On the role of selective attention in visual perception

STEVEN J. LUCK* AND MICHELLE A. FORD

Department of Psychology, University of Iowa, 11 Seashore Hall E, Iowa City, IA 52242-1407

ABSTRACT What is the role of selective attention in visual perception? Before answering this question, it is necessary to differentiate between attentional mechanisms that influence the identification of a stimulus from those that operate after perception is complete. Cognitive neuroscience techniques are particularly well suited to making this distinction because they allow different attentional mechanisms to be isolated in terms of timing and/or neuroanatomy. The present article describes the use of these techniques in differentiating between perceptual and postperceptual attentional mechanisms and then proposes a specific role of attention in visual perception. Specifically, attention is proposed to resolve ambiguities in neural coding that arise when multiple objects are processed simultaneously. Evidence for this hypothesis is provided by two experiments showing that attention—as measured electrophysiologically—is allocated to visual search targets only under conditions that would be expected to lead to ambiguous neural coding.

The Problem of the Locus of Selection

Because of its finite computational resources, the human brain must process information selectively in a variety of domains. For example, we may limit processing to a subset of the many possible objects that could be perceived, a subset of the many possible decisions that could be made, a subset of the many possible memories that could be accessed, and a subset of the many possible actions that could be performed. Although the fundamental need for selective processing is present at each of these stages, there are important differences in the computations performed by different cognitive systems, and it is therefore likely that substantially different attentional mechanisms are responsible for selective processing at each stage. However, attention is often treated as a unitary cognitive process, and few researchers have attempted to isolate and characterize the different mechanisms of attention that operate within different cognitive systems (for some exceptions, see refs. 1 and 2). This is due, in part, to the fact that most studies of attention have relied on measures of behavioral output that reflect the combined effects of many different cognitive systems, making it difficult to determine which system was responsible for a given change in response speed or accuracy. The techniques of cognitive neuroscience, in contrast, make it difficult to avoid dividing attention into different components, because these techniques naturally tend to subdivide cognitive processes on the basis of their timing and/or neuroanatomical substrates. In this paper, we will discuss the general use of these techniques in making a coarse distinction between attentional mechanisms that influence perception and attentional mechanisms that operate postperceptually. We will then provide a more detailed description of our present understanding of the role of attention in visual perception.

Consider a visual search task in which an observer is presented with an array of 15 green letters and 1 red letter and must report whether the red letter is a T. A task such as this can generally be performed quite easily because the distinctive color of the potential target item allows the observer to focus attention onto a single item. It is possible to demonstrate that attention is used by occasionally asking the observer to report the identities of the green items after the offset of the array. Recall is found to be quite poor for these items, which indicates that attention was indeed focused on the one red item. But why, exactly, is the observer unable to report the green letters? One possibility is that attention operates at a very early stage such that the observer literally does not see the green items [this is the sort of model implied by Broadbent (3)]. Another possibility is that low-level, feature-based processing occurs for all items, but integrated object representations are formed only for attended items, making it impossible for the observer to report anything more than the basic features of the green items [this alternative corresponds roughly to Treisman’s feature integration theory (4)]. A third possibility is that every item in the array is fully identified, but only attended items are stored in working memory so that they can be reported [this sort of model has been promoted by Duncan et al. (5–7)]. This general issue is typically called the “locus-of-selection” question, with some investigators taking the “early-selection” position that attentional selection operates at the level of perceptual processing, and other investigators taking the “late-selection” position that attention operates after stimulus identification is complete.

How can these three models of attention be empirically distinguished? There are a variety of sources of evidence indicating that low-level features are extracted preattentively (8–11), but it is very difficult to determine whether observers fail to report an item because it was not identified or because it was not stored in working memory. The most common approach to this problem has been to use indirect measures of identification such as priming and interference to determine whether ignored items have been identified. In particular, many experiments have used variants of the Stroop interference paradigm, in which responses are slowed when an unattended source of information conflicts with the attended source of information. Proponents of late selection have provided many clear examples of such interference and have concluded that these results indicate that unattended stimuli are fully identified (12, 13). This conclusion seems logical because a stimulus could not cause interference if it was not identified. The problem with this sort of evidence, however, is that results of this nature could simply mean that the stimuli
and task did not engender a highly focused state of attention and that some information about the to-be-ignored stimuli “leaked through” the attentional filter. Indeed, other investigators have provided examples of experiments in which attention was highly focused and no interference from the to-be-ignored stimuli was present, from which they concluded that the to-be-ignored stimuli were not perceived (14, 15). However, this conclusion is also problematic because it is possible that the to-be-ignored stimulus was identified but then blocked from influencing behavior at a postperceptual stage (as argued in refs. 16 and 17). Thus, neither the presence nor the absence of interference provides clear evidence about the locus of selection.

