Simons, D. J., Mitroff, S. R., & Franconeri, S. L. (2003). Scene perception: what we can learn from visual integration and change detection. In M. Peterson & G. Rhodes (Eds.), *Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes* (335-355). Oxford University Press.

12

Scene Perception

What We Can Learn from Visual Integration and Change Detection

DANIEL J. SIMONS, STEPHEN R. MITROFF,
AND STEVEN L. FRANCONERI

As we view our world, each eye movement causes a dramatic shift in the retinal position of every object in a scene, every blink interrupts visual sensation, and moving objects occlude other objects. Somehow our perceptual system makes sense of this varying information and we experience a stable visual world. How do we sense stability in the face of such disruptions? What sorts of representations underlie this experience?

Scenes can be described in terms of the quantity and quality of the information available for representation. To adequately describe the nature of the representations underlying our experience, we must ascertain both how many items in a scene are retained and how much detail about each item is retained. Although these dimensions are not necessarily independent, after all, both are forms of visual information. The distinction will prove useful in characterizing the possible inferences we can draw from studies of visual representations. Almost all models of visual stability concur that some information from a scene is preserved from one view to the next, but such models vary dramatically in their proposals for both the *completeness* and the *precision* of these representations. That is, the proposed representations vary in how many items they include and how much detail about those items is retained. A complete representation is one that includes all the items in a scene. Note, though, that a complete representation could be as vague as a brief verbal description of the items or as detailed as a photographic representation of infinite precision. Maximally precise representations preserve all of the information about encoded items, although not all items in the scene would necessarily be encoded. Representations can be complete, but imprecise as well as precise, but incomplete.

Many early explanations for our experience of visual stability assumed the need for representations that are both complete and precise replicas of the world. Ancient Greek models assumed that any demonstration of contact between the self (or soul) and objects in the external world was sufficient evidence for such a representation of the world. For example, Democritus was satisfied that accurate perception was possible because he noted the complete reflection of the world on the pupil (Lindberg, 1976). Such models simply assumed that representations were inherently accurate and largely ignored discrepancies between the visual world and its retinal projection. Given this assumption, they also did not consider the need to adjust representations in the face of eye movements. In contrast, medieval explanations (e.g., those of Al-Kindi, Kepler), which also accepted the importance of storing a complete and precise representation of the world, noted the existence of disparities between the retinal image and the world. These theorists searched for mechanisms that could correct for such discrepancies (Lindberg, 1976). More recently, Descartes proposed the need for a correction to produce an accurate representation of the world in the motions of the pineal gland (see Gibson, 1966; Van Hoorn, 1972) and Helmholtz argued that unconscious inference (or cerebration) allows accurate representations of the three-dimensional world from the roughly two-dimensional retinal image (Helmholtz, 1866/1924).

In contrast to these beliefs in a complete and precise internal representation, other theorists argue that our internal representations minimize the information preserved from one glance at a scene to the next (e.g., Irwin, 1991; Rensink, 2000a). These models argue that our experience of a stable, continuous visual world derives, not from a complete internalization of the world, but instead from the *absence* of such a representation. If our visual system simply assumes that the world is stable and unchanging, then there is no need to represent it internally in order to experience it as stable; the world itself can serve as a memory for most visual details (O’Regan, 1992). Only those items that are needed for action from one view to the next (Hayhoe, 2000) or that are the current focus of attention (Horowitz & Wolfe, 1998; Rensink, 2000a) need to be preserved across views. These models vary in the degree of precision in the representations of this small set of items, some positing relatively precise representations of many features (Luck & Vogel, 1997) and others arguing for more abstract representations (Irwin, 1991).

