¹ This theory begins with the proposal that the visual system initially decomposes the visual world into maps of simple features such as color, orientation, and direction of motion. The feature extraction process is thought to occur independently for different points in space and for different types of features; in other words, feature analysis is proposed to occur in parallel and without capacity limitations. These statements are generally supported by psychophysical and neurophysiological studies of visual processing, and are not very controversial. However, feature integration theory goes further, postulating that the separately coded features of a given object cannot be accurately localized or combined unless attention is focused on that object. Without attention, this theory proposes, an observer can obtain only a coarse estimate of the pooled activity across an entire feature map and has no access to the spatial arrangement of activity in that map or to the relationships between features in different maps. For example, an observer may be able to determine that there is some activity in the map corresponding to the feature "red" even without the use of attention, but that observer would be unable to determine which of several objects is red without focusing attention onto the red object. Similarly, an observer may be able to determine that an array contains something horizontal and something green without attention, but would be unable to determine whether these features belonged to a single object or to different objects. Attention is thus proposed to serve the computational role of combining together the separate features of an object so that features from different objects will not be mistakenly combined together.

In the present article, we will describe a theory of attention that provides an explanation for this role of attention in terms of the anatomy and physiology of the nervous system (for an earlier discussion of this model, see Luck & Beach, in press). At its core, this theory proposes that the primary computational role of selective attention in the visual system is to resolve ambiguities in neural coding that arise when multiple stimuli are processed simultaneously, and we therefore refer to this model as the *ambiguity resolution theory*. This theory is very much in the same spirit as feature integration theory (see especially Treisman, 1996), but it is much more directly concerned with the visual system's anatomy and physiology and is therefore at a somewhat lower level of analysis. This theory is also closely related to the biased competition

¹ Similar models have also been proposed by other investigators (see, e.g., Cohen & Ivry, 1989; Cohen & Ivry, 1991; Prinzmetal *et al.*, 1995).

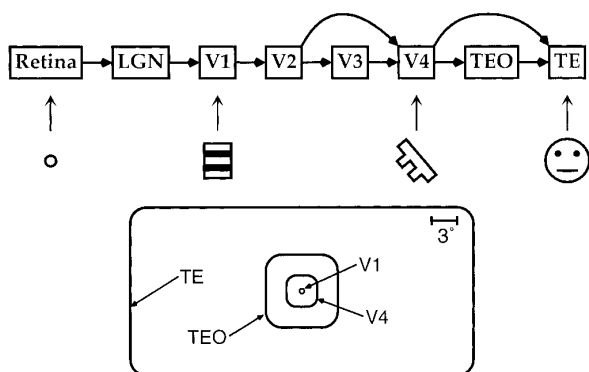


FIG. 1. Simplified diagram of the flow of information into and through the ventral object recognition pathway of the macaque monkey. The arrows represent major feedforward connections between areas; for every feedforward connection shown, there is also a feedback connection. Representative stimuli for which typical neurons are selective are shown for several areas. Typical receptive field sizes are also shown, but it should be noted that receptive field sizes vary as a function of eccentricity and also between different neurons at the same eccentricity.

model that has been developed by Desimone, Duncan, and their colleagues (Desimone & Duncan, 1995; Desimone, Wessinger, Thomas, & Schneider, 1990; Duncan, Humphreys, & Ward, in press).

Key Attributes of the Primate Visual System

The ambiguity resolution theory's starting point is the architecture of the ventral object recognition pathway of the primate visual system (a simplified diagram of this pathway is shown in Fig. 1). After passing through the retina and the lateral geniculate nucleus, visual information enters the cortex in area V1 and then travels forward through areas V2, V3, and V4 of the occipital lobe and into areas TEO and TE of the inferior temporal lobe. As information travels along this pathway, two obvious changes in neural response properties can be observed. First, there is a substantial increase in the complexity of the features coded by individual neurons. Second, there is a dramatic increase in receptive field sizes, which is accompanied by a decrease in the spatial resolution of individual neurons. These changes may be functionally related, because the increase in featural complexity in higher-level areas leads to an enormous increase in the number of features that must be represented, and an implausibly high number of neurons would be required to code each of these features separately for each small location in space.

The changes in neural response properties between area V1 and area TE appear to be related to the general problem of perceptual constancy. Specifically, the image cast on the retina by a given object at a single moment varies according to temporary factors such as the position and distance of the object, the position and composition of the source of illumination, and the presence

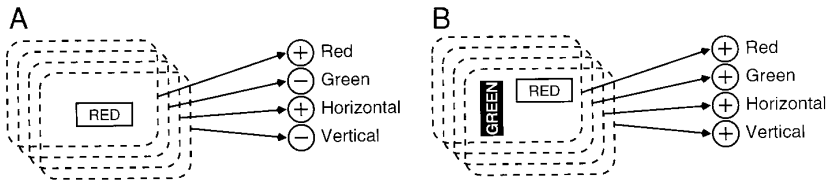


FIG. 2. Example of the ambiguity in neural coding that can arise in an area such as V4 when multiple stimuli are present inside the receptive field of a given neuron. The regions enclosed in broken lines represent the receptive fields of four hypothetical V4 neurons, and the plus and minus symbols inside the circles represent whether a given neuron would respond (plus) or not respond (minus) to the stimulus or stimuli shown inside the receptive fields.

of occluding objects; the visual system must be able to recognize an object in a consistent manner despite these great variations in the retinal image. The transformations that occur between area V1 and area TE appear to convert the temporary image cast on the retina into a representation that is relatively independent of factors such as viewpoint and illumination. An important implication of this organizational framework is that the information coded in area V1 is not directly useful for most real-world visual tasks and is probably not directly accessible for performing most visual tasks (Crick & Koch, 1995; He, Cavanagh, & Intriligator, 1996).

Another important characteristic of the visual system is that objects are coded by means of distributed representations. Specifically, although the features coded by individual neurons become progressively more complex at higher levels of the visual system, complex objects are not coded by the activation of highly specialized “grandmother” cells, but are instead represented by the parallel activation of large numbers of neurons that code the object’s features (see, e.g., Fujita, Tanaka, Ito, & Cheng, 1992; Komatsu & Ideura, 1993; Llinás & Paré, 1996). Distributed representations also appear to be used for making fine discriminations of simple features, which are coded by the pattern of activity over many coarsely tuned neurons rather than a few narrowly tuned neurons. For example, discriminating between a 45° line and a 50° line does not require the presence of individual neurons that respond selectively to a range of 5° or less, but can instead be accomplished by comparing the pattern of activity over several neurons, each of which is broadly tuned to a different range of orientations.

The Ambiguity Resolution Theory

Although the coarsely coded, distributed representations used by the visual system are an efficient means of coding information, they may lead to substantial computational problems when multiple objects are presented simultaneously, as in visual search experiments and most real-world visual tasks. This is illustrated in Fig. 2, which shows the hypothetical outputs of a set of four

idealized neurons that have overlapping receptive fields, but are selective for different features. When a single object is presented to these neurons, as in Fig. 2A, the neural outputs appropriately code the features of that object. However, when two objects are presented simultaneously, as in Fig. 2B, all four neurons are activated and the representation becomes ambiguous, because it is not possible to determine from these neurons which features belong together as attributes of a single object. In Treisman's terms, the presence of multiple stimuli within the receptive fields of these cells may lead to "illusory conjunctions" (see Treisman & Schmidt, 1982). This is one important way in which neural representations at the higher stages of cortex can become ambiguous when multiple stimuli are presented simultaneously, and it is a special case of the more general "binding problem" that arises at several levels of the nervous system (Damasio, 1989; Luck & Beach, in press; Singer & Gray, 1995; Treisman, 1996).