**Neurophysiological Evidence for Early Selection**

The main difficulty in determining the locus of selection from studies of behavior is that behavior reflects the output of processing and does not directly reveal the individual steps that led to that output. The techniques of cognitive neuroscience, however, naturally tend to subdivide processing into different stages on the basis of neuroanatomy and/or timing. For example, if attention can be shown to influence the initial neural activity in sensory processing regions, then this would provide clear evidence for early selection. In contrast, if attention influences only late neural activity in high-level processing regions, then this would provide clear evidence for late selection (assuming that the stimuli and task engendered a highly focused attentional state). Note, however, that it is important to assess both the timing and the neuroanatomical site of the effects of attention. For example, finding that attention influences neural activity in primary visual cortex does not necessarily indicate that attention operates before stimulus identification is complete, because it is possible that neurons in this area also participate in postperceptual processes such as working memory. Similarly, finding that attention begins to modulate neural activity beginning 150 ms after stimulus onset does not prove that attention modulates perceptual processing, because it is possible that postperceptual processes have begun by this time. Extreme cases, however, may provide fairly compelling evidence. In particular, any effects of visual attention observed before 100 ms poststimulus or observed in the retina are very likely to reflect modulations of perceptual processing. In addition, attention effects that occur at a relatively early time and at a relatively early neuroanatomical locus are likely to reflect early selection.

Cognitive neuroscience techniques have been applied to the study of visual attention for more than 20 years, and these studies have generally indicated that focusing attention onto a location in space leads to a modulation of perceptual processing, although not until a significant amount of early sensory analysis has taken place. This conclusion is based on event-related potential (ERP) and positron emission tomography (PET) studies in humans and single-unit recordings in monkeys. Many of these studies have used variations on the paradigm shown in Fig. 1A. In this paradigm, the subjects are instructed to attend to one location during some trial blocks and to a different location during others, and they are required to detect occasional target stimuli at the attended location that are interspersed among nontarget stimuli presented at both attended and unattended locations. As shown in Fig. 1B, the initial “C1” wave of the ERP waveform is not influenced by whether the evoking stimulus is presented at the attended location or at the ignored location, but the subsequent “P1” and “N1” waves are larger for the attended-location stimuli.

Several studies have indicated that the C1 wave is generated in area V1 (18–20), so the absence of an attentional modulation of this component suggests that attention operates after this very early stage of processing. A recent PET study has indicated that the P1 wave arises in extrastriate areas of visual cortex (21), and the combination of this anatomical information with the fact that the P1 attentional modulation begins before 100 ms provides excellent evidence that visual–spatial attention influences perceptual processing. It is also important to note that these effects are identical for both target and nontarget stimuli (22), which provides further evidence that attention operates before perceptual processing is complete.

More precise evidence has been obtained from single-unit recordings in area V1 (primary visual cortex) and in area V4 (an intermediate stage in the object recognition pathway) (23). As shown in Fig. 1C and D, responses in area V4 were found to be larger for attended-location stimuli than for ignored-location stimuli, but no effects of attention were observed in area V1. Importantly, both the initial stimulus-evoked activity

---

**Fig. 1.** (A) Common experimental design for neurophysiological studies of attention. The outline squares are continuously present and mark the two locations at which the solid square can be flashed. (B) Example occipital ERPs recorded in a paradigm of this nature (data from ref. 42). Note that the C1 wave (generated in area V1) shows no attention effect, whereas the P1 and the N1 waves (generated in extrastriate cortex) are larger for the attended stimuli. (C) Single-unit responses from area V4 in a similar paradigm (data from ref. 23). Note that the response is larger for attended compared with ignored stimuli. (D) Single-unit responses from area V1 (data from ref. 23) showing no effect of attention.
and the attention effect in area V4 began at 60 ms poststimulus, indicating that attention modulates the initial afferent volley in this area. In addition, identical effects were observed for both target and nontarget stimuli. Thus, several sources of evidence converge on the conclusion that visual–spatial attention modulates perceptual processing in extrastriate visual areas, beginning within 100 ms of stimulus onset.