To distinguish between all of these contrasting models of the completeness and precision of our representations, we must determine our capacity to store and integrate visual details from one view to the next. The clearest empirical evidence suggesting the existence of complete and precise representations comes from work on visual pattern masking. Pattern masking occurs when normal processing of one stimulus is disrupted by the presentation of a second stimulus. Without a masking stimulus, briefly presented visual stimuli persist with great precision in an iconic memory for up to 300 msec after the stimulus has been removed (Sperling, 1960). However, if a mask appears shortly after the stimulus and prior to the report, the stimulus and the mask are integrated and further independent processing of either the stimulus or the mask is impaired (Breitmeyer, 1984; Kahneman, 1968). Optimal masking occurs when the mask shares features in common with the target stimulus such that when the two stimuli are superimposed, observers cannot easily separate them. The fact that the mask is most effective when it has the same features as the target suggests that relatively detailed information about those features is preserved, at least for a brief period.

A more stringent test of the completeness and precision of the representation requires observers to visually integrate two displays and then to report the emergent result. In a classic visual integration task, observers viewed two successive, complementary 12-dot patterns drawn from a 5 × 5 array. When these 12-dot arrays were integrated into a single 24-dot array, only one location in the 5 × 5 array never contained a dot. When presented sequentially, success in locating the empty location depends on integrating the two patterns over time. Consistent with the integration occurring during pattern masking, observers can readily perform this task when the two images appear successively at the same location in space and are projected to the same retinal position (i.e., when observers are fixating; Di Lollo, 1980).

The existence of such precise and complete representations raises the possibility that our experience of visual stability arises from the merging of consecutive views of a scene into a single representation. However, the perceptual demands of a masking experiment are quite different from those of the real world. As we view scenes in the real world, we move our eyes (saccade) 3–4 times each second and visual information is processed primarily when the eyes are stable during a fixation—relatively little new visual information is processed during saccades (e.g., Campbell & Wurtz, 1978). As the eyes move, the projection of each object in the world shifts to a new part of the retina, in contrast to the relatively stable retinal projections in visual masking experiments. Consequently, even if our visual system does store a complete and precise representation in an iconic memory, visual integration with subsequent fixations would require an adjustment for such eye movements so that the two representations would be aligned. That is, the visual system would need to create a single representation of the world based on many different images on the retina; an object occupying a single location in space, but projected to different locations on the retina, needs to be perceived as the same object.

The primary purpose of this chapter is to consider the nature of the

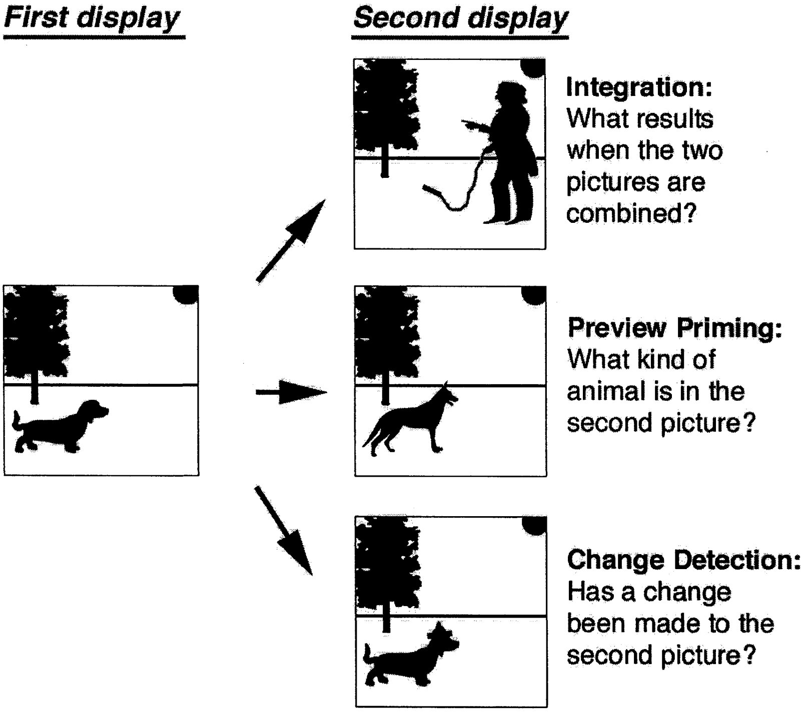

FIGURE 12.1. A cartoon illustration of three methods used to study the nature of the information preserved across views.

