29 Short-Term Memory for the Rapid Deployment of Visual Attention

KEN NAKAYAMA, VERA MALJKOVIC, AND ARNI KRISTJANSSON

ABSTRACT We describe a short-term memory system useful for the rapid deployment of focal visual attention. The memory system is primitive, temporary, cumulative, and efficacious. It automatically links separable features of objects to the act of attentional deployment so that visual perception and performance are greatly enhanced. The effects of the memory cannot be overridden by higher level knowledge. Its properties can account for those aspects of searching and foraging behavior currently attributed to "search images." The learning described here is likely to be general, not confined to the rapid deployment of focal visual attention.

Blindness as we commonly understand it is caused by damage to the eye or to some part of the visual pathway. This definition encompasses diseases of the retina as well as strokes affecting the visual cortex. There is another form of blindness that is not due to specific peripheral mechanisms but occurs in the absence of visual attention (McConkie and Zola, 1979). Grimes (1996) showed that if we look at a complex natural scene on a computer display and make a saccadic eye movement, we are highly insensitive to experimenter-induced changes during the very short interval of the saccade. Buildings and individuals can disappear from the screen or move large distances, yet subjects often report that they see no change. This phenomenon, known as *change blindness*, dispels the naive belief that we "see" all or most things in our visual world (Simons and Levin, 1997; O'Regan, Rensink, and Clark, 1999). Further confirmation of change blindness came from Rensink and colleagues (1997), who showed that if two, almost identical scenes are alternated sequentially with a short time gap between presentations, observers sometimes have great difficulty picking out prominent differences between the two pictures. Mack and Rock (1998) have dubbed this inability to see without

attention *inattentional blindness*. Their inescapable conclusion is that for humans to see an object, we must deploy our attention to it. Thus, it has been argued that vision is an active sense with greater similarities to the exploring hand than to a picture (Gibson, 1966). We must visually grasp an object with our attention in order to perceive it consciously (Nakayama, 1990).

If vision is active and attention is important for seeing, how is it that we do not notice this striking characteristic of vision in our everyday lives? Such "blindness" might suggest that we should often miss very important things. There are many reports of lapsed attention causing industrial and vehicular accidents. Such mishaps are very rare, however, and we do pretty well at staying out of trouble, even if research shows that we are blind to unattended objects. We must have a set of mechanisms that guide our attention effectively. This chapter examines one such mechanism in detail, a short-term memory mechanism that helps deploy attention to features of objects recently attended.

First we will review some current ideas on how visual attention might be appropriately deployed. Much of our everyday world is fairly predictable, and we can learn to direct our attention strategically in accordance with our knowledge and goals. In the case of driving, for example, experienced drivers pay attention to the roadway and the car they are following, but they also attend to cars farther ahead, to driveways, and to upcoming intersections. The full range of such attentional deployments is reached only after extended practice, along with an understanding of the driving environment.

Even when confronting less ordered or less well understood scenes, we are often assisted by low-level mechanisms that ensure that we pay attention to relevant things. The phenomenon of pop-out is one example. In pop-out, an odd feature among a homogeneous array of other elements attracts our attention (Treisman and Gelade, 1980). Most useful in everyday life is the motion of an object. In a large crowd, you can get a friend's attention by waving your arms. Related to motion are sensory transients; the

KEN NAKAYAMA Department of Psychology, Harvard University, Cambridge, Mass.

VERA MALJKOVIC Department of Psychology, University of Chicago, Chicago, Ill.

ARNI KRISTJANSSON Institute of Cognitive Neuroscience, London, U.K.

appearance or blinking of a light hardly ever fails to grab our attention.

Functional role of attention for survival

If attention determines whether or not we see, it should not be surprising that attention is consequential in life-or-death situations. Much has been written on the sensory capacities of animals and their role in determining whether the animal will survive or not (Lythgoe, 1979). The eagle has very good visual acuity, which it uses to detect and identify small prey at large distances. Cats are nocturnal animals and have an abundance of rods, so that they can see under very low illumination. Thus, basic visual capacities are important, and the ecology of animals is clearly related to sensory performance.

Less has been said about the role of attention and other cognitive capacities of animals in determining their population numbers, generally manifested in whether they will catch prey or avoid predation (but see Dukas, 1998). L. Tinbergen (1960), in accounting for the distribution of birds and insects in a Dutch woodland, suggested that the ecological balance between predator and prey depended critically on the cognitive capacities of the predator, in particular its ability to allocate attention effectively. In a much cited article, he suggested that

The intensity of predation depends . . . on the use of specific search images. This implies that the birds perform a highly selective sieving operation on the visual stimuli that reach their retina . . . birds can only use a limited number of different search images at the same time. (Tinbergen, 1960, pp. 332–333)

In support of the search image hypothesis, Dawkins (1971) found that newly hatched chicks, when presented with two different colors of grain, would peck for long runs of one color and then another, almost as if the bird were looking for one color for long stretches at a time, then for another. The search image hypothesis has been a key concept in the study of foraging, explaining the patterns by which animals obtain food successfully. Reid and Shettleworth (1992) demonstrated the power of the search image construct in successfully predicting that pigeons would overselect more frequently presented color grains. As such, the notion of the search image is explicitly mentalistic, a cognitive top-down process invoked to explain behavior. In this chapter we describe a short-term memory phenomenon, more akin to priming, that is well suited to serve the same purpose: to direct attention to features in scenes that have been recently attended to.

Transient attention and its flexible deployment

Before we describe the memory phenomenon, we first need to describe the properties of focal attention itself, in partic-

ular, a fast, transient attentional system that briefly enhances perception and performance (Reeves and Sperling, 1986). In a series of experiments (Nakayama and Mackeben, 1989), subjects were asked to report on the shape and color of a briefly presented probe target in a multiple-element array that appeared within a cue object. The cue appeared at various times (cue lead time) before the target and distracter array. The brief target array was immediately followed by a large, long-duration mask. The cue, the target array, and the mask are shown in figure 29.1A, which also shows the time relations of the various sequentially presented frames. Figure 29.1B shows performance on the discrimination task as a function of cue lead time. Subjects were instructed to fixate the central cross. The targets always fell within peripheral vision, requiring subjects to direct attention away from fixation without moving the eye. Subjects could maintain fixation easily, but just to make sure they did, we monitored eye movements using an infrared detector system. The data indicated by the open circles are from a session during which eye movements were monitored and were found to be neglibible. The solid circles represent data from a session in which eye movements were not monitored. The lack of any eye movement and the same basic form of the function in each case indicate that the pattern of results is reproducible within and between subjects (Nakayama and Mackeben, 1989). The function invariably showed a rapid rise in performance from zero to about 100 ms, which we interpret as the time that attention needs to go to a location away from fixation.