The conclusion that visual–spatial attention can influence perceptual processing is now very well supported, but it is important to note three limitations on this conclusion. First, there are many circumstances under which attention does not operate at an early stage. Indeed, there is growing evidence that early attention effects are observed primarily under conditions of high perceptual load (24), which is sensible given that there is no point in suppressing the identification of irrelevant objects unless the visual system is so overloaded that the irrelevant objects interfere with the identification of relevant objects. Second, even when attention does operate at an early stage, it may simultaneously operate at a late stage (for a particularly clear example of this, see ref. 25). Finally, although visual–spatial attention may operate at an early stage, other types of attention may be restricted to postperceptual stages. For example, when an observer detects a target within a rapid stream of stimuli, the allocation of attention to this target causes a temporary decline in the ability to detect subsequent targets (26, 27). Although it is tempting to suppose that this “attentional blink” effect reflects a perceptual filtering of the stimuli that follow the first target, ERP studies have shown that these stimuli are fully identified even though they cannot be accurately reported (28, 29). Thus, “time-based” attentional selection may occur only after perception is complete and may operate to control the encoding and consolidation of information in working memory rather than the identification of stimuli.

**Visual Attention and the Binding Problem**

Now that we have considered the general problem of isolating perceptual-level attentional mechanisms, we turn to the specific role of attention in visual perception. It is important to realize, however, that visual perception is extremely rich and complex in primates, and attention may actually play several distinct roles within different visual subsystems. Thus, in addition to the possibility of different attentional mechanisms that operate at different stages of processing, it is also necessary to consider the possibility that several different attentional mechanisms may operate within a stage, especially if something as complex as visual perception is considered as a single stage. Thus, although this section will focus on a single attentional mechanism that operates during visual perception, there are likely to be other perceptual-level attentional mechanisms as well.

To isolate separate mechanisms of attention within a stage, we have taken the approach of trying to understand the specific computational problems that selective attention may be used to solve. Here we focus on a role that attention may play in solving the “binding problem” (30–32). The binding problem is particularly salient within the ventral object recognition pathway when the visual input consists of multiple concurrently presented objects. Neurons at the higher levels of this pathway tend to have large receptive fields, which is a useful property insofar as it reduces the number of neurons that are necessary to cover the visual field and allows objects to be coded in a relatively position-independent manner (33). However, these large receptive fields lead to a significant problem when multiple objects fall inside a given receptive field (as is virtually always true in natural visual scenes). Specifically, the response of a neuron is ambiguous when multiple objects fall inside the receptive field, because it is not clear which of the objects is responsible for the neuron’s response. Consider, for example, a color-selective neuron that responds to red stimuli but not blue stimuli. If a red square and a blue circle fall inside this neuron’s receptive field, the neuron will fire, indicating that a red stimulus is present, but it will not be clear which of the two stimuli is red. In this manner, neural responses can be ambiguous, and accurate perception of objects such as these therefore requires a mechanism for linking together features that are coded by different neurons (for additional discussion, see refs. 34 and 35).

There are two general classes of mechanisms that have been proposed for solving the binding problem. One solution uses the temporal microdynamics of neural activity to link together the neurons that code a given object. For example, if action potentials from all of the neurons coding the red square occurred simultaneously, but at different times from the neurons coding the blue square, the timing of the action potentials could be used to bind together the many features of an object (31, 36). Although this proposed binding mechanism has a great deal of merit, it is probably not sufficient to explain human perceptual performance (for a detailed discussion, see ref. 35). We will therefore focus on a second means of solving the binding problem, namely serial processing. Specifically, because the binding problem arises from the simultaneous presence of multiple objects, it can be solved by using selective attention to restrict the responses of a set of neurons to a single object at a time. For example, attention might be focused initially onto the blue circle such that the red square was prevented from producing any responses and the neural activity reflected only the features of the blue circle. After the blue circle is identified, attention can then be shifted to the red square, such that the neurons become responsive only to the features of that one object and are no longer ambiguous.