information retained from one view to the next. Does the visual system retain a precise and complete representation and integrate it across a saccade? Or, does our experience of stability derive from the absence of such representations? We address these questions by exploring evidence from two distinct, but related lines of research: (1) studies of visual integration and facilitation across eye movements and (2) studies of change detection. Most studies of visual integration focus entirely on the consciously accessible products of visual integration. Given that measures dependent on explicit awareness might well underestimate the amount of information preserved across views, we also examine evidence from more sensitive implicit measures of integration such as cross-saccade priming. Priming studies infer the nature of the information preserved across views by measuring the influence of information presented on one view on processing of the subsequent view (see figure 12.1 for a schematic illustration of these methods of studying the representations preserved from one view to the next).

Following our discussion of visual integration, we provide a more thorough discussion of evidence from studies of change detection. If observers can detect a change from one view of a scene to the next, then they must have represented some information from the first view and compared it to the subsequent view. As in studies of visual integration, most studies of change detection rely on reports of the consciously accessible product of the comparison process as a dependent measure. Given that explicit reports might well underestimate the precision and completeness of the preserved information, we also evaluate the possibility that more sensitive, implicit measures might reveal evidence for change detection in the absence of awareness. In the final section, we speculate about the sorts of information that might be expected to be preserved across views, and we argue that in the absence of awareness, representations likely are limited to those aspects of our visual world that are needed from one instant to the next.

Visual Integration Across Saccades

In order for visual integration of complete and precise representations to underlie our experience of visual stability, observers must be able to combine information from one fixation to the next, and the visual system must somehow account for the corresponding change to the retinal projection. In one intuitive visual integration model (McConkie & Rayner, 1976), information from one fixation enters a visually integrative buffer, much like an iconic store. Then, after a saccade, visual information from the next fixation also enters this buffer. Rather than displacing the current contents, the second fixation is integrated with the previous fixation, much as two overhead transparencies can be combined when overlaid. To accomplish this integration, each fixation would first be adjusted to match an internal model of the locations of objects in the world. All subsequent fixations could then be compared to this internal model. This precise and complete model would thereby allow visual stability because each fixation could be mapped to a single representation (see also Trehub, 1991, 1994).

If this sort of visual integration underlies our experience of stability, observers should be able to integrate two patterns that appear successively in the same spatial position, even if the patterns are projected to different retinal positions on each fixation. A number of studies have used the missing-dot task (Di Lollo, 1980) described above to explore this possibility (Bridgeman & Mayer, 1983; Irwin, Yantis, & Jonides, 1983; Jonides, Irwin, & Yantis, 1982, 1983; Rayner & Pollatsek, 1983). As in the studies of visual integration of patterns within a fixation, observers are asked to integrate two 12-dot patterns, thereby determining the location of the one dot missing from the 5×5 array. Recall that when the two dot patterns are presented to the same retinal location within a single fixation, observers can integrate the patterns. To study integration across saccades, the two complementary 12-dot patterns are presented to different retinal locations across an eye movement. One pattern is presented parafoveally and observers move their eyes to fixate it. During the saccade, the original pattern is replaced by its complementary 12-dot pattern. Once the saccade is complete, observers try to identify the location of the missing dot.

Initial evidence using this task supported the claim that observers maintain a precise and complete representation and that they can integrate it with subsequent patterns following an eye movement (Jonides et al., 1982). However, these initial findings resulted from a display artifact in which the original pattern persisted on the monitor, allowing integration of patterns projected to the same retinal location. (Bridgeman & Mayer, 1983; Irwin et al., 1983; Jonides et al., 1983; Rayner & Pollatsek, 1983). Despite repeated attempts over the next 10 years, no subsequent studies have provided convincing evidence that observers can integrate such complex patterns across an eye movement (e.g., Irwin, 1991; Irwin, Brown, & Sun, 1988; Irwin, Zacks, & Brown, 1990). Even though the two patterns occupy the same location in space, observers cannot combine them to form a single, coherent representation.