When only one stimulus falls inside a neuron's receptive field, as in Fig. 2A, the neuron's output unambiguously signals the features of that stimulus; the binding problem therefore primarily arises when multiple stimuli are presented simultaneously inside the receptive field of a given neuron. For example, feature binding is not typically a problem in area V1, where receptive fields are small and different features can be linked together by virtue of being represented by neurons that code the same small portion of space. In addition, this problem will not arise at an intermediate-level area such as V4 if the stimuli are placed so far apart that only a single stimulus falls inside the receptive field of a given neuron. However, binding will almost always be problematic in area TE, where receptive fields are so large that a given neuron almost always receives inputs from multiple stimuli.

A possible solution to the binding problem was suggested by Moran and Desimone (1985), who recorded single-unit responses from areas V1, V4, and TE in macaque monkeys while they performed an attention task. When an attended stimulus and an ignored stimulus were presented simultaneously inside the receptive field of the neuron being recorded, the neuron's output was found to reflect the features of the attended stimulus but not the features of the ignored stimulus. For example, a neuron that was selective for red stimuli would fire at a high rate when the attended-location stimulus was red and the ignored-location stimulus was green, but would not fire when the ignored-location stimulus was red and the attended-location stimulus was green. However, attention had no effect when one stimulus was presented inside the receptive field of the cell being recorded and the other stimulus was presented outside the receptive field. This surprising finding may reflect the fact that a neuron's output is unambiguous when only a single stimulus falls inside its receptive field. Moran and Desimone also found that no attention effects could be obtained in area V1, where the receptive fields were always too small to contain both an attended stimulus and an ignored stimulus and the neural responses were presumably unambiguous.

These findings may help to explain a common discrepancy between theories of attention and our phenomenal impression of the visual world. Specifically, although many theories of attention posit that stimuli at ignored locations are suppressed when attention is focused on a specific object, people do not generally have the phenomenal impression that the rest of the visual input disappears when attention is focused on a particular object. This discrepancy may be explained by the finding that only those neurons whose receptive fields contain both an attended stimulus and an ignored stimulus exhibit attentional suppression, which leaves many unsuppressed neurons available to provide a subjective impression of the unattended portions of the visual input.

We have now provided enough background to allow a summary of the main elements of the ambiguity resolution theory. First, this theory proposes that the coding of features by neurons in visual cortex becomes ambiguous when multiple stimuli are presented simultaneously, especially when multiple items fall within the receptive field of a given neuron. Second, although ambiguous coding is assumed to be minimal in low-level areas with small receptive fields, such as V1, we assume that the information coded in these low-level areas is not directly available for report. This assumption is supported by the lack of anatomical projections from area V1 to higher stages of cortex (Van Essen, 1985), and is motivated computationally by the fact that the coding of information in V1 is neither viewpoint nor illumination independent, making it inappropriate for direct use in most real-world visual tasks. Third, we propose that the primary role of selective attention is to resolve ambiguous coding in the higher-level areas. At present, we assume that this is accomplished by a mechanism that simply suppresses inputs from unattended objects such that a given neuron unambiguously codes the features of one object at a given moment (for more sophisticated attentional mechanisms, see Niebur, Koch, & Rosin, 1993; Singer & Gray, 1995). We further assume that the degree of attentional suppression exhibited by a given neuron depends on a combination of top-down control signals and locally determined information about the presence of multiple inputs to that neuron.²

Up to this point, we have mainly considered the effect of ambiguous coding on the visual system's ability to combine together the different features of an object. However, fine discriminations of simple features may require information from several coarsely coded neurons to be combined, which may also lead to ambiguity when multiple stimuli are presented simultaneously. Thus, the ambiguity resolution theory predicts that attention may be necessary for

² For the sake of simplicity, this account of the ambiguity resolution theory neglects the important contribution of the parietal lobes to visual selective attention. There are two likely roles that the parietal lobes may play. First, they may provide the top-down signals that indicate where attention should be directed within ventral-pathway areas such as V4 and TE. Second, they may play a role in constructing "object tokens" that collect together the features that remain after attention has been used to suppress the features that belong to distractor items.

the accurate discrimination of simple features when the target is surrounded by competing distractor items, especially when the target and distractor items are similar (for additional discussion of this issue, see Treisman, 1996; Treisman & Gormican, 1988). Similarly, the large RFs in areas such as V4 and TE may make it difficult to report the precise location of a simple feature in the absence of focused attention.

Empirical Evidence: Performance in Psychophysical Tasks

The ambiguity resolution theory can readily explain the main psychophysical results that have been taken as evidence for Treisman's feature integration theory. First, the ambiguity resolution theory predicts that, in the absence of attention, observers may accidentally combine features that belong to different objects, resulting in the perception of illusory conjunctions (see Treisman & Schmidt, 1982). Second, the search for a conjunction target will typically require the serial application of attention to individual items in order to suppress features from other items that might create illusory conjunctions (although the possibility of guided search may sometimes obviate the need for a serial search—see Wolfe, Cave, & Franzel, 1989). Third, attention should be unnecessary for detecting highly salient features, because feature-selective neurons should unambiguously code the presence of a given feature even when multiple stimuli are present inside the receptive field (see Treisman & Gormican, 1988). For example, the presence of the feature "red" in Fig. 2B could be determined by assessing the extent to which the red-selective neuron is active. Fourth, because location is coded rather coarsely in higher-level areas such as V4 and TE (in the absence of attention), observers may be able to report the presence or absence of a feature without being able to report the feature's precise location (see Treisman & Gelade, 1980).

Although the ambiguity resolution theory makes many of the same predictions as feature integration theory, it also makes some diverging predictions. First, some coarse information about stimulus position is available in the absence of attention even at the highest levels of the visual system (Ito, Tamura, Fujita, & Tanaka, 1995). As a result, the ambiguity resolution theory would predict that features can be coarsely localized in the absence of attention, whereas feature integration theory posits that no veridical information about feature locations is available without focused attention. Consistent with this prediction, Johnston and Pashler (1990) found that features were at least coarsely localized whenever they were identified correctly. In addition, neural representations become ambiguous only when multiple stimuli are present within a given receptive field, so the ambiguity resolution theory would predict that illusory conjunctions should occur only when the stimuli are close together; consequently, an observer should be able to discriminate conjunctively defined targets without attention in arrays containing a relatively small number of widely dispersed stimuli. In support of these predictions, Cohen and Ivry (1989) found that illusory conjunctions occurred primarily between items that

were close enough to fall within a single receptive field in higher-level areas such as V4 and TE, and Cohen and Ivry (1991) found that visual search for conjunctions was more attention demanding when the items were presented in close proximity to each other (see also Prinzmetal, Henderson, & Ivry, 1995). Similarly, Sohn, Liederman, and Reinitz (1996) recently reported that although illusory conjunctions occur frequently between objects within the same hemifield, they are eliminated when the objects are presented in opposite hemifields. This may be explained by the fact that excitatory receptive fields in area V4 rarely extend more than 1° across the vertical meridian, but may extend over several degrees within a hemifield (Desimone, Moran, Schein, & Mishkin, 1993; Desimone & Schein, 1987; Gattass, Sousa, & Gross, 1988).