We call this the ambiguity resolution theory of visual selective attention (note that this theory is closely related to Treisman’s feature-integration theory; see ref. 8).

We have previously described several pieces of evidence that support this role of attention in resolving the ambiguous neural representations that lead to the binding problem (23, 34, 37). Several of these experiments tested the important prediction that perceptual-level attentional mechanisms should be necessary primarily when the observer must combine information from neurons that code different features of the target. For example, attention should be necessary if the observer is required to detect a blue square among red squares and blue circles, because without attention it will be difficult to know whether a blue object is a square or a circle. However, perceptual-level attention should not be necessary if the observer is simply required to report the presence or absence of blue in the array, because this does not require the feature to be linked to a specific object or location. It is important to note, however, that higher-level attentional mechanisms may be required for both of these tasks (e.g., for both tasks, it is probably necessary to transfer information about the stimulus array into working memory).

In a recent series of experiments, we compared the allocation of attention in a conjunction discrimination task with the allocation of attention in a feature detection task, and we predicted that attention would be present only when a conjunction discrimination was required (34). Consistent with this prediction, greater attentional allocation was observed during the conjunction discrimination task than during the feature discrimination task. However, although attentional allocation was reduced for the feature task, significant attention effects were still observed. One possible explanation for this pattern of results is that attention does not play a special role in feature binding and that attentional requirements simply vary in a continuous manner as a function of the difficulty of the task (5, 6). Another possibility, however, is that attention is indeed unnecessary for feature detection, but that the subjects in these experiments occasionally allocated attention to the simple target.
feature targets even though attention was not necessary for performing the task. In other words, because there was no penalty for allocating attention to the visual search targets, this previous study did not really address the necessity of attention for performing the feature and conjunction tasks. We have therefore conducted a new set of experiments in which an attention-demanding central task was added to the visual search task; this was intended to discourage the subjects from unnecessarily using attention to perform the visual search task.

New Experimental Evidence

The present experiments were designed to test the hypothesis that perceptual attentional mechanisms are necessary for discriminating conjunctions but not for detecting features, and it was therefore important to use a specific measure of the operation of attention at the level of perception. Consequently, we chose to measure attention by means of the N2-posterior-contralateral (N2pc) component of the ERP waveform. The N2pc component is a negative-going deflection in the time range of the N2 complex, and it is typically observed at posterior scalp sites contralateral to the location of a visual search target. Previous studies have indicated that this component reflects the focusing of attention onto a target to filter out interfering information from nearby distractor items (38), and we have recently shown that this component strongly resembles attentional suppression effects that have been observed in single-unit recordings from area V4 and from inferotemporal cortex (34, 39). In addition, the N2pc component is sensitive to sensory factors such as target position and distractor density, and it is therefore likely to reflect a perceptual-level attentional mechanism.

The stimuli used in our first experiment are shown in Fig. 2A. Subjects performed a visual search task with arrays of 12 squares, 10 of which were gray and 2 of which were colored, and they were required to report whether a particular color was present in each array. In some trial blocks subjects performed this task alone, and in other trial blocks this task was combined with a central task in which subjects had to report whether a letter presented at fixation was a consonant or a vowel. The letter was degraded with random visual noise, and the intensity of this noise was adjusted so that accuracy on this task was approximately 85% correct. The subjects were told that the central task was to be considered the primary task and that they should devote all of their attention to this task. The central letter and the visual search array were presented simultaneously for a duration of 100 ms; this timing regimen was designed to prevent the subjects from shifting attention to the visual search target after discriminating the central target. Performance of the visual search task was nearly perfect when it was performed alone (99% correct) and only slightly worse when combined with the central task (97% correct).

The ERPs from this experiment are shown in Fig. 2B. A small but consistent N2pc component was observed when the visual search task was performed in isolation, which replicates previous studies (34, 40). When the central task was added, however, the N2pc component was completely eliminated. This result indicates that the perceptual-level attentional mechanism reflected by the N2pc component is not necessary for accurate feature detection.