However, these results do not support the conclusion that we lack any representation of the pre-saccade pattern. Even if observers do not have an internal representation that is sufficiently precise and complete to perform the dot-integration task, they might still preserve some information from one view to the next. In fact, one recent study provides some support for a more limited form of visual integration (Hayhoe, Lachter, & Feldman, 1991). In this study, observers attempted to perceive a shape that was visible only if they were able to combine information across several fixations. Three consecutive views each presented a single dot that defined a corner of a triangle, and observers were asked to judge whether the top angle formed by integrating the patterns was acute or obtuse. Observers succeeded in judging the angle, suggesting that they retained enough information from one fixation to the next to determine the spatial relationships among the dots. This study provides evidence that some information can be preserved across views. However, it provides relatively little support for the idea that integration of complete and precise representations of entire scenes underlies our experience of stability in the real world—the task required only that a minimal amount of information be preserved.

Taken together, these studies suggest that some information might be preserved across saccades, but that information might not be precise or complete enough to support the visual integration model of visual stability. The typical inference from these findings is that we do not maintain a precise and complete visual representation across an eye movement. Note, however, that this inference is not logically supported by these experiments. Although successful visual integration requires a precise and complete representation, the failure to integrate two patterns does not imply the absence of a representation. All of these studies of visual integration depend on explicit reports of the features of the combined percept. Observers might well have a precise and complete representation but simply fail to perform the comparison (Angelone, Levin, & Simons, 2001; Levin, Simons, Angelone, & Chabris, in press; Simons, 2000). Or, more importantly, the results of the comparison might simply be unavailable to awareness. In other words, evidence from these studies could only confirm the existence of preserved information if that information were accessible to conscious report. Such explicit reports of preserved information could underestimate the amount of information that actually is preserved and integrated from a glance at a scene, thus more sensitive measures are needed to adequately test for the presence of preserved information.

Implicit Measures of Visual Integration: Priming from a Prior View

Priming is one potentially more sensitive approach to exploring the nature of the information preserved across views. Priming occurs when the prior presentation of one stimulus leads to more rapid or accurate processing of a subsequent stimulus. Note that priming measures do not require visual merging of the details of the two stimuli. Rather, facilitation provides evidence that some information extracted on one fixation influences the perception of an object on the next fixation.

In a typical priming study, observers respond only to the second stimulus, and the effects of the first stimulus result in changes in the response latency or error rate. Observers initially fixate the center of a screen and a preview picture is presented away from fixation. They then initiate an eye movement from fixation to this preview, and during the saccade, the original picture is replaced with a target picture that observers are asked to name (see Rayner, McConkie, & Ehrlich, 1978; Rayner, McConkie, & Zola, 1980).

In such priming studies, naming of the target picture is faster if the preview is identical to the target than if it is just a location marker such as an asterisk (Henderson, 1992; Henderson, Pollatsek, & Rayner, 1987, 1989; Pollatsek, Rayner, & Collins, 1984; Pollatsek, Rayner, & Henderson, 1990). This finding alone demonstrates that some information about the preview object is represented across views. Furthermore, the preview benefit does not seem to require explicit awareness of the relationship between the preview and the target. Through systematic variation of the similarity of the preview to the target we can determine how complete and precise a representation is preserved across the saccade. If the preview and target must be identical in order to produce facilitation, then if priming occurs, the representation of the preview must be complete and precise. Such a representation is more likely to be based on something akin to a visual image than on an abstract, nonvisual encoding of a few features of the object; if all of the details must be preserved with absolute precision in order to produce priming, then it seems implausible that an observer could rapidly abstract (e.g., verbally label) all of those details into a nonvisual representation. In contrast, if the preview produces priming when the preview and target are visually distinct exemplars of a category (e.g., a collie and a doberman, when observers are required to name the object at the basic level, dog), then the representation need not incorporate visual details of the preview. Instead, an abstract conceptual or verbal code would suffice.