The ambiguity resolution theory proposes that ambiguous coding is particularly likely when a distractor item is near a target item, because distant items are more likely to be coded by separate populations of neurons. This leads to the prediction that attentional suppression should be greatest for locations near the target and should be minimal at distant locations. This prediction appears to conflict with many studies that have found an opposite pattern, with increasing rather than decreasing impairments in performance as a function of distance from the attended location (e.g., Downing & Pinker, 1985; LaBerge & Brown, 1989; Shulman, Wilson, & Sheehy, 1985). However, these effects may reflect contributions from multiple attentional mechanisms, some of which may produce local inhibition and others of which may show gradients in resource allocation. Indeed, the apparent spatial distribution of attention has been shown to depend on whether accuracy or reaction time is the dependent variable (Handy, Kingstone, & Mangun, in press). The common finding of decreasing rather than increasing performance as a function of distance from the target should not, therefore, be viewed as problematic for the ambiguity resolution theory. In addition, recent studies have shown that it is possible to obtain the pattern predicted by the ambiguity resolution theory under some conditions (Cave & Zimmerman, submitted; Suzuki & Cavanagh, in press).

Empirical Evidence: Single-Unit Recordings

The single-unit recordings of Moran and Desimone (1985) provided the original impetus for the ambiguity resolution theory, and we have recently replicated and extended these results in a set of recordings from areas V1, V2, and V4 (Luck, Chelazzi, Hillyard, & Desimone, 1997). In this new study, consistent attentional modulations were again obtained in area V4 only when both the attended and ignored stimuli were located inside the receptive field of the neuron being recorded. We also replicated Moran and Desimone's finding of no attention effects in area V1, where receptive fields were too small to contain both attended and ignored stimuli. Similarly, the receptive fields of most neurons in V2 were also too small to contain both stimuli, and attention effects were absent in these neurons as well. However, we were able to find a subset of V2 neurons with relatively large receptive fields, and

in these neurons we found that attention effects could be observed when both the attended and ignored stimuli were located inside the receptive field. We also demonstrated that the attention effects in area V4 were larger when the attended and ignored stimuli were presented simultaneously rather than sequentially, presumably because the potential for ambiguous coding is greater under conditions of simultaneous presentation.

Additional supporting evidence has also been reported by Chelazzi and his colleagues (Chelazzi & Desimone, 1994; Chelazzi, Miller, Duncan, & Desimone, 1993), who recorded from neurons in areas V4 and TE³ in a modified visual search task with two items in each stimulus array. The stimuli were complex, multicolored images for some neurons and simple color patches for others, and they were chosen such that one item in a given test array was effective in eliciting a response from the neuron being recorded and the other item was ineffective. For example, a typical neuron in area V4 might respond to green stimuli and not to red stimuli, and for this neuron the test array would consist of a green patch and a red patch. If attention operates to filter information from ignored stimuli, then the response of a neuron to two stimuli presented simultaneously should be similar to the neuron's response to the attended stimulus presented alone. Consequently, when green and red stimuli are presented simultaneously inside the receptive field of a neuron that normally responds only to green, a large response would be expected when the green item is attended and a small response would be expected when a red item is attended. This logic has been used in several previous single-unit studies of visual attention (e.g., Luck *et al.*, 1997; Moran & Desimone, 1985; Treue & Maunsell, 1996).

Large attention effects were obtained by Chelazzi *et al.* (1993, 1994) under the conditions that would be predicted to yield the largest effects of attention: (a) both target and nontarget items were inside the receptive field of the cell being recorded; (b) the target and nontarget items were complex stimuli that could not easily be discriminated on the basis of a simple, salient feature; and (c) the observers were required to make a saccade to the target, which required precise localization of the target. As shown in Fig. 3A, the search array initially elicited a large response that simply reflected the presence of an effective stimulus inside the receptive field and did not vary as a function of whether the effective stimulus was the target or the distractor. Beginning approximately 175 ms after the onset of the test array, however, the responses began to diverge as a function of which stimulus was the target. Specifically, if the effective stimulus was the target, the neuron continued to fire at a high rate, but if the ineffective stimulus was the target, the neuron's output became

³ The results that we attribute to area TE also included recordings from sites in perirhinal cortex, which is just medial to area TE. The neurons in this area have similar receptive field and feature selectivity properties to those in area TE, and we have therefore used the label "TE" for the sake of simplicity in describing these results.

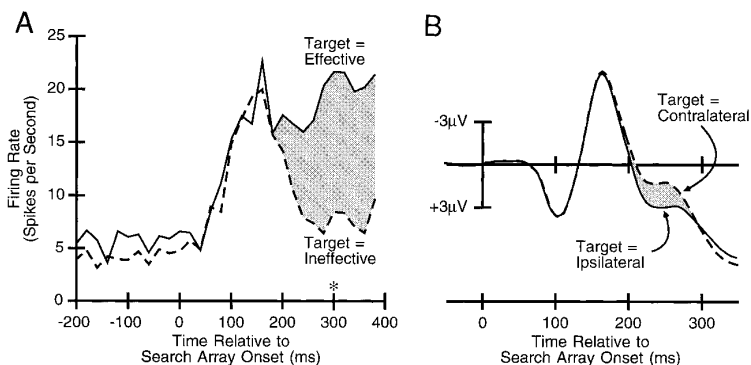


FIG. 3. (A) Poststimulus histograms averaged over a group of 22 neurons in area TE in the study of Chelazzi *et al.* (1993). These data were recorded while the monkeys performed a task in which they were required to make a saccade toward a complex visual search target. Time zero represents the onset of the test array, and the asterisk represents mean saccade onset latency. The effect of attention is indicated by the shaded area. (B) Example of the N2pc component (indicated by the shaded region), which is defined as a negative-going potential for trials on which the target is contralateral to the hemisphere of the recording site relative to trials on which the target is ipsilateral. Note that, by convention, negative is plotted upward.

highly suppressed. If this pattern of activity is extrapolated to the entire local population of neurons, these data suggest that only the neurons that coded the features of the target stimulus were active by approximately 250 ms after the onset of the stimulus array. This is exactly what would be expected if attention serves to minimize ambiguous coding by allowing only one item to be processed at a given moment (for a detailed model of the mechanisms underlying these results, see Usher & Niebur, 1996).

The magnitude of the attention effect in the experiments of Chelazzi *et al.* (1993, 1994) depended on the nature of the stimuli and the response. The large attention effect shown in Fig. 3A was obtained when the target was a complex stimulus and the monkey was required to foveate the target; a somewhat smaller modulation was observed when the monkey made a lever-release response to the target rather than an eye movement. When the target was a simple color patch, the attention effect was greatly reduced for the foveate task and was completely eliminated for the lever task. Thus, attention effects were greater for complex stimuli than for simple stimuli and were also greater when the monkey was required to foveate the target (presumably because this required target localization as well as detection). It should also be noted that, as in the studies of Moran and Desimone (1985) and Luck *et al.* (1997), these attentional modulations were greatly reduced in area V4 when only one of the two items in the test array was inside the receptive field. In addition, although both items were always inside the receptive field in area TE, the attentional modulations were reduced or eliminated when one item was in

the ipsilateral hemifield and the other was in the contralateral hemifield, presumably because the responses of TE neurons are typically dominated by the features of the contralateral stimulus when ipsilateral and contralateral stimuli are presented simultaneously. These results are in accord with the ambiguity resolution theory, because ambiguous coding should be most likely to occur and most problematic for accurate task performance when: (a) multiple items compete for access to the receptive field of the cell being recorded; (b) a conjunction of features must be reported; and (c) features must be precisely localized.