A second experiment was also conducted in which the visual search task was changed so that it required the discrimination of a color-orientation conjunction. As shown in Fig. 3A, each item in the visual search array was a horizontally or vertically oriented rectangle, and the subjects were required to discriminate the orientation of the one rectangle that was drawn in a particular prespecified color. This is a conjunction discrimination insofar as it requires the observer to link two separate features—color and orientation—to the same object or location. As in experiment 1, the visual search task was conducted by itself (leading to 98% accuracy) or in combination with the central letter-discrimination task (leading to 96% accuracy). Because the visual search task in this experiment required the subjects to combine color and orientation information, we predicted that attention would be necessary for accurate performance.

The ERPs from this experiment are shown in Fig. 3B. An N2pc component was observed when the visual search task was performed alone, and it was larger than the N2pc component observed for feature targets in the first experiment. In addition, when the central task was added, the N2pc component was not eliminated, but was simply shifted to a later time. This suggests that subjects were indeed focusing attention onto the central task, but they were able to shift attention to the iconic image of the visual search array after completing the central discrimination task. Thus, even when the subjects were strongly motivated not to use attention, they still used attention when the task required them to conjoin a color and an orientation. These results are consistent with the proposal that one major role of attention in visual perception is the resolution of ambiguities in neural coding that arise when multiple objects are present and the observer is required to bind together neural responses that belong to a given object.

FIG. 2. (A) Example of the stimuli used in experiment 1. (B) ERPs from experiment 1, recorded at lateral occipital scalp sites ipsilateral and contralateral to the location of the target stimulus. The presence of an N2pc wave (shaded area) when the feature detection task was performed alone indicates that visual–spatial attention was allocated to the target in this condition. In contrast, although the observers were able to accurately detect the feature target while performing the concurrent letter discrimination task, no N2pc was observed, indicating that attention was not necessary for feature detection.
the type of central task used by Joseph et al. (41) leads to postperceptual impairments in which items that are fully identified fail to be stored in working memory (28, 29), and it is very likely that the impairment in feature detection performance observed by Joseph et al. (41) reflects the operation of a postperceptual attentional mechanism. Thus, the present experiments indicate that feature identification can be accomplished without the use of perceptual-level attentional mechanisms, whereas the results of Joseph et al. indicate that attention may be required to make the identified features available to awareness.

We would like to acknowledge the important role that several individuals have played in formulating these ideas and in conducting the experiments, including Steven A. Hillyard, Leonardo Chelazzi, Robert Desimone, Edward K. Vogel, Kimron L. Shapiro, Massimo Girelli, and Michele T. McDermott. Preparation of this manuscript and several of the studies described here were supported by Grant 95-38 from the McDonnell-Pew Program in Cognitive Neuroscience and by Grant 1 R29 MH56877-01 from the National Institute of Mental Health.


FIG. 3. (A) Example of the stimuli used in experiment 2. (B) ERPs from experiment 2, recorded at lateral occipital scalp sites ipsilateral and contralateral to the location of the target stimulus. A large N2pc wave (shaded area) can be observed for the conjunction task when it was performed alone (larger than for the feature task in experiment 1). When the conjunction task was combined with the central letter discrimination task, the N2pc was delayed, suggesting that the observers shifted attention to the iconic image of the target after completing the central letter task and could not perform this task without the use of spatially focused attention.

It is useful to contrast these results with the results of an analogous psychophysical study that was recently reported by Joseph et al. (41). As in the present experiments, this study combined a visual search task with a central attention-demanding task and assessed the extent to which the allocation of attention to the central task interfered with the visual search task. However, the central task in the Joseph et al. study was more complex than the task used in the present study and involved the detection of a target item within a rapid stream of nontarget items (all presented at the fovea). In contrast to the present study, Joseph et al. (41) found that observers could not accurately detect visual search targets defined by a simple feature when performing the concurrent central task. From these results, the authors concluded that there is no direct route from feature coding to awareness and that even simple features must pass through a limited-capacity attentional stage to reach awareness. Although this conclusion seems very sound, it would be easy to draw an unwarranted additional conclusion from these results; namely that attention is necessary for the accurate identification of simple features (as opposed to being necessary for making overt responses on the basis of the feature identities). We have previously shown that...