Most surprising and unexpected, however, was the finding of a very reliable drop in performance as cue lead time increased beyond about 100 ms. Performance beyond about 200 ms is well below the peak performance seen for shorter cue lead times. Even though the position of the target location was fully specified for these longer cue lead times and the cue remained on when the target was visible, performance still declined. This finding suggests that something more is at play than the subject's knowledge of the target location, because the strongest manifestation of attention occurred early and then dropped, whereas knowledge of cue location rose quickly and remained high. Supporting this view, Nakayama and Mackeben (1989) found just as large a transient boost in performance when the cued location was the same from trial to trial and thus was always known.

Nakayama and Mackeben interpreted a rapid rise and noticeable fall in performance as a combination of two attentional processes, a weak, sustained component, which is related to the subject's knowledge of the target location, and a stronger transient component, which rises and quickly falls. The magnitude of transient attention is approximately twice that of the sustained component. At first glance, it might seem that this transient component is a reflex, because it is very fast and has a finite duration, and does not correspond to the subject's knowledge of the target location.
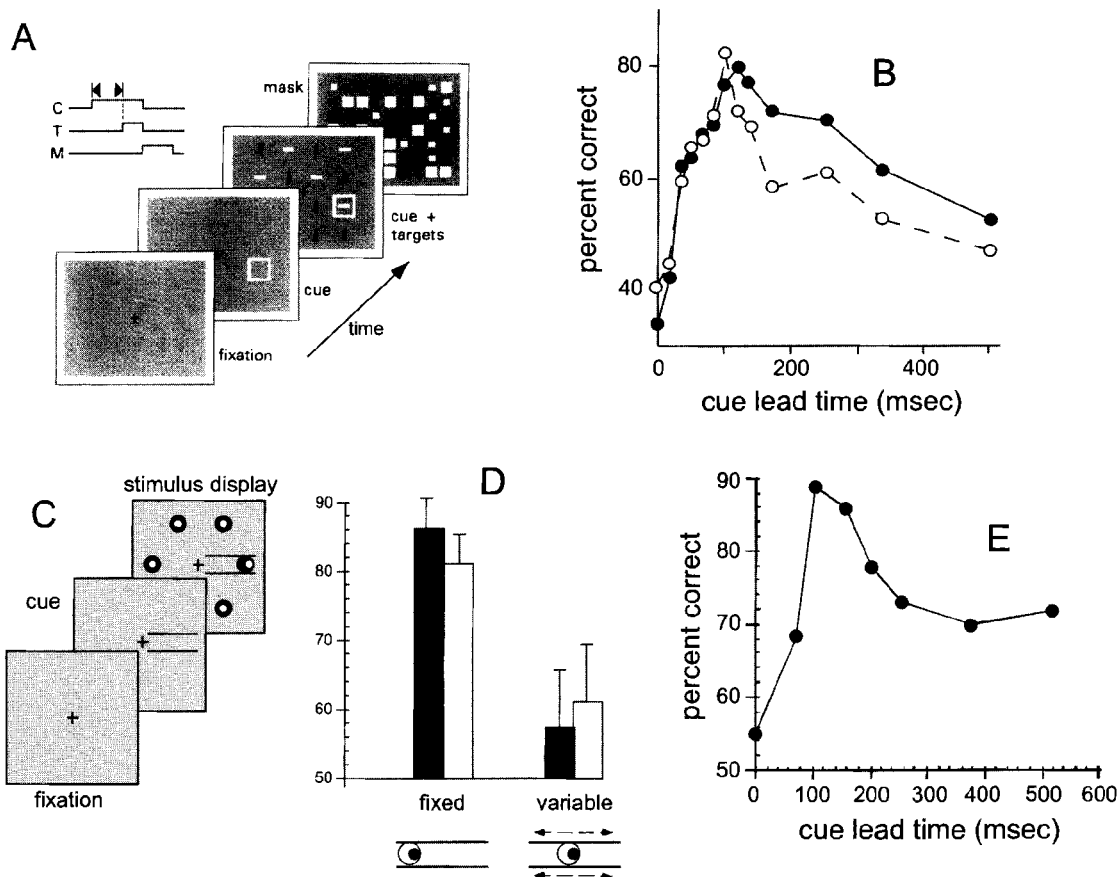
FIGURE 29.1 Transient attention and its flexible deployment. (A) The cuing paradigm originally used to characterize transient attention. A series of frames, fixation, cue, target (with cue and distracters), followed by a mask, is depicted. (B) Performance (percent correct) plotted as a function of cue lead time. Open circles represent data for sessions which eye movements were monitored, solid circles represent data from sessions in which eye movements were not monitored (C) stimulus situation to show the learning of tran-

sient attention. Note that the cue object is larger than the target. (D) Performance as a function of whether the target has a fixed or variable relation to the cue object, that is, whether it is on the same side or varies from trial to trial. (E) Performance in the fixed condition as a function of cue lead time. (Reprinted from Nakayama and Mackeben, 1989; Kristjansson, Mackeben, and Nakayama, 2001.)

Although the transient component does have these characteristics, it is much more flexible and can learn very simple relations. It can show dramatic changes in accordance with past experience.

To document this flexibility, we used an array very similar to the one just described, but with a few changes (Kristjansson, Mackeben, and Nakayama, 2001). First, the target and distracter array were arranged to form a circle (figure 29.1C). The target probe was an "eye" that could look in four possible directions. Most important, the cue was much larger than the target itself, consisting of two parallel horizontal lines within which the eye could be displayed (figure 29.1C). The subject's task was to report in which direction the eye was pointing. If the process of summoning transient attention were a simple reflex, we might expect that the benefits of attention to be bestowed equally within the cue object. There should be no preferred location where the

eye target would be the most visible. To examine this issue, we ran the experiment under two conditions. In the fixed condition, we ran a block of trials during which the target always appeared at one end of the cue (either right or left). In the variable condition, the probe target could move from trial to trial within the cue object on a random basis. A schematic of the fixed and variable conditions is shown below the horizontal axis of the graph in figure 29.1D. In this experiment, the cue lead time was fixed at 100 ms, where it was expected to capture the peak of the transient attention function (as depicted in figure 29.1B).