Although the single-unit recording studies described above found large and consistent attentional modulations only when both the attended and ignored stimuli were presented simultaneously inside the receptive field, it should not be concluded that no effects of attention can be observed when only a single stimulus is present inside the receptive field. Indeed, several single-unit studies of spatial attention have found modest attentional modulations with only a single stimulus inside the receptive field (e.g., Connor, Gallant, Preddie, & Van Essen, 1996; Motter, 1993; Spitzer, Desimone, & Moran, 1988; Treue & Maunsell, 1996). However, all known studies that have compared a condition with only one stimulus inside the receptive field to a condition with multiple stimuli inside the receptive field have found much larger attention effects in the latter case (Chelazzi & Desimone, 1994; Luck *et al.*, 1997; Moran & Desimone, 1985; Treue & Maunsell, 1996). Interestingly, the attentional modulations observed with only a single stimulus inside the receptive field were present primarily when difficult perceptual tasks were used, and these effects may reflect a different computational problem than the one addressed by the ambiguity resolution theory.

Overview of the Present Study

As discussed above, the ambiguity resolution theory is consistent with many psychophysical studies of visual search performance and also with several studies of the effects of attention on individual neurons in macaque visual cortex. However, these two sources of evidence were acquired from different species under very different conditions, and they represent different levels of analysis. Faced with these substantial differences, it is important to ask whether these two sources of evidence can be combined in a straightforward manner. For example, although the human and macaque visual systems are remarkably similar up to the level of area V1 (De Valois, Morgan, & Snodderly, 1974), it is not yet known whether higher level areas such as V4 and TE have the same properties in both species or even if the human and monkey brains contain the same visual areas. In addition, human psychophysical studies of attention typically rely on naive observers who participate in a single session, whereas monkey electrophysiological data are obtained from trained observers who have received months or years of practice. Extended training has been shown to have a large impact on the performance of visual

tasks (see, e.g., Schneider & Shiffrin, 1977; Treisman, Vieira, & Hayes, 1992; Wolfe *et al.*, 1989), and this makes it difficult to compare results from naive and trained observers. The purpose of the present study, therefore, was to attempt to “bridge the gap” between monkey electrophysiological studies and human psychophysical studies by testing whether the pattern of attentional modulations observed during visual search in monkeys would also be observed in humans.

In this study, we used an electrophysiological measure of attention that appears to reflect the same type of attentional processing that is indexed by the single-unit recordings, namely the “N2pc” component of the event-related potential (ERP) waveform. As illustrated in Fig. 3B, the N2pc component occurs in the latency range of the N2 family of components (200–300 ms poststimulus) and is distinguished by its unique posterior contralateral scalp distribution (“N2pc” is an abbreviation of “N2-posterior-contralateral”). Specifically, the N2pc is observed as a negative-going voltage deflection at scalp sites contralateral to the position of a target that is embedded in a bilateral stimulus array. This contralaterality relative to the target distinguishes the N2pc component from the early sensory components, which typically have bilateral scalp distributions for bilateral stimulus arrays, and also distinguishes it from the late cognitive components, which generally have bilateral distributions for both unilateral and bilateral stimuli. There is substantial evidence linking the N2pc component with the focusing of attention (Luck & Hillyard, 1994a, 1994b).

Different mechanisms of attention may operate at different stages of processing, and the N2pc component was chosen as a measure of attention in the present study because several sources of evidence suggest that it may reflect the same type of attentional modulation that was observed in the single-unit studies of Chelazzi *et al.* (1993, 1994). First, both the N2pc component and the single-unit modulations appear to reflect the selection of the target item and the suppression of distractor items during visual search tasks. Second, both of these effects begin at approximately 175 ms poststimulus. Third, the N2pc component is strongly contralateral, which is consistent with the fact that neurons in V4 respond almost exclusively to contralateral stimuli and the fact that the responses of neurons in area TE are typically dominated by contralateral stimuli. For these reasons, the N2pc component was chosen as a means of assessing whether the operation of attention in visual cortex is similar for humans and monkeys.

EXPERIMENTS 1 AND 2

Experiments 1 and 2 tested three hypotheses based on the single-unit recordings of Chelazzi *et al.* (1993, 1994). First, we predicted that the need for focused attention would be greater—and hence the N2pc component would be larger—when the observers were required to perform a complex conjunction discrimination task rather than a simple feature detection task. Second, we

predicted that the need for focused attention would be greater—and the N2pc would be larger—when the visual search display contained a distractor item in close proximity to the target item such that both the target and the distractor would fit within a typical V4 receptive field. Third, based on the geometry and topographic mapping of the presumed human homolog of macaque area V4, we predicted that the N2pc component would be substantially larger when the items in the search array were in the lower visual field rather than in the upper visual field. This prediction is based on the assumptions that (a) the representation of the lower and upper visual fields are located in the dorsal and ventral portions of human area V4, respectively, as they are in macaque V4, and (b) the locations of the dorsal and ventral portions of area V4 are roughly similar in both species (both of which appear to be a plausible assumptions—see Clarke & Miklossy, 1990; Sereno, Dale, Reppas, Kwong, Belliveau, Brady, Rosen, & Tootell, 1995). If these assumed homologies are valid, then the lower visual field representation in V4 should be located in dorsal-lateral occipital cortex, directly beneath the occipital-temporal electrode sites, and activation of the lower visual field should therefore create a large amplitude ERP deflection at these sites. In contrast, the upper visual field representation in V4 should be located on the ventral surface of the occipital lobe (see, e.g., Fig. 2 of Sereno *et al.*, 1995), which is much more distant from the scalp and would therefore be expected to produce smaller ERP amplitudes in recordings from typical scalp electrodes. In addition, the left and right hemisphere locations of the ventral portion of area V4 are relatively close together and are oriented such that electrical dipoles in these areas would be approximately parallel to each other, factors that together would diminish any differences in amplitude between contralateral and ipsilateral amplitudes at the scalp.

To test these hypotheses, two separate experiments were conducted, as illustrated in Fig. 4. The participants performed a relatively complex conjunction discrimination task in Experiment 1 and a simple feature detection task in Experiment 2. In both experiments, the target was sometimes crowded by a nearby distractor and sometimes presented in isolation, and the stimuli were presented in the upper visual field on some trials and the lower visual field on others. Because ERPs are highly sensitive to subtle changes in stimulus parameters, the stimuli and task had to be designed so that it would be possible to compare ERPs elicited by exactly the same stimuli, but under different psychological conditions (i.e., different instructions). This constraint was satisfied by using stimulus arrays containing two colored items, one in the left visual field (LVF) and one in the right visual field (RVF), and by asking the observers to discriminate one color in some trial blocks and the other color in other trial blocks. For example, a stimulus array might contain a green item in the LVF and a red item in the RVF, and the participants would discriminate the green item in some trial blocks and the red item in others. The ERP elicited by this stimulus array could therefore be compared for the