The preview benefit does not appear to require a precise visual representation. Although naming latency is 100–130 ms faster when the preview is identical to the target than when it is just an asterisk, naming is equally fast when the preview picture is 10% larger or smaller than the target (Pollatsek et al., 1984). Thus, at a minimum, the preserved representation does not maintain precise scale information. Furthermore, naming is speeded by approximately 90 ms when the preview and target are different examples of the same category (e.g., two different pictures of horses; see also Pollatsek et al., 1990). Similarly, previews with the same name but different appearances (i.e., homonym pairs such as "bat") speed naming by approximately 50 ms. Apparently, relatively abstract, nonvisual representations (e.g., verbal, phonological, and semantic information) underlie much if not all of the preview benefit, suggesting that representations need not be precise.

Summary

Although some information clearly is preserved across saccades and used to facilitate processing on subsequent fixations, evidence from visual integration studies, both explicit pattern merging and implicit naming facilitation, provide little or no support for the preservation of complete and precise visual information across saccades. The convergence of the results from these tasks provides some support for the notion that we lack detailed internal representations altogether. However, this conclusion depends on confirming a null result, so the logical inference that we lack such representations is unsupported. Successful integration and facilitation require a representation, but failed integration and facilitation do not eliminate the possibility that observers do represent all of the information from a scene with perfect precision. The preview effects suggest that some information is preserved, even when observers are unaware of the nature of the preview or its relationship to the target. Thus, the more sensitive measures did reveal a greater degree of integration than did the explicit measures. This finding raises the possibility that even more sensitive measures might reveal the existence of more complete and precise representations. One such measure might be the ability to detect changes from one view to the next.

Using Change Detection to Explore Preserved Representations

In a visual integration task, observers must retain all of the details from one view and merge them with a subsequent view. If the representation is incomplete, they will be unable to perform the task. In contrast, detecting a change from one view to the next requires only that observers retain and compare the information that changed, even if that information was only a small subset of the potentially changing information. Successful change detection requires a representation of the pre-change object and a comparison to the post-change object, but observers need not maintain a complete and precise representation of the entire scene. Thus, change detection provides a less stringent and possibly more sensitive measure of whether any precise information is retained from one view to the next.

In other respects, most change detection studies rely on methods similar to those used in studies of visual integration. Observers view the original and changed display and then report whether or not they consciously detected a difference between the displays. Consistent with the failure to integrate displays and with the notion that observers do not maintain a complete and precise representation of each fixation, observers often fail to detect large changes to scenes from one view to the next.

Inferring the Precision of Representations

As in studies of integration and preview effects, most early studies of change detection focused on the ability to compare information from sequential fixations by making a change during a saccade (saccade-contingent changes). In one of the more dramatic early examples, observers read lines of text that AlTeRnAtEd CaSe with each letter (McConkie & Zola, 1979). Periodically, during a saccade, every letter changed case. Observers did not notice changes that coincided with an eye movement, even though the changes were impossible to ignore if they occurred during a fixation. Similarly, observers often miss large, saccade-contingent changes made to photographs of natural scenes (Grimes, 1996; Henderson & Hollingworth, 1999; McConkie & Currie, 1996).

More recent studies have extended these findings of "change blindness" to forms of visual disruption other than saccades (for recent overviews, see Rensink, 2000a; Simons, 2000; Simons & Levin, 1997). For example, change blindness occurs when a change occurs during a blink (O'Regan, Deubel, Clark, & Rensink, 2000), a flashed blank screen (Pashler, 1988; Rensink, O'Regan, & Clark, 1997; Simons, 1996), and during a cut or pan in a motion picture (Hochberg, 1986; Levin & Simons, 1997; Simons, 1996). Change blindness occurs for simple, artificial displays (e.g., Pashler, 1988; Rensink, 2000b; Simons, 1996) as well as for photographs of natural scenes (e.g., Rensink et al., 1997), and it can even occur during a real-world interaction (Simons & Levin, 1998).

Interestingly, change blindness can sometimes occur even in the absence of visual masking of the change location by a disruption. For example, change blindness occurs when distracting elements are briefly flashed over the display without covering the change location (O'Regan, Rensink, & Clark, 1999; Rensink, O'Regan, & Clark, 2000). The additional elements apparently draw attention away from the signal caused by the change. Furthermore, change blindness can occur in the absence of any visual disruption whatsoever. When an original photograph of a scene gradual fades into a modified one, observers often fail to notice the change (Simons, Franconeri, & Reimer, 2000); the change signal is visible when attention is focused on the change location, but the change occurs too slowly to attract attention.