Again, if the summoning of transient attention were a simple reflex, one might expect the results of the two conditions to be the same. The results shown in figure 29.1D do not support this view. For both observers (one designated by black bars, the other by white bars), there is a large difference in performance between the fixed target case and the

variable target case. With the target fixed, performance is approximately 80%–85% correct, whereas for a variable target, performance is very poor, approximately 50%–60% correct. The poor showing in the variable condition suggests that attention is diluted over the extent of a cue and that only when there is a consistent relation between a specific location within the cue object is there much benefit. We concluded that the difference between the two conditions is the result of experience. It is as if the system responsible for the deployment of transient attention has learned to go to a particular position within the cue object. This result is quite surprising, because the rapid deployment of attention and its transient nature suggest that deployment of transient attention might be a reflex, summoned by an exogenous cue, yet our results suggest that transient attention shows unsuspected flexibility and that it could adapt by going to just one portion of an object on the basis of past experience.

The existence of this unsuspected learning raises many questions that we sought to answer. First, we needed to establish whether the benefit seen for a consistent location within a larger cue (the fixed condition) reflected the learning capacity of transient attention, not sustained attention. Would the flexible deployment of attention (to go to a particular portion of a larger object) show the transient peak? To determine whether this was the case, we used the paradigm described in figure 29.1 C and varied cue lead time. If transient attention were operative, we would expect to see a similar function, a fast rise followed by a fall in performance with increasing cue lead times. Figure 29.1 E shows that this is the case. Performance is highest for intermediate lead times and shows a discernible falloff for durations greater than a few hundred milliseconds. This indicates it is the very rapid process of transient, seemingly reflexive, attention that can learn to go to a specific position within a larger object.

The learning curve

To examine the learning systematically, we conducted experiments to characterize its time course. First, we asked how rapidly the learning might take place. Would it grow slowly and incrementally, or would it occur more suddenly? To characterize how attention might learn to go to a specific place within a larger object, we used the paradigm shown in figure 29.1 C, modifying it slightly. The probe target "eye" could be located on the left or the right end of the cue object, always remaining within the two parallel lines. The position of the target varied from trial to trial but was biased so that on average, there were more repetitions than changes of target location within the cue. Thus, there were streaks of random length where for a number of trials, the target would be on the left side of the cue, followed by a sequence of trials in which the target would be on the right. If learning occurred, we would expect to see some change in per-

formance as the position within a "same-position" streak was taken into account.

To document the hypothesized build-up of learning, we plotted the performance (percent correct) as a function of a trial's position within a streak of same target locations. If learning were taking place, we would expect that as the position within a same-position streak increased, performance would improve. Figure 29.2 A shows that this was the case. As we look at the position within the streak, performance (percent correct) rises dramatically. For the first position within a streak (where the preceding trials had a different target position), performance is very poor, hovering around 50% for the two subjects. Note, however, as the position within the streak increases, there is a conspicuous rise in performance such that by the third repetition, performance has reached asymptote (at about 80% correct). This indicates that the presumed learning did indeed occur, and it occurred very quickly. Therefore, we have shown that the difference seen in figure 29.1 D is very likely the product of rapid learning. In a very few trials, attention can learn to go to a particular part of an object (left or right), boosting performance dramatically.

It seems that we have discovered a malleable process for the directing of focal attention. We have shown that it can be flexibly and rapidly deployed to either end of a larger object. It is an object-relative phenomenon of learning, as the cue object can appear in many places around the circle, as depicted in figure 29.1 C. At this point, it seems natural to determine whether this system can learn something else about objects. Could it learn to go to a portion of an object designated by some other property besides its relative left versus right position? For example, could attention learn to go to a distinctly colored region of an object even if the object changed its orientation, such as being flipped horizontally, as in a right-left reversal? For this purpose, we employed two kinds of distinctions within an object, color and shape. For the color case, the object was a pair of horizontal lines that were colored red at one end and green on the other. For the case of shape, the object had a keyhole form, so that one end was round and the other was angular. Most important for the experiments was the fact that these two features, color and shape, could vary randomly in position, left versus right. Thus, the green end of the cue could be on the left or on the right of the object in a streak of color-constant locations. The insets of figure 29.2 B and C show two possible relations between the cue and target location, where the target is at the green or red end for the color experiment and at the round or angular end for the shape experiment.

The results, shown in figure 29.2 B and C, indicate a dramatic increase in percent correct for the case where the target relation to color or shape repeats. This observation indicates that the transient attentional system can learn to
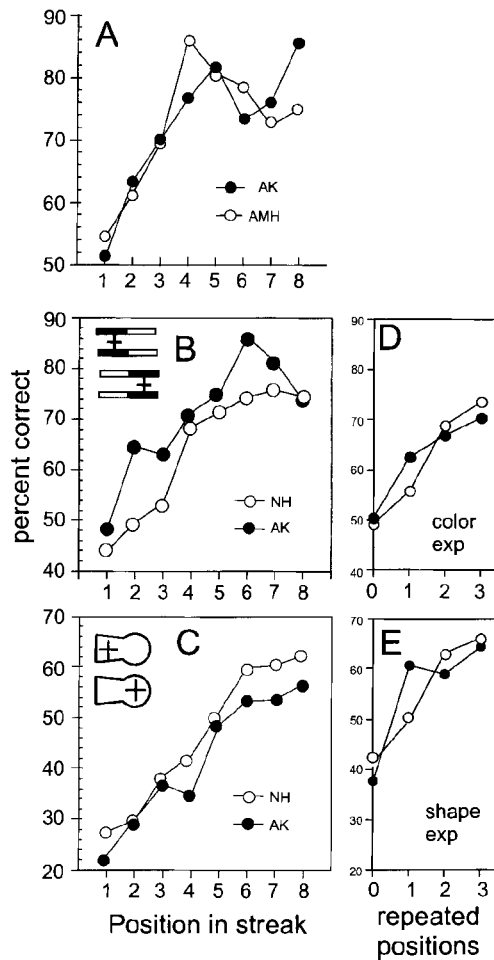
FIGURE 29.2   Learning curves showing the advantage of repeated positions, local colors, and local shapes on learning of mobe objects. (*A*) Effect of repeated position (left or right) as a function of the ordinal position in a streak length. (*B*) Similar effects were obtained for repeated color. Schematic at top left shows the repeated linkage of the target to a specific local color within the larger cue object even as the relation of local colors flips, left to right. (*C*) Similar effects were obtained for repeated local shape. Schematic at top left shows repeated linkage of target to a specific local shape (round vs. angular) over a series. (*D*) Evidence that during color learning, learning of relative position was also occurring. (*E*) Evidence that during shape learning, position learning was also occurring.

go quickly and efficiently to a given color or shape within an object, even for an object that flips horizontally.