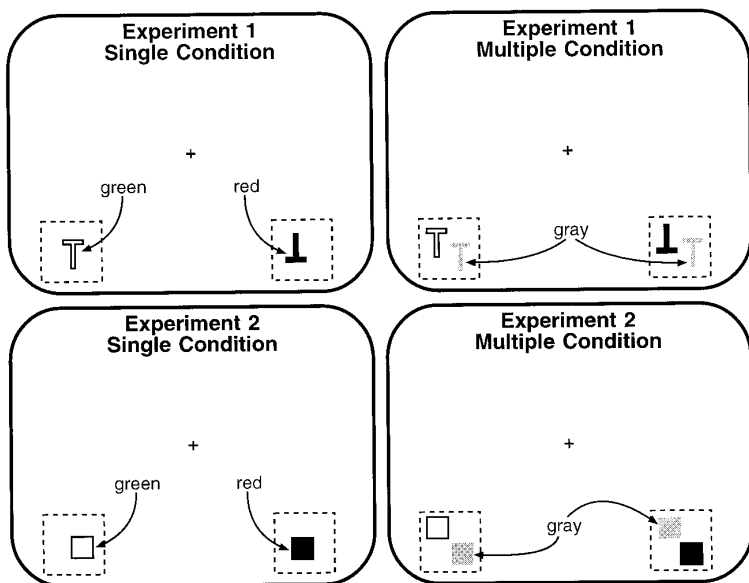


FIG. 4. Examples of the stimuli used in Experiments 1 and 2. In both experiments, the search array items were presented at random locations within the rectangular areas shown in broken lines (these rectangles were not visible to the participants). These areas were $1.4^\circ \times 1.4^\circ$ in size, and were centered 2.7° below and 2.7° to the left and right of the fixation point. The Ts in Experiment 1 were 0.6° high and 0.5° wide, and the squares in Experiment 2 were $0.5^\circ \times 0.5^\circ$. The colored search items were presented in red (CIE 1976 coordinates $u' = 0.431$, $v' = 0.529$), green ($u' = 0.145$, $v' = 0.588$), and blue (Experiment 2 only; $u' = 0.167$, $v' = 0.306$), with luminances within 3% of 18.0 cd/m^2 . The luminance of the gray search items was also within this luminance range. The search arrays were presented on a gray background with a luminance of 2.6 cd/m^2 .

condition in which attention was directed to the LVF versus the condition in which attention was directed to the RVF.

Method

Participants. The participants in this study were 20 neurologically normal students (10 per experiment), ranging between 18 and 32 years of age.

Stimuli and procedure: Experiment 1. The stimuli in Experiment 1 were displayed on a video monitor located 70 cm from the observer's eyes. A fixation cross remained visible at the center of the display at all times. Each stimulus array contained one red T and one green T, located in opposite visual fields (see Fig. 4 for details of stimulus size, position, and color). In the "single" condition, these two items were presented in isolation; in the "multiple" condition, an additional gray T was placed next to each colored T at a distance of 0.7° – 1.6° (measured center-to-center). The red T was in the LVF on 50% of trials and in the RVF on 50% of trials, randomly intermixed within a given trial block; the green T was always in the opposite hemifield from the red T. The orientation (upright or inverted) of each T within an array was randomly determined. The items within a given stimulus array were either all in the upper visual field or all in the lower visual

field, and these two types of stimuli were presented in separate trial blocks. The duration of each array was 750 ms and the interval between successive array onsets varied randomly between 1350 and 1650 ms. Each trial block consisted of 160 trials.

The participants were instructed to maintain fixation on the fixation cross at all times. Prior to the start of each trial block, participants were told which color (red or green) would be the target color for that block. The task was to press one of two buttons to indicate whether the T drawn in the target color was upright or inverted. The participants used the index finger of the preferred hand to indicate "upright" and the middle finger of the same hand to indicate "inverted." To be considered valid, a response had to occur between 200 and 1200 ms poststimulus. Each participant received six blocks of the single condition and six blocks of the multiple condition in counterbalanced order, evenly divided between attend-red and attend-green.

Stimuli and procedure: Experiment 2. The stimuli and procedure in Experiment 2 were identical to those in Experiment 1 except for the following changes. Each search array contained two colored squares, one in each hemifield, and the two colors in a given array were selected at random from the set of red, green, and blue. As a result, a given color was present in two-thirds of the arrays (the addition of a third color was necessary so that target-absent trials would be possible while maintaining two colored items in each array). At the beginning of each trial block, one of the three colors was designated as the target, and the participants were required to press an index-finger button on target-present trials and a middle-finger button on target-absent trials. Each block consisted of 216 trials, and each participant received six blocks in the single condition and six in the multiple condition, with each color serving as target in two of the six blocks in each condition.

Recording and analysis. The electroencephalogram (EEG) was recorded from 29 nonpolarizable tin electrodes mounted in an elastic cap. The electrode array included several sites from the International 10/20 system (F3, F4, C3, C4, T3, T4, T5, T6, P3, P4, O1, O2, Fz, Cz, Pz) and several nonstandard sites (CP1, CP2, CT5, CT6, INz, IN3, IN4, IPz, PO1, PO2, TO1, TO2, IT5, and IT6, as defined in Girelli & Luck, in press). Electrodes were also placed on the left and right mastoids. The right mastoid served as the reference electrode, and all scalp recordings were algebraically referenced offline to the average of the left and right mastoids. The horizontal electrooculogram (EOG) was recorded as the potential between electrodes located 1 cm lateral to the external canthus of each eye, and the vertical EOG was recorded from an electrode located beneath the left eye, referenced to the right mastoid. All trials in which eye movements, blinks, or incorrect behavioral responses were made were excluded from the averaged ERP waveforms (EOG artifacts led to the rejection of 17% of trials in Experiment 1 and 20% of trials in Experiment 2; behavioral errors are described below). The EEG and EOG were amplified by an SA Instrumentation amplifier with a bandpass of 0.01–80 Hz (half-power cutoffs of Butterworth filters with roll-offs of 12 dB per octave) and digitized at 250 Hz. To improve the signal-to-noise ratio, the waveforms from Experiment 1 were collapsed across red and green targets and also across upright and inverted targets, and the waveforms from Experiment 2 were collapsed across red, green, and blue targets (ERPs elicited by nontarget stimuli will not be considered here).

N2pc amplitude was quantified as the mean voltage between 170 and 280 ms poststimulus relative to a 200 ms mean prestimulus voltage and was measured at the six most posterior pairs of lateral electrode sites (PO1/2, O1/2, TO1/2, T5/6, IN3/4, and IT5/6). The N2pc measurements from Experiments 1 and 2 were concatenated into a single data set for analysis in a mixed-model analysis of variance (ANOVA) with a between-subjects factor of target complexity (Experiment 1 versus Experiment 2) and within-subjects factors of display type (single versus multiple), array position (upper versus lower visual field), contralaterality (ipsilateral versus contralateral target relative to the electrode site), electrode hemisphere (left versus right hemisphere electrode site), and within-hemisphere electrode site. All *p* values reported below reflect the Greenhouse-Geisser correction for heterogeneity of variance and covariance (Jennings & Wood, 1976).

Results and Discussion

Behavioral results. Reaction time (RT) and accuracy are summarized for both experiments in Table 1. In general, mean RT was between 400 and 550

TABLE 1
 Mean Reaction Time (RT, in ms) and Accuracy (Acc, in Percentage) Values
 from Experiments 1 and 2

	Experiment 1		Experiment 2	
	Single	Multiple	Single	Multiple
Upper field RT	495 (10)	542 (12)	427 (10)	427 (10)
Lower field RT	493 (10)	541 (11)	413 (10)	420 (10)
Upper field Acc	95.6 (1.2)	95.9 (0.6)	98.8 (0.3)	98.7 (0.3)
Lower field Acc	95.2 (1.0)	95.7 (0.6)	98.9 (0.3)	99.1 (0.3)

Note. Standard errors in parentheses.

ms, and mean accuracy was greater than 95% correct. Responses were generally faster and more accurate for the feature detection task of Experiment 2 than for the conjunction discrimination task of Experiment 1, leading to a significant main effect of target complexity in analyses of RT and accuracy, $F(1,18) = 22.38, p < .001$ and $F(1,18) = 9.54, p < .01$, respectively. The presence of a distractor item adjacent to the target item had no effect on either RT or accuracy for the simple feature detection task, but led to an increase in RT of almost 50 ms for the conjunction discrimination task, yielding a significant interaction between the target complexity and display type factors, $F(1,18) = 12.64, p < .005$. The presence of distractors had no significant effect on accuracy in either experiment, although this may be because accuracy was near ceiling.