The recent introduction of the flicker paradigm (Rensink et al., 1997) has allowed a more systematic exploration of the completeness and precision of the information retained from one view to the next. In the flicker paradigm, an original and modified display alternate repeatedly, separated by a blank screen (thereby giving the display its flickering appearance). The blank screen serves as a visual disruption, and successful change detection often requires many cycles of alternation. As in other change detection studies, in the absence of a transient signal, change detection requires observers to represent and compare the original and modified feature. Measuring the latency of change detection also allows an assessment of the completeness of the representation. If observers could represent all of the information in a display and compare it all to the modified display, change detection should happen rapidly, typically within a cycle or two. However, if representations are less complete, observers will need more cycles to detect the change because they will be able to compare less information on each cycle.

One recent series of studies explored the completeness of representations by varying the number of items in the display (Rensink, 2000b). If the number of items in the display exceeds the holding capacity of visual short-term memory, observers should need additional cycles to detect the change. This methodology allows an estimate of the number of items observers can retain and compare across views. These studies found that observers could retain and compare approximately four to six items with each exposure to the display (Rensink, 2000b), suggesting that the representation of each view is limited to the amount of information that can be encoded with focused attention (note that this number was somewhat larger for some change types, perhaps owing to grouping strategies). These studies further suggest that representations are relatively incomplete, limited to a small subset of the potentially changing items in the scene. However, representations of these items might be relatively precise. In a change detection task using simple shapes (Luck & Vogel, 1997), subjects were unable to remember more than four objects across a disruption. However, they could remember many features of each object (e.g. size, color), suggesting that their representations of those objects were reasonably precise.

Given that successful change detection requires some preserved information from the initial view, failed change detection has been taken to suggest the absence of such representations. However, as for explicit measures of visual integration, this inference rests on faulty logic (see Simons, 2000). Change blindness does not imply the absence of a representation;

rather, it shows only that observers do not have *explicit* access to the relevant contents of the previous view. Furthermore, change blindness could also result from a failed comparison process (Angelone et al.; 2001; Simons, 2000). As for studies of integration, one way to explore the possibility that explicit reports underestimate the completeness and precision of our representations is to adopt more sensitive measures of change detection. Given that successful detection requires a representation, if more sensitive measures reveal detection, then they also reveal the existence of a representation. Recently, a number of studies have explored whether more information is preserved than findings of change blindness initially suggested.

Changed Detection Without Awareness

Most forms of visual disruption serve to mask or eliminate the change signal that would otherwise direct attention to the change automatically (Simons & Levin, 1997). In the absence of such a signal, how is attention eventually focused on the change location? The two major accounts for successful change detection each lead to different implications for how the visual system represents and processes changes. First, change detection might be based entirely on explicitly available representations and comparisons. In the flicker task, observers shift attention from one item to the next until they happen to encode and compare the feature that is changing (Mitroff & Simons, in press). According to this hypothesis, explicit measures of change detection are appropriate estimates of the amount of information retained and compared across views, and changes will not be detected without awareness. That is, implicit measures of change detection will provide no additional evidence for the existence of complete and precise representations. Alternatively, change detection might occur without awareness, thereby revealing the existence of more complete and precise representations (Rensink, 2000a).

Evidence for implicit change detection falls into three categories: *identification*, *localization*, and *registration*. Implicit identification requires that observers represent information about the identity of the changed item even if they lack awareness of the change itself (Thornton & Fernandez-Duque, 2000). Such a representation requires some precision in the representation of the features of the original and/or changed object. Implicit localization requires that observers be drawn to or gain access to the change location even if they do not consciously detect the change (Fernandez-Duque & Thornton, 2000; Smilek, Eastwood, & Merikle, 2000). However, such representations need not include much precision about the object's properties or features. Implicit registration is the weakest claim: the change need only influence performance in some way in the absence of awareness of the change. Evidence for registration requires minimal completeness and precision and does not specify any functional role for implicit detection in the eventual explicit detection of the change

(Hayhoe, Bensinger, & Ballard, 1998; Rensink, 1998; Williams & Simons, 2000). Here we briefly discuss the existing evidence for each of these claims and then argue that, in fact, none of the evidence supports the claim that changes are implicitly detected (see Mitroff & Simons, in press; Mitroff, Simons, & Franconeri, 2001, for details).