Equally important, many of these learning processes can coexist and operate independently and simultaneously. Just as transient attention shows a learning curve for repeated shapes or colors, it also exhibits a simultaneous learning curve for position. Recall that to plot the learning curves for color or shape (as shown in figure 29.2*B*, *C*), we manipulated the sequence, to ensure long streaks of the same shape or color. At the same time, we arranged that probe target position (right or left) would flip randomly as a Bernoulli series, with no intertrial dependencies. By chance, however, such sequences have short runs of repeated target positions. Thus the target can appear on the right for a few trials, then on the left, even as we probe for the effects of longer streaks of color and shape. To show the learning of position, we used the exact same data sets and plotted performance in terms of the number of repeats of the same positions (right or left). These scores are plotted to the right of the color and position learning curves (figure 29.2*E* and *F*, respectively), showing that the experiments originally designed to get the learning curves for color and position also furnish learning curves for repeated position. Each shows the data from the same subject from the exact same series of trials, plotted as a function of number of repeats of the same position (left or right). Although we were only able to obtain reliable data from short streaks with a maximum of three repeats, the same rapid rate of learning is apparent. From this, we conclude that the learning of position, color, and shape occur independently and simultaneously.

It should be pointed out that this very rapid and robust learning takes place without any effort on the subject's part. Subjects are simply instructed to do the best they can in identifying the probe target when it appears within the cue. They are not told that there are streaks of trials during which the target will appear at one end or another, and they are not obviously aware of this relation.

Transient attention cannot learn simple contingent relations

Because learning occurs so easily and automatically, it might represent part of a specialized, dedicated system that lacks the flexibility of more general learning systems. To address this issue, we made the learning requirements slightly more complex. We ask whether attention could go to opposite ends of objects that are obviously distinct. Would it be possible for attention to go quickly to the left end of object A and to the right end of object B? We conducted two experiments using object properties (color and shape) that have already been shown to be used by this learning system (as demonstrated in figure 29.2*B*, *C*). For the shape experiment there were two objects, one rounded, one rectangular. In the
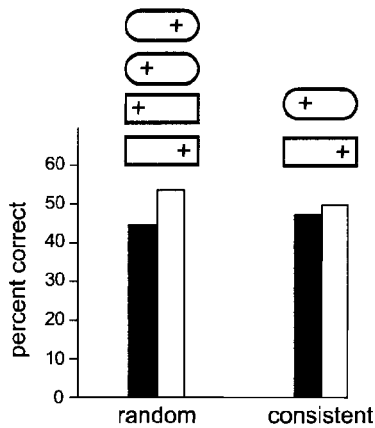
Figure 29.3 Demonstration that transient attention cannot learn to go to the left of one object and the right of another. Two very distinguishable cue objects, one rounded and the other rectangular, have either random or consistent cue-target relations. There is no advantage to consistent presentations, indicating the lack of learning of this relation. (From Kristjansson and Nakayama, 2003.)

experimental condition, which we call "consistent," the target was always on the left side if the object was rounded and on the right side if the object was rectangular. Schematics of these objects and this contingency relation are shown in figure 29.3. For comparison we had a random control condition in which there was no specific assignment of the target to the ends of particular objects. If transient attention could learn to go to a specific place within a given object, then we should expect to see improved performance in the consistent condition. The results as depicted in figure 29.3 show no difference in performance between the two conditions. Even though the two shapes are highly discriminable, there is no sign of learning. The attentional system, or at least this fast transient system, cannot learn this simple relation. We also found that there was no learning when we used color to distinguish the objects (Kristjansson and Nakayama, 2003). Transient attention could not learn to go the right of a red object and to the left of a green one.

At this point we should pause and reflect on the nature of the learning that we have characterized so far. The learning is very rapid, taking place over very few trials, but surprisingly powerful, going from near chance to reliable performance and doing so hundreds of milliseconds faster. Yet the suprising lack of the ability to learn contigent relations suggests that the learning is not registering information about the nature of the objects, but something more primitive. The learning is only concerned with the recurring features of objects, not with the objects themselves.

Learning to deploy attention in a visual search array

At the beginning of this chapter we reviewed the concept of a search image, according to which construct an animal

searching for food possesses a top-down template specific to the items sought. This construct could explain otherwise puzzling findings of Dawkins (1971), indicating that animals have runs of grain types in a free choice situation and often overselect for the more prevalent food type (Reid and Shettleworth, 1992).

Our experimental findings on cued attention in humans (Nakayama and Mackeben, 1989; Kristjansson, Mackeben, and Nakayama, 2001; Kristjansson and Nakayama, 2003) might seem only distantly related to these issues of search and foraging, yet the learning mechanism that we have delineated provides an alternative explanation, replacing the notion of the search image. To make this view more plausible, we will now describe the results obtained with a visual search paradigm in which the learning of attentional deployment is also evident. This paradigm also allows us to more readily make the case that the learning mechanism we have outlined might apply to searching and foraging. As originally derived from the work of Bravo and Nakayama (1992), the goal of the visual search paradigm was to develop a visual search task that would require the subject to deploy focal attention to an item in an array. For this purpose, we devised a task in which the subject had to discriminate a subtle shape difference of an oddly colored singleton that was presented along with two other colored distracters (Maljkovic and Nakayama, 1994). Thus, the subject's task was to report the shape of the odd-colored target while maintaining fixation on a central cross (figure 29.4A).

Maljkovic and Nakayama (1994) manipulated the length of runs of same-color odd targets and found a dramatic decrease in reaction time as the target's position within a streak increased. Figure 29.4B shows the results obtained in a representative observer. The reaction time fell systematically as the number of repetitions within a streak of same-color targets increased. So, just as in the attentional cuing experiments shown in figure 29.2, there is a strong benefit derived from repeating the color of the object of attention. This observation confirms the existence of a comparably rapid learning process in a very different paradigm and suggests that the results obtained with our cued attention experiments can be more widely generalized.