Electrophysiological results. Figure 5 shows the ERPs recorded at temporal-occipital scalp sites. Three main effects of the experimental manipulations can be observed in Fig. 5. First, the N2pc component was generally larger for arrays presented in the lower visual field than for arrays presented in the upper visual field, and it was even reversed in polarity for upper-field stimuli at some electrode sites. Second, the N2pc component was generally larger for the multiple condition than for the single condition. Third, the N2pc component was substantially larger and longer lasting for the conjunction discrimination task used in Experiment 1 than for the feature detection task used in Experiment 2.⁴

⁴ It might be objected that the use of different stimuli in Experiments 1 and 2 might be responsible for the differences in N2pc amplitude between Experiments 1 and 2 rather than the differences in tasks. To rule out this possibility, we conducted a control experiment in which the participants performed the feature detection task with the T-shaped stimuli of Experiment 1 in one condition and the square stimuli of Experiment 2 in a different condition. As was the case for the feature detection task used in Experiment 2, very little N2pc activity was observed in this control experiment, and N2pc amplitude did not differ between the T stimuli and the square stimuli.

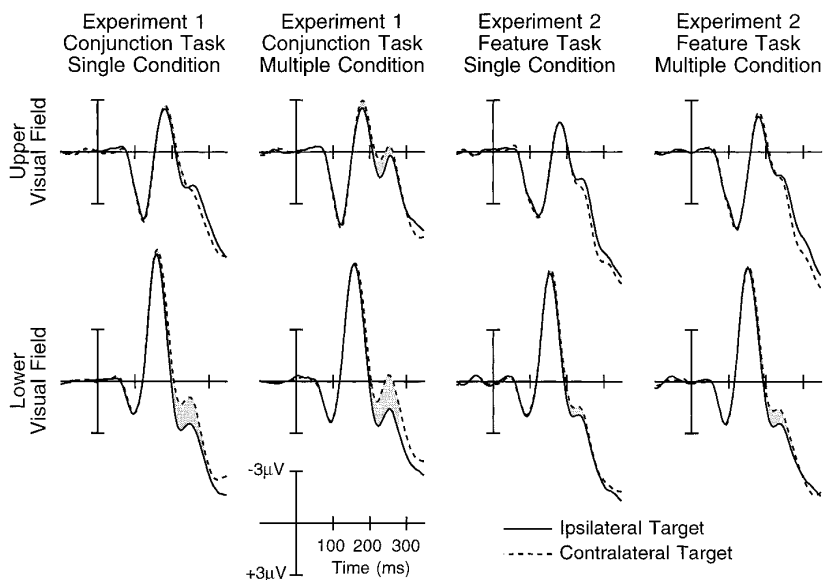


FIG. 5. ERP waveforms from the temporal-occipital electrode sites (TO1 and TO2) in Experiments 1 and 2, averaged across participants. Separate waveforms are shown for ipsilateral and contralateral targets, relative to the electrode site, and collapsed across the left and right hemisphere sites. The N2pc component is indicated by the shaded regions. These waveforms were digitally low-pass filtered (Gaussian impulse response function, standard deviation of 6 ms, half-amplitude cutoff at approximately 35 Hz).

In the statistical analysis, the size of the N2pc component was represented by the contralaterality factor, which reflected the degree to which the voltage in the N2 latency range was different for contralateral versus ipsilateral targets at a given electrode site. The variations in N2pc amplitude described above therefore led to significant interactions between the contralaterality factor and the factors of: (a) array position, $F(1,18) = 36.71$, $p < .001$; (b) display type, $F(1,18) = 4.76$, $p < .05$; and (c) target complexity, $F(1,18) = 7.02$, $p < .02$. There were no significant higher-order interactions between these factors.

The virtual elimination of the N2pc component for upper-field targets suggests that the N2pc component is generated in one or more topographically mapped areas, such as V4, and is not generated in the human homologue of macaque area TE, because this area does not contain separate maps of the upper and lower fields. However, this pattern of results may instead reflect a functional asymmetry in the allocation of attention for the upper and lower fields. Support for this alternative was provided recently by He *et al.* (1996), who found an analogous psychophysical asymmetry in attention.

The pattern of results from Experiments 1 and 2 generally mirrors the pattern of single-unit results obtained by Chelazzi *et al.* (1993, 1994). There

were three primary areas of concordance: (a) the larger N2pc observed in Experiment 1 than in Experiment 2 mirrors the larger single-unit modulations observed for complex stimuli than for simple color patches; (b) the larger N2pc for the multiple condition than for the single condition mirrors the larger single-unit modulations observed when multiple stimuli fell inside the receptive field of the neuron being recorded; and (c) the larger N2pc for lower-field arrays than for upper-field arrays would be expected on the basis of the presumed geometry and topographic mapping of area V4. These results therefore indicate that the single-unit attention effects observed in monkeys provide a reasonable model of the operation of attention in humans and provide additional support for the ambiguity resolution theory.

There is one aspect of the present results that is not entirely consistent with the single-unit experiments. Specifically, the N2pc component was never completely eliminated by reducing the number of distractors or by using a simple feature detection task, at least for stimuli in the lower visual field. In contrast, the single-unit attention effects were completely eliminated by the use of simple color patches and were almost completely eliminated when the test items were presented in separate hemifields without any nearby distractors. This discrepancy may be due to the use of naive observers in the present study, as opposed to the highly overtrained monkeys typically used in single-unit recordings. Support for this proposal was provided by Cave and Zimmerman (submitted), who examined the effects of training on the suppression of distractors in human observers and found that suppression at locations near the target is greater than suppression at distant locations—consistent with the local suppression effects observed in monkeys—but only after the observers received extensive training. Similarly, even if focused attention is completely unnecessary for detecting the presence of a salient color patch, the use of attention may lead to greater confidence and an untrained observer may therefore focus attention on a simple target to verify the correctness of the discrimination.

EXPERIMENT 3

The attention effects observed in visual cortex by Chelazzi *et al.* (1993, 1994) were consistently larger when the monkeys were required to foveate the target rather than making a manual detection response, presumably because the foveate task required the monkey to determine the precise location of the target. In Experiment 3, we performed an analogous manipulation in which we compared the color detection task used in Experiment 2 with a task in which the participants were required to foveate the target. We predicted that, like the single-unit modulations, the N2pc component would be larger for the foveate task than for the detection task.

Method

Participants. The participants were 12 neurologically normal students between 18 and 32 years of age.

Stimuli and procedure. The stimuli in this experiment were identical to those used in Experiment 2, except that the signal-to-noise ratio was increased by using only lower-field arrays and only the multiple condition. As in Experiment 2, one of the three colors was designated the target at the beginning of each trial block. In the “button-press” condition, the participants were required to press a button with the index finger of the preferred hand when they detected the target color; no response was required for target-absent arrays. In the “foveate” condition, the participants were required to make a saccade to the location of target color for target-present arrays and to make no eye movement for target-absent arrays. Each trial block consisted of 90 trials, and each participant received six blocks in the button-press condition and six blocks in the foveate condition, evenly divided among the three target colors. The order of conditions was counterbalanced across participants.