IMPLICIT IDENTIFICATION

Only one published study has claimed support for implicit identification in the absence of awareness (Thornton & Fernandez-Duque, 2000). In these studies, observers were faster to judge the orientation of a cued item when a changed item in the display had the same orientation as the cued item, even if observers reported being unaware of the change. Although this finding suggests that observers processed the identity of the changed item without awareness, the study did not control for the spatial relationship between the cued and the changed items. On the critical trials in this experiment, the changed item was always diametrically opposite the target item. Thus, the relative positions of the changed item and the target item were predictable, and observers could become aware of the spatial contingency. When we replicated this finding but eliminated the consistent and predictable spatial relationship between the changed item and the cued item, the effect of the changed item on the orientation judgment was eliminated (Mitroff et al., 2001). That is, when the predictable spatial relationship between the cued target and the changed item was random, the presence of an unreported change no longer influenced judgments about the target. If the change were implicitly detected in the original experiment, then the predictability of the spatial relationship should have been irrelevant. Furthermore, given that observers might have been aware of this spatial relationship, it seems reasonable to suggest that the performance benefit from undetected changes can be attributed to an explicit strategy of attending to the diametrically opposite item rather than to implicit change detection. It is possible, of course, that future studies will reveal the existence of implicit identification, but our evidence suggests that such representations might not surface with more sensitive measures.

IMPLICIT LOCALIZATION

As for implicit identification, claims of implicit localization can also be explained by taking explicit strategies into account. The primary study supporting the existence of implicit localization found better-than-expected guessing of the change location when observers did not detect the change; given a two-alternative choice, observers showed better than 50% selection of the changed item when they were unaware of the change (Fernandez-Duque & Thornton, 2000). Yet, the study did not control for the possibility that observers could improve their performance by explicitly adopting a strategy of eliminating items known not to have changed. If observers determined that one of the two alternative items had not changed, then they would "guess" the correct item. This explicit strategy of eliminating one item predicts the levels of guessing found in the original study. Furthermore, the number of items observers are able to eliminate accurately predicts guessing performance across repeated cycles of the change (Mitroff et al., 2001). Other recent findings suggesting implicit localization (Smilek et al., 2000) can also be explained entirely by explicit mechanisms, with no need for implicit detection (Mitroff et al., 2001). Consequently, studies of implicit localization provide no evidence that scene representations are any more complete or precise than would be expected based on explicit measures of detection.

IMPLICIT REGISTRATION

As for implicit identification and localization, most findings supporting claims of implicit registration can be attributed to explicit mechanisms. For example, no-change responses are sometimes slowed by the presence of an unnoticed change (Williams & Simons, 2000). Yet, such effects could be attributed to variations in the confidence of these responses. When observers are more confident they tend to respond more rapidly, and observers might have generally been more confident when they correctly responded no-change. In fact, after taking into account the contributions of confidence, the presence or absence of a change accounted for only 2.5% of the remaining response time variance (Mitroff et al., 2001). Although confidence might just be another measure of implicit detection, it could also reflect explicit detection. Consequently, the finding that response-time differences can be attributed to confidence raises the possibility that such effects do not require implicit detection.

A similar line of research has found that fixation duration is affected by the presence of an undetected change (e.g., Hollingworth, Williams, & Henderson, in press): observers fixate longer on a changed object, relative to trials with no change, even when they did not detect the change. This finding could be consistent with implicit registration of a change, but it does not provide strong evidence for detection without awareness. Fixation duration in this case might simply be a more sensitive measure of explicit change detection in that observers might have been somewhat aware of the change. Likewise, the finding could result from confidence differences or from other explicit strategies. The concern that a purported implicit measure is just a more sensitive measure of explicit detection applies to other evidence of detection as well. A convincing demonstration of implicit detection would need to dissociate performance from sensitive measures of explicit detection (Reingold & Merikle, 1988), ideally showing that the result could not be attributed to awareness of the change. Given this ambiguity, both lines of research are consistent with the claim that implicit registration does not exist, and performance is no different than would be expected based on explicit mechanisms alone.