Because the experiment is more like a visual search task, we can now ask whether the results might be accounted for by a process akin to that involved in the search image. Does the subject form a mental template that aids the perception of targets that fit this template or search image? The implication in the search image literature is that a search image is adopted as a formed and structured template (corresponding to the desired target), which would explain the long streaks of trials of a given food selected (Dawkins, 1971) and the overselection for common colored foods (Reid and Shettleworth, 1992). A search image also suggests an active process, one that reflects the expectancy of the upcoming
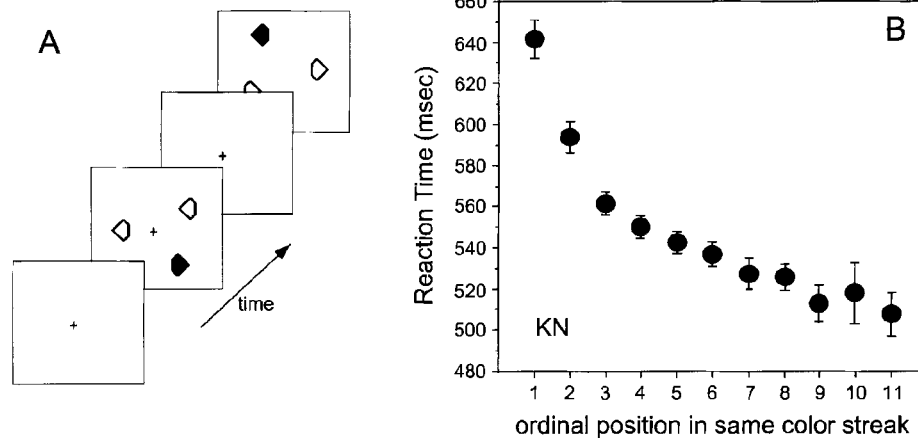
FIGURE 29.4 The learning of attentional deployment in a visual search task. (A) In a sequence of trials, the observer's task was to identify the shape of an odd-colored target. The color of the target could either remain the same or change, as could the target's position. (B) Reaction time as a function of target position in a same-color target streak. (From Maljkovic and Nakayama, 1994.)

target color. After a long run of same-color trials, expectancy for yet another similar trial would be high. Such a notion could plausibly explain our data, at least those shown in figure 29.4*B*.

Our aim, however, is to formulate a more mechanistic alternative to the search image hypothesis. We argue that the learning could consist of a passive association of isolated characteristics of previously attended sites and the "response" of attentional deployment, explicitly adopting the active hand metaphor for attention. In this scheme, expectancy plays no role. Attention is driven by its association with simple features. This accords with the piecemeal quality of the learning described for cued attention, where the learning of position (left or right), color, and shape occur independently and in parallel. The effects of learning are not overcome by knowledge, expectancy, or intention.

To evaluate these two different notions, we devised an experiment in which the predicted results from an expectancy hypothesis would be very different from those based on passive association. We used a deterministic sequence of trials that provided ample opportunity for both expectancy and passive learning to be evident. The sequence consisted of one pair of a set of "same" color trials, followed by the other pair of "same" color trials. Thus, all sequences had run lengths of two: targets were presented as two greens, two reds, two greens, and so on. The experiment was run under two sets of instructions. In the first set, we asked the subjects simply to do the experiment and not think about the pairs of same-color trials. We called this the passive condition. In the second case, we asked subjects to be fully aware of the completely predictable nature of this sequence and asked that they actively anticipate the color of each upcoming trial by subvocally voicing the upcoming color prior to each trial. We called this the active and expectant condition.

The results are shown in figure 29.5. In the case of the passive condition, with no conscious effort on the part of subjects to anticipate the color of the upcoming trial, there was clear evidence of learning. Reaction times were significantly shorter for the second trial of the sequence, confirming the same sequence of effects as shown in figures 29.2 and 29.4. Most critical are the results of the active condition. Expectancy is plainly evident, because the sequence is both simple and deterministic. Furthermore, expectancy should be the same for all positions within the sequences. Thus, we predict that the difference between the first and second trial of the sequence should be abolished. The lower curves in Figure 29.5 shows that this is not the case; the difference between trial 1 and trial 2 for the active condition is as great as that in the passive condition, even though the expectancies for trial 1 and trial 2 of the sequence should be identical.[1] This experiment shows that performance cannot be influenced by prior knowledge or expectancy. Whether a subject knows, pays attention to, or completely ignores the sequence of target colors is irrelevant. The process of learning proceeds independently of these higher-order cognitive factors. As such, the process of rapid attentional deployment is autonomous and machine-like, impervious to other mental processes and reflecting only passive repetition.

An elemental memory event

These results suggest that with each attentional deployment to a given feature (color, in this case), there is a primitive

[1] The overall reaction times are consistently faster for this active condition, perhaps reflecting the higher motivation of the subject, given the additional task of anticipating the color of the upcoming trial.
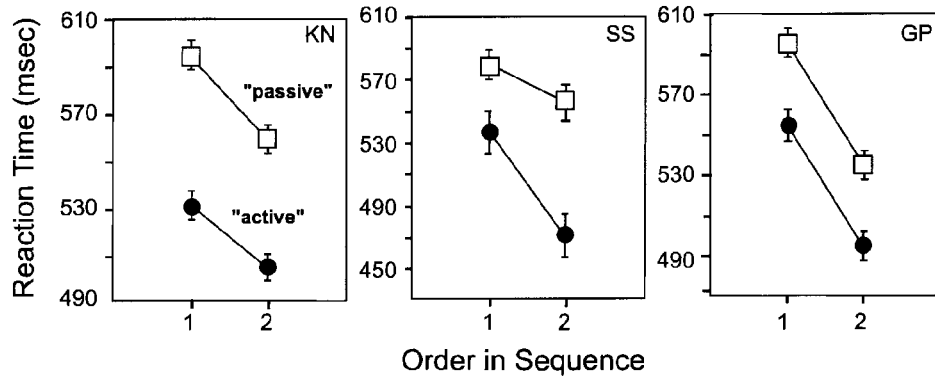
FIGURE 29.5 Irrelevance of expectancy, knowledge, and intention. Using the same stimulus display as shown in figure 29.4, a deterministic series of trials was presented, doubly alternating red and green so that the streak length was always two. Subjects could either passively just do the task (passive) or could consciously anticipate each trial (active). The learning was the same in each case. (From Maljkovic and Nakayama, 1994.)

process that effectively associates the deployment of attention with this feature and this process can be incremented over long sequences to boost performance significantly. How might such performance enhancement occur? One way to examine this is to ascertain the influence of a single episode of attentional allocation, to see how it, by itself, might influence future allocations. To explore such a hypothesized associative process in isolation, we devised a method to measure the influence of a single trial in the past, and thus to determine its influence on the future. We used the same search task schematized in figure 29.4A. As mentioned earlier, if the learning of attentional deployment is the result of the simple association of a feature to a past attentional deployment, we might be able to isolate this process and track its time course. Thus, our goal was to measure the influence of arbitrarily distant trials in the past. In the simplest case, we would expect to see an influence from the immediately preceding trial. If the previous trial's target color was the same, we would expect reaction times to be shorter than if the previous color was different. We can generalize this question beyond the consideration of a previous trial and consider the influence of any trial in the past. Because we used a Bernoulli sequence of target colors, such that each trial's target color was independent of all others, such performance differences can provide unbiased estimates for any given trial's influence. Computing this difference for up to 15 trials in the past and for a smaller number in the future (to gain a sense of noise of our measurements), we obtained a memory kernel function, the influence of such a single trial (Maljkovic and Nakayama, 1994). Figure 29.6 shows this function for two observers. The vertical dashed line separates past from future. It is apparent that the immediate previous trial has the largest influence on the current trial, such that reaction times are approximately 40–50 ms faster for same versus different color targets. But there is also is considerable residual influ-

ence from single trials much more distant in the past, as indicated by the extent to which this function deviates from the baseline of zero difference. This means that any given trial will have a lasting influence up to about six to eight trials into the future. Yet this influence, while very strong, does ultimately decay, and after a dozen or more trials it is effectively abolished.