Recording and analysis. The recording and averaging procedures for Experiment 3 were the same as in Experiments 1 and 2. The N2pc measurement procedures had to be modified, however, because the eye-movement responses in the foveate condition produced lateralized electrical potentials that sometimes occurred within the latency range of the N2pc component, which could artificially inflate the N2pc amplitude measurements. The majority of eye movements in this experiment began after the time range of the N2pc component (i.e., after 300 ms poststimulus), but there was significant variability in saccade onset time both between and within participants, and some eye movements began as early as 200 ms poststimulus. We therefore used a strict artifact rejection procedure that eliminated all trials with eye movements that began within 250 ms of stimulus onset, and we also limited the N2pc measurement window to 200–250 ms poststimulus. The artifact rejection procedure resulted in the elimination of between 1 and 44% of trials from the single-subject ERP waveforms in the foveate condition, with a mean of 16%, compared to between 0 and 18% in the button-press condition, with a mean of 7% (these values include trials eliminated due to blinks as well as saccades, because blinks and vertical saccades could not be differentiated by our artifact-rejection software).

Most of the trials rejected in the foveate condition were trials with fast but correct responses, and the exclusion of such trials only in the foveate condition could confound any comparisons between the foveate and button-press conditions. As a result, we also eliminated the trials with the fastest responses in the button-press condition in order to obtain similar proportions of rejected trials in the two conditions. Specifically, for each participant we estimated the proportion of trials that were excluded because of fast saccades in the foveate condition by subtracting the proportion of rejected trials in the button-press condition from the proportion of rejected trials in the foveate condition. We then adjusted the bottom end of the range of acceptable reaction times in the button-press condition until the proportion of trials discarded because of fast reaction times in the button-press condition was equal to the estimated proportion of trials discarded because of fast saccades in the foveate condition. This procedure was performed individually for each participant, and the minimum allowable reaction times ranged between 240 and 455 ms, with a mean of 336 ms (the maximum allowable reaction time was held constant at 1350 ms for all participants). We also analyzed the data without rejecting trials with fast button presses, and the results were virtually identical.

The N2pc measurements from Experiment 3 were entered into a within-subjects ANOVA with factors of task (button-press versus foveate), contralaterality (ipsilateral versus contralateral target relative to the electrode site), electrode hemisphere (left versus right hemisphere electrode site), and within-hemisphere electrode site.

Results and Discussion

Mean reaction time for the button-press condition was 454 ms, and mean accuracy was 99.6% correct. Software was not available for quantifying the timing and accuracy of the saccades in the foveate condition, but performance was monitored online during the course of each session, and very few errors

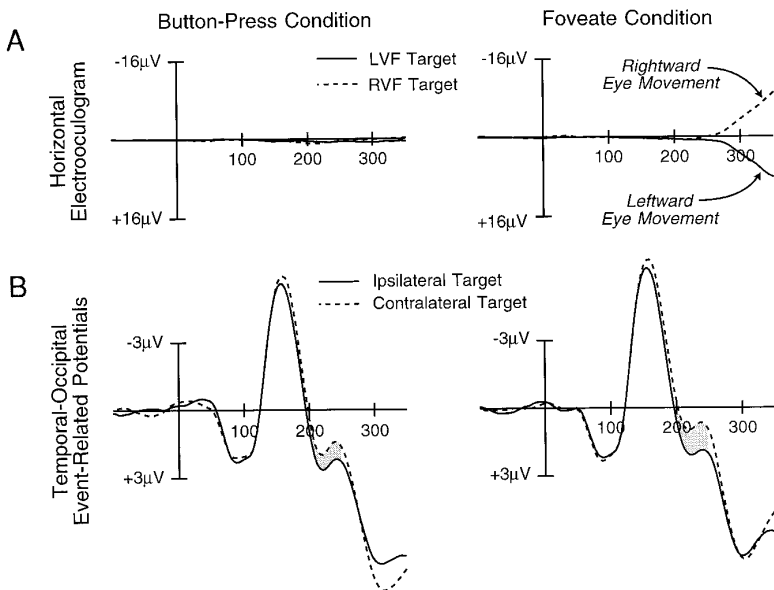


FIG. 6. (A) Horizontal EOG waveforms from Experiment 3, averaged across participants, for LVF and RVF targets. Each degree of eye rotation creates a deflection of approximately 16 μV (Lins *et al.*, 1993). Note that there was no substantial deflection toward the target until after the N2pc measurement window (i.e., after 250 ms poststimulus). (B) ERP waveforms at the temporal–occipital sites, averaged across participants and divided according to whether the target was ipsilateral or contralateral to the electrode site. The N2pc component is indicated by the shaded regions. All waveforms in this figure were digitally low-pass filtered as in Fig. 5.

were observed. The time course of the saccade-produced voltages is displayed in Fig. 6A, which shows the averaged horizontal EOG waveforms for LVF and RVF targets (these waveforms were computed after the application of the artifact rejection procedures described above). As shown in Fig. 6, there was no substantial voltage difference between trials with LVF and RVF targets until after 250 ms. It should also be noted that less than 5% of the horizontal EOG voltage would be expected to propagate as far back as the posterior temporal and occipital scalp sites at which the N2pc was measured (Lins, Picton, Berg, & Scherg, 1993).

Figure 6B shows the ERPs elicited by targets at the temporal–occipital electrodes sites. As predicted, the N2pc component was larger in the foveate condition than in the button-press condition, resulting in a significant task X contralaterality interaction, $F(1,11) = 5.74, p < .05$. However, as in Experiment 2, a small but reliable N2pc component was observed in the button-press condition as well as the foveate condition. This was verified statistically by entering the N2pc measurements for these two conditions into separate ANOVAs. For both conditions, the main effect of contralaterality was signifi-

cant, $F(1,11) = 10.09$, $p < .01$ for the button-press condition and $F(1,11) = 18.72$, $p < .005$ for the foveate condition.

The larger N2pc component observed in the foveate condition was analogous to the larger single-unit effects observed by Chelazzi *et al.* (1993, 1994) in a similar comparison. These results thus provide additional evidence that the same attentional mechanisms that produce modulations of single-unit responses in monkeys also operate in human observers.

GENERAL DISCUSSION

As described above, the effects of attention observed by Chelazzi *et al.* (1993, 1994) in areas V4 and TE of monkey visual cortex are consistent with the ambiguity resolution theory, and the purpose of the present study was to determine whether attention operates similarly in human visual cortex.

The experiments presented here provide four separate pieces of evidence indicating that the operation of attention in the human visual system is at least qualitatively similar to the operation of attention in monkey visual cortex. First, the N2pc component was smaller when the search array contained only a single item in each hemifield than when two items were placed close together within each hemifield. This parallels the finding of larger single-unit attention effects when multiple stimuli are presented together inside the receptive field of the neuron being recorded (Chelazzi & Desimone, 1994; Chelazzi *et al.*, 1993; Luck *et al.*, 1997; Moran & Desimone, 1985; Treue & Maunsell, 1996). Second, the N2pc component was larger when the observers performed a complex conjunction discrimination task rather than a simple feature detection task, which parallels the finding of larger single-unit attention effects when monkeys perform a complex discrimination task rather than a simple color detection task (Chelazzi *et al.*, 1993, 1994). Third, the N2pc component was larger for simple feature targets when the observers were required to foveate the targets than when they simply made a manual detection response, which parallels the finding of larger single-unit effects when monkeys are required to foveate the target (Chelazzi & Desimone, 1994). Finally, the substantial reduction in N2pc amplitude for upper-field targets compared to lower-field targets is consistent with the proposal that the N2pc component arises, at least in part, from a topographically mapped area such as V4 (although it may instead reflect a functional asymmetry in attentional allocation). These findings are all in accord with the ambiguity resolution theory, and they are also consistent with several related cognitive-level theories of attention (see, e.g., Cohen & Ivry, 1991; Prinzmetal *et al.*, 1995; Treisman, 1996).