Subsequent research has found that this decay is event driven: it is the intervening allocations of focal attention, not
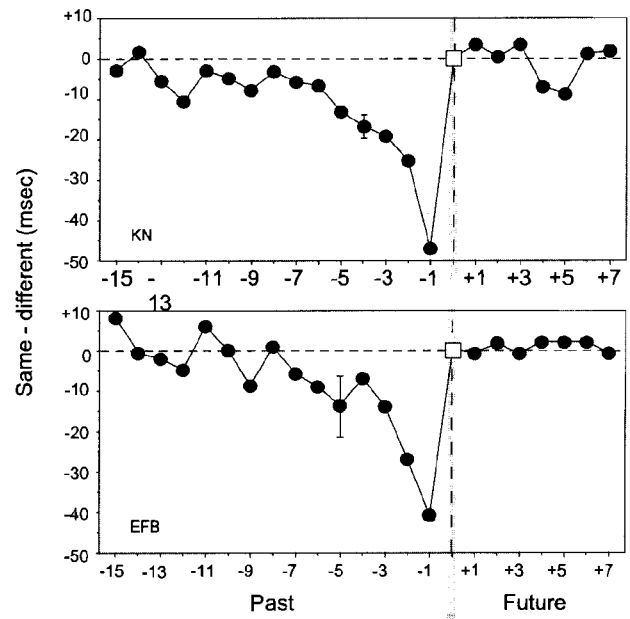

FIGURE 29.6 Influence of a single trial or event for two observers. Using the same stimulus display as shown in figure 29.4A, and with a series of random colored targets presented over time (a Bernoulli sequence), we showed the influence of past (or future) trials on the present. (From Maljkovic and Nakayama, 1994.)

simply time, that cause the memory to decay (Maljkovic and Nakayama, 2000). In addition, we also found that any deployments of focal attention to unrelated targets (neither red nor green) would also deplete the memory.

Summary of findings

We have demonstrated the existence of a very powerful attentional system that we have called transient attention, a brief heightening of visual processing that seems to require an exogenous cue. Its benefits are much stronger than those of sustained attention, and it seems curiously uncoupled from higher-order knowledge of a target's upcoming location. As such, it seems to have characteristics very similar to a reflex. Yet it is a flexible reflex. It can learn to go to a specific portion of a larger object, identified by color, shape, or position, and the associations formed are piecemeal, each is independent, and all occur simultaneously. But the piecemeal nature of learning has inherent bounds. Transient attention cannot learn to go to the left end of one identifiable object and to the right end of another. Thus the learning has a nonintegrative quality: it can learn the simplest first-order associations, but no more. Within these bounds, the learning is very strong, and its effects cannot be overcome by conscious knowledge of upcoming target properties. Finally the memory kernel functions (see figure 29.6) show that the learning from a single trial is temporary, decaying over a variable period up to six to eight trials.

The basic findings described here have been replicated by others (Hillstrom, 2000; Goolsby and Suzuki, 2001; Olivers and Humphreys, 2003). They have also been extended to eye movements and to manual pointing. McPeek, Maljkovic, and Nakayama (1999) showed the phenomenon in human saccadic latencies, and the same phenomenon has also been demonstrated for monkey saccades (McPeek and Keller, 2001; Bichot and Schall, 2002). Song and Nakayama (2003) showed that manual pointing with the finger exhibited the same benefit of feature repetition.

Discussion: Comparison with other forms of learning and possible neural substrates

Our findings indicate that with each allocation of attention, there is an increase in the association to those features that were at the site of attentional deployment. The influence from a single allocation of attention can summate, so that after a few repetitions of deployment of attention to particular features, performance is greatly enhanced. The diagram in figure 29.7 is a pictorial representation of the process suggested. After each attentional allocation to a given feature on an object, there is a transient boost in memory that will gradually fall off with increasing deployments of attention (figure 29.7A). With repeated deployments of attention to
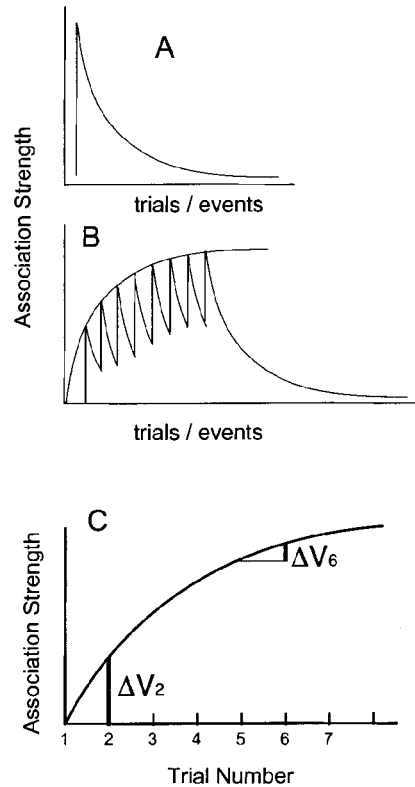

FIGURE 29.7 Contrasting of the learning of transient attention with other types of learning processes. A reproduces the memory kernel function in schematic form, while B shows how it is constant and summates. In C, we contrast this the with acquisition of more conventional learning, where each increment in learning decreases as learning proceeds. B and C show the negatively accelerating learning curves common to both transient attention and other forms of learning, respectively.

the same feature in a streak, such effects combine, so that effects can be quite large (figure 29.7B), accounting for the improved performance.

Superficially, there seems to be some similarity to all other forms of learning, including classical conditioning. In particular, the graded and incremental nature of the learning seems identical. With repeated deployment of attention to the same features over time, this system displays the familiar negatively accelerating learning curve (see figures 29.2, 29.4B) seen in most studies of learning and conditioning (figure 29.7C). Yet from our results, it is evident that the underlying process is likely to be very different.

Current theories of associative learning, such as Rescorla and Wagner's (1972) model, predict similar negatively accelerating learning functions. These formulations are based on a delta rule, according to which the increment of learning at any instant is variable, decreasing during the course of learning. This occurs as the amount of unused associative capacity is exhausted. The reduced increment in learning

strength for such types of learning is depicted in figure 29.7C, which shows incremental reductions on later trials. This situation is to be contrasted to the learning of attentional deployment, where it is likely that the memory increment is constant for each attentional deployment and the negatively accelerating function is due to quick decay of the elementary memory event, which precludes summation beyond its own, very limited duration (figure 29.7B). Thus, the very temporary nature of the learning and its inevitable decay mark it as very different from more traditional forms of learning.

Another difference is the piecemeal nature of the learning. This was evident with the independent influence of position, color, and shape associations in facilitating speeded attentional cuing. They combine similarly for the visual search task (Maljkovic and Nakayama, 1996) and add linearly with separable effects (Norton and Maljkovic, 2001). To provide a framework to understand the independent nature of the underlying components, we suggest that the memory system is a distributed one, with separate associations to many simple features. In keeping with the analogy of attention being like a mobile hand, we depict the association to features as either positive or negative. For example, studies by Maljkovic and Nakayama (1994, 1996) show that there are equally strong inhibitory effects, such that colors and positions not attended to on a given trial are more slowly attended to in the future. Because the associations are to separate features, it is thus not surprising that the learning cannot be extended to anything of higher complexity. Thus it cannot learn, for example, to go to the right end of object A and the left end of object B.

We have described the rapid learning of transient attention as associative because from a behavioral perspective, stronger associations between particular features and the act of attentional deployment are evident. Yet if we are to think of possible neural mechanisms underlying such learning, we need to be cautious in drawing on known mechanisms of associative learning and memory. Most obvious is the very temporary nature of the learning. Being temporary, it does not require the cellular mechanisms of memory storage that are longer lasting. Consequently, the mechanisms of conditioning and other forms of permanent learning are unlikely to be applicable. The learning described here is also very automatic and noneffortful. This suggests that the learning is very different from the concept of working memory, and Maljkovic and Nakayama (2000) showed that indeed it is, that the properties of working memory cannot account for our results.

A very different type of process is required to explain the flexibility seen here. We suggest that phenomenon described should be regarded as modulatory, activating or suppressing the specific neural circuits that have been activated very recently. As such, those processes that enabled attention to

go to particular portions of an object are facilitated or inhibited temporarily. Seen in this light, the learning can be considered more as a form of priming of existing associations. It is not a full-blown learning system by itself but a process that can modulate existing associations.

Where in the nervous system might such learning or modulation of associations occur? We suggest it occurs at some level removed from the early retinotopic stages of the visual system. This is mandated by the object-relative nature of the learning, that is, the fact that attention can learn to go to the left or right part of an object no matter where it appears in the visual field (see figures 29.1, 29.2; also Maljkovic and Nakayama, 1996).

That the learning is nonintegrative, piecemeal, and distributed suggests that at any given instant, the learning could occur across many different objects, and is not linked to specific stored object representations. This situation would suggest that the learning we have described is not found in parts of the cortex that represent specific objects. It argues against the inferotemporal cortex as the site of this particular type of learning.

Parts of the brain subserving the allocation of attention and eye movements are of particular interest, including frontal, parietal, and midline cortical areas (Posner and Raichle, 1994). Olson and colleagues (Olson, Olson, and Gettner, 1993; Olson, 2003) have shown that there are neurons in the supplementary area of the frontal cortex that fire before a monkey makes an eye movement to a specific position within an object, independent of its visual field position. Neurons in these frontal regions are specialized, so that some fire just before the monkey looks to the left part of an object, whereas other neurons fire preferentially just before a monkey looks to the right part. Because eye movements are preceded by shifts in focal attention (Kowler et al., 1995), the very specific properties of these object-relative neurons suggest that they could be the possible mediators of the learning effects described here.

More direct evidence regarding the likely locus of learning comes from the work of Bichot and Schall (2002). They characterized the same temporary learning of repeated target colors measuring saccadic eye movement latencies in the monkey. Most important, they showed the expected changes in frontal eye field movement neuronal firing. Such cells are strong candidates to mediate the effects that we have described here. It remains for future research to delineate the exact locus of learning, whether it is confined to these neurons themselves or to their specific inputs.

We now consider other types of learning that could be related. Of interest is the phenomenon of worsened performance when a subject changes from one task to another (task switching). The decrement in performance associated with task switching could be the flip side of short-term temporary learning. For example, Allport and Wylie (2000) report large

costs of task switching, with performance restored to optimum asymptotic values only after three to eight repetitions. They found this under a very wide range of tasks and interpreted the phenomenon as a form of temporary associative learning between particular stimuli and responses. With these diverse tasks and task changes, it is not evident that all of these examples fall in the same category of learning that we have outlined. All are unlikely to be so easily explained by the learning of attentional deployment. In fact, it is clear that sequence effects can occur that are not directly tied to the learning of attentional deployment but are more related to surface formation or perceptual grouping (McCarley and He, 2001; Kristjansson, Wang, and Nakayama, 2002). As such, rapidly acquired temporary learning may be a very widespread phenomenon, not confined to the deployment of attention. We argue that it deserves much more investigation with respect to its scope and possible cellular mechanisms.

Short-term memory and priming as an alternative to search imagery

What advantages does our phenomenon have for the situation described earlier in the chapter, that of foraging? The short-term memory system could provide a more mechanistic replacement for the more cognitive couse of the search image. The properties of the temporary learning we have described are not those of the search image posited by Tinbergen (1960), but the two constructs may serve the same purpose. Rather than calling for a "highly selective sieving operation," we suggest a collection of passive features associatively linked to the act of attentional deployment. As soon as attention is deployed somewhere, it becomes more probable that it will be deployed to objects that share the same features or positions. This could explain much of the data supporting the the search image concept, especially the long runs of selecting foods with a certain feature (Dawkins, 1971) and the overselection of certain features with greater than base probability (Reid and Shettleworth, 1992). All of these behaviors could be explained by the more piecemeal temporary associations seen here. And because the learning is very temporary, there is no risk that the animal will be caught perserverating on a feature that was attended to a while back.

In a similar vein, the passive and piecemeal modulatory process that we have described might also supplant, at least in part, theories postulating top-down processes for visual search. For example, the guided search theory of Wolfe and colleagues (1989) has similarities to the notion of the search image by employing a top-down process to explain a variety of observations about a visual search, many of which observations could also be accounted for by the more mechanistic processes described here.