Although the present findings and many previous studies are consistent with the ambiguity resolution theory, it is important to recognize that this theory is not a complete account of visual selective attention. For example, there have been several reports of attention effects in visual cortex under conditions that would not be expected to yield ambiguous neural representations (e.g., Connor *et al.*, 1996; Motter, 1993; Spitzer *et al.*, 1988; Treue &

Maunsell, 1996), and some other role of attention must be postulated to explain these results. Similarly, although this theory can explain how the visual system combines information about separable features such as color and orientation, it cannot easily explain how information is combined when it is necessary to determine the spatial relationships between object parts, such as when an observer searches for an inverted T in an array of upright Ts (see Logan, 1994; Luck & Beach, in press). In addition, although the ambiguity resolution theory is intended to provide an explanation of attention in terms of the response properties of neurons in visual cortex, we have not yet provided a detailed account of how the proposed attentional mechanism could be implemented in neural circuitry. Additional work will be necessary to address these limitations.

REFERENCES

- Cave, K. R., & Zimmerman, J. M. *Flexibility in spatial attention before and after practice*. Manuscript submitted for publication.
- Chelazzi, L., & Desimone, R. (1994). Responses of V4 neurons during visual search. *Society for Neuroscience Abstracts*, **20**, 1054.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, **363**, 345–347.
- Clarke, S., & Miklosy, J. (1990). Occipital cortex in man: Organization of callosal connections, related myelo- and cytoarchitecture, and putative boundaries of functional visual areas. *Journal of Comparative Neurology*, **298**, 188–214.
- Cohen, A., & Ivry, R. (1989). Illusory conjunctions inside and outside the focus of attention. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 650–663.
- Cohen, A., & Ivry, R. B. (1991). Density effects in conjunction search: Evidence for a coarse location mechanism of feature integration. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 891–901.
- Connor, C. E., Gallant, J. L., Preddie, D. C., & Van Essen, D. C. (1996). Responses in area V4 depend on the spatial relationship between stimulus and attention. *Journal of Neurophysiology*, **75**, 1306–1308.
- Crick, F., & Koch, F. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, **375**, 121–123.
- Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, **33**, 25–62.
- De Valois, R. L., Morgan, H., & Snodderly, D. M. (1974). Psychophysical studies of monkey vision III. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Research*, **14**, 75–81.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, **18**, 193–222.
- Desimone, R., Moran, J., Schein, S. J., & Mishkin, M. (1993). A role for the corpus callosum in visual area V4 of the macaque. *Visual Neuroscience*, **10**, 159–171.
- Desimone, R., & Schein, S. J. (1987). Visual properties of neurons in area V4 of the Macaque: Sensitivity to stimulus form. *Journal of Neurophysiology*, **57**, 835–868.
- Desimone, R., Wessinger, M., Thomas, L., & Schneider, W. (1990). Attentional control of visual perception: Cortical and subcortical mechanisms. *Cold Spring Harbor Symposium on Quantitative Biology*, **55**, 963–971.
- Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. P. a. O. S. Marin (Ed.), *Mechanisms of Attention: Attention and Performance XI* (pp. 171–187). Hillsdale, NJ: Erlbaum.

- Duncan, J., Humphreys, G., & Ward, R. (in press). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*.
- Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, **360**, 343–346.
- Gattass, R., Sousa, A. P. B., & Gross, C. G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *Journal of Neuroscience*, **8**(6), 1831–1845.
- Girelli, M., & Luck, S. J. (in press). Are the same attentional mechanisms used to detect visual search targets defined by color, orientation, and motion? *Journal of Cognitive Neuroscience*.
- Handy, T. C., Kingstone, A., & Mangun, G. R. (in press). Spatial distribution of visual attention: Perceptual sensitivity and response latency. *Perception & Psychophysics*.
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, **383**, 334–337.
- Ito, M., Tamura, J., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology*, **73**, 218–226.
- Jennings, J. R., & Wood, C. C. (1976). The e-adjustment procedure for repeated-measures analyses of variance. *Psychophysiology*, **13**, 277–278.
- Johnston, J. C., & Pashler, H. (1990). Close binding of identity and location in visual feature perception. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 843–856.
- Komatsu, H., & Ideura, Y. (1993). Relationship between color, shape, and pattern selectivities in the inferior cortex of the monkey. *Journal of Neurophysiology*, **70**, 677–694.
- LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, **96**, 101–124.
- Lins, O. G., Picton, T. W., Berg, P., & Scherg, M. (1993). Ocular artifacts in EEG and event-related potentials I: Scalp topography. *Brain Topography*, **6**, 51–63.
- Llinás, R., & Paré, D. (1996). The brain as a closed system modulated by the senses. In R. Llinás & P. S. Churchland (Eds.), *The Mind–Brain Continuum*. Cambridge, MA: MIT Press.
- Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 1015–1036.
- Luck, S. J., & Beach, N. J. (in press). Visual attention and the binding problem: A neurophysiological perspective. In R. Wright (Ed.), *Visual Attention*. London: Oxford University Press.
- Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, **77**, 24–42.
- Luck, S. J., & Hillyard, S. A. (1994a). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, **31**, 291–308.
- Luck, S. J., & Hillyard, S. A. (1994b). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 1000–1014.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, **229**, 782–784.
- Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2 and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, **70**, 909–919.
- Niebur, E., Koch, C., & Rosin, C. (1993). An oscillation-based model for the neuronal basis of attention. *Vision Research*, **33**, 2789–2802.
- Prinzmetal, W., Henderson, D., & Ivry, R. B. (1995). Loosening the constraints on illusory conjunctions: The role of exposure duration and attention. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 1362–1375.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing. I: Detection, search and attention. *Psychology Review*, **84**, 1–66.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen,

- B. R., & Tootell, R. B. H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, **268**, 889–893.
- Shulman, G. L., Wilson, J., & Sheehy, J. B. (1985). Spatial determinants of the distribution of attention. *Perception and Psychophysics*, **37**, 59–65.
- Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, **18**, 555–586.
- Sohn, Y.-S., Liederman, J., & Reinitz, M. T. (1996). Division of inputs between hemispheres eliminates illusory conjunctions. *Neuropsychologia*, **34**, 1057–1068.
- Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, **240**, 338–340.
- Suzuki, S., & Cavanagh, P. (in press). Focused attention distorts visual space: An attentional repulsion effect. *Journal of Experimental Psychology: Human Perception and Performance*.
- Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, **40**, 201–237.
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, **6**, 171–178.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, **95**, 15–48.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, **14**, 107–141.
- Treisman, A., Vieira, A., & Hayes, A. (1992). Automaticity and preattentive processing. *American Journal of Psychology*, **105**, 341–362.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97–136.
- Treue, S., & Maunsell, J. H. R. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, **382**, 539–541.
- Usher, M., & Niebur, E. (1996). Modeling the temporal dynamics of IT neurons in visual search: A mechanism for top-down selective attention. *Journal of Cognitive Neuroscience*, **8**, 311–327.
- Van Essen, D. C. (1985). Functional organization of primate visual cortex. In A. Peters & E. G. Jones (Eds.), *Cerebral Cortex* (pp. 259–329). New York: Plenum.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 419–433.

(Accepted February 12, 1997)