REFERENCES

ALLPORT, A., and G. WYLIE, 2000. Task switching, stimulus-response bindings and negative priming. In *Attention and Performance*, S. Monsell and J. Driver, eds. Cambridge, Mass.: MIT Press.

BICHOT, N. P., and J. D. SCHALL, 2002. Priming in macaque frontal cortex during popout visual search: Feature-based facilitation and location based inhibition of return. *J. Neurosci.* 22:4675–4685.

BRAVO, M., and K. NAKAYAMA, 1992. The role of attention in different visual search tasks. *Percept. Psychophys.* 51:465–472.

DAWKINS, M., 1971. Shifts of "attention" in chicks during feeding. *Anim. Behav.* 19:575–582.

DUKAS, R., 1998. Constraints on information processing and their effects on behavior. In *Cognitive Ecology*, R. Dukas, ed. Chicago: University of Chicago Press, pp. 89–127.

GIBSON, J. J., 1966. *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin.

GOOLSBY, B. A., and S. SUZUKI, 2001. Understanding priming of color-singleton search: Roles of attention at encoding and "retrieval." *Percept. Psychophys.* 63:929–944.

GRIMES, J., 1996. On the failure to detect changes in scenes across saccades. In *Perception. Vancouver Studies in Cognitive Science*, vol. 5, K. Akins, ed. New York: Oxford University Press, pp. 89–109.

HILLSTROM, A. P., 2000. Repetition effects in visual search. *Percept. Psychophys.* 62:800–817.

KOWLER, E., E. ANDERSON, B. DOSHER, and E. BLASER, 1995. The role of attention in the programming of saccades. *Vision Res.* 35:1897–1916.

KRISTJANSSON, A., M. MACKEBEN, and K. NAKAYAMA, 2001. Rapid, object-based learning in the deployment of transient attention. *Perception* 30:1375–1387.

KRISTJÁNSSON, Á., and K. NAKAYAMA, 2003. A primitive memory system for the deployment of transient attention. *Percept. Psychophys.* 65:711–724.

KRISTJANSSON, A., D. WANG, and K. NAKAYAMA, 2002. The role of priming in conjunctive visual search. *Cognition* 85:37–52.

LYTHGOE, J. N., 1979. *Ecology of Vision.* New York: Oxford University Press.

MACK, A., and I. ROCK, 1998. *Inattentional Blindness.* Cambridge, Mass.: MIT Press.

MALJKOVIC, V., and K. NAKAYAMA, 1994. Priming of popout. I. Role of features. *Mem. Cogn.* 22:657–672.

MALJKOVIC, V., and K. NAKAYAMA, 1996. Priming of pop-out. II. The role of position. *Percept. Psychophys.* 58:977–991.

MALJKOVIC, V., and K. NAKAYAMA, 2000. Priming of popout. III. A short term implicit memory system beneficial for rapid target selection. *Vis. Cogn.* 7:571–595.

McCARLEY, J. S., and Z. J. HE, 2001. Repetition priming of 3-D perceptual organization. *Percept. Psychophys.* 63:195–208.

McCONKIE, G. W., and D. ZOLA, 1979. Is visual information integrated across successive fixations in reading? *Percept. Psychophys.* 25:221–224.

McPEEK, R. M., and E. L. KELLER, 2001. Short-term priming, concurrent processing, and saccade curvature during a target selection task in the monkey. *Vision Res.* 41:785–800.

McPEEK, R. M., V. MALJKOVIC, and K. NAKAYAMA, 1999. Saccades require focal attention and are facilitated by a short-term memory system. *Vision Res.* 39:1555–1566.

NAKAYAMA, K., 1990. The iconic bottleneck and the tenuous link between early visual processing and perception. In *Vision: Coding and Efficiency*, C. Blakemore, ed. Cambridge, Engl.: Cambridge University Press, pp. 411–422.

NAKAYAMA, K., and M. MACKEBEN, 1989. Sustained and transient components of focal visual attention. *Vision Res.* 29:1631–1647.

NORTON, J., and V. MALJKOVIC, 2001. Accumulation and decay properties of implicit short-term memory suggest separate mechanisms code object features and position. *Soc. Neurosci. Abstr.* 27,76.5.

OLIVERS, C. N. L., and G. W. HUMPHREYS, 2003. Attentional guidance by salient feature singletons depends on intertrial contingencies. *J. Exp. Psychol. Hum. Percept. Perform.* 29:650–657.

OLSON, C. R., 2003. Brain representation of object-centered space in monkeys and humans. *Annu. Rev. Neurosci.* 26:331–354.

OLSON, C. R., and S. N. GETTNER, 1995. Object-centered direction selectivity in the supplementary eye field of the macaque monkey. *Science* 269:985–988.

O'REGAN, J. K., R. A. RENSINK, and J. J. CLARK, 1999. Change-blindness as a result of "mudsplashes." *Nature* 398:34.

POSNER, M. I., and M. E. RAICHLE, 1994. *Images of Mind.* New York: Scientific American Library.

REEVES, A., and G. SPERLING, 1986. Attention gating in short term visual memory. *Psychol. Rev.* 93:180–206

REID, P. J., and S. J. SHETTLEWORTH, 1992. Detection of cryptic prey: Search image or search rate? *J. Exp. Psychol. Anim. Behav. Proc.* 18:273–286.

RESCORLA, R. A., and A. R. WAGNER, 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In *Classical conditioning.* Part II, *Current research and theory,* A. Black and W.F. Prokasy, eds. New York: Appleton-Century-Crofts, pp. 64–99.

RENSINK, R. A., J. K. O'REGAN, and J. J. CLARK, 1997. To see or not to see: The need for attention to perceive changes in scenes. *Psycholog. Sci.* 8:368–373.

SIMONS, D. J., and D. T. LEVIN, 1997. Change blindness. *Trends Cogn. Sci.* 1:261–267.

SONG, J.-H., and K. NAKAYAMA, 2003. The role of focal visual attention in a manual pointing task [Abstract]. *J. Vision* 3:264a.

TINBERGEN, L., 1960. The natural control of insects on pinewoods. I. Factors influencing the intensity of predation by songbirds. *Arch. Neerland. Zool.* 13:265–343.

TREISMAN, A. M., and G. GELADE, 1980. A feature-integration theory of attention. *Cogn. Psychol.* 12:97–136.

WOLFE, J. M., K. R. CAVE, et al., 1989. Guided search: An alternative to the feature integration model for visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 15:419–433.