Summary

Apparently, change detection may not occur in the absence of awareness, and explicit measures of change detection might not necessarily underestimate the completeness and precision of our representations. Our claim that change detection, at least in the absence of a change signal, depends only on an explicit comparison process (Mitroff & Simons, in press; Mitroff et al., 2001) is consistent with the idea that observers are generally unaware of much of their environment. If observers were able to access and compare all of the items in a scene from one view to the next, then they would readily detect changes when looking for them. However, the failure to detect such changes does not imply the absence of a representation (Simons, 2000). Change blindness could well result from a failure to compare an existing, complete, and precise representation of one view to the subsequent view. In fact, more sensitive measures might well reveal the presence of such representations even in the face of change blindness. For example, observers typically failed to report the removal of a striped basketball, even when they were asked repeatedly whether anything had changed. However, when asked specifically if a ball had been present before, they discovered that they did have a representation of the ball and even reported some details correctly (e.g., see Simons, Chabris, Schnur, & Levin, in press). They failed to detect the change, even though they represented some details from the pre-change scene. If such detailed representations exist, then change blindness does provide evidence that detailed and complete representations are not always compared from one view to the next. Combined with evidence against implicit change detection, these studies suggest the existence of a limited capacity comparison process (see also Scott-Brown, Baker, & Orbach, 2000; Scott-Brown & Orbach, in press).

Much of visual perception occurs without awareness, and although we may not need to store the results of such percepts internally, there are almost certainly some cases in which performance would be enhanced by representing, however transiently, some of the contents of a scene. In the concluding section we consider what aspects of scene perception tend to occur without awareness and whether these effects imply complete or precise representations.

Representations, Actions, and the Experience of Stability

Findings from studies of visual integration and change detection provide relatively little evidence that we achieve an experience of visual stability by forming and combining complete and precise representations. However, the tasks used to study integration and change detection measure aspects of perception that might not require precise representations. Most judgments about objects do not require rapid and accurate detection of changes.

One area that does appear to require relatively precise information is action. In order to act on our world, we must represent the shapes and positions of objects with sufficient precision to avoid colliding with them or to successfully grasp them. Such actions certainly require precise percepts, and they might well require precise representations and comparisons from one view to the next. Consequently, visually guided action might be an appropriate domain to explore the possibility that representations are both complete and precise and that they can be integrated across views.

More importantly, visually guided action can be measured without explicit verbal reports and often is more precise than such reports. For example, observers better estimate the slope of a hill when they respond by manually adjusting a board than when they provide a verbal estimate (Proffitt, Bhalla, Gossweiler, & Midgett, 1995). People also adjust their actions in response to visual feedback without explicit awareness of the corrections. For example, expert table tennis players make small paddle adjustments during their stroke that correct the trajectory of the swing (Bootsma & van Wieringen, 1990), but they are not ordinarily aware of these fine adjustments. Similarly, initial saccades to a target are often inaccurate, but they are quickly followed by corrective saccades that bring the projection of the target into the fovea (e.g., Deubel, Wolf, & Hauske, 1982; Shebilske, 1976). Again, observers typically are unaware of both the error and the correction (Becker & Fuchs, 1969; Deubel et al., 1982). Such implicit corrections occur even when the target of a saccade is slightly displaced just after the beginning of the eye movement (Bridgeman, Hendry, & Stark, 1975). In fact, observers are often unaware of their eye movements altogether (e.g., Bridgeman, 1992).

These implicit corrections are useful in that they allow the eye movement targeting system to operate more efficiently. However, it is not clear that they require a complete or precise representation; they could be driven solely from the perceptual information currently at hand, and not from the representation of information across the saccade. In a test of this possibility, observers made an eye movement to a target and then tried to report the direction that the target shifted during the eye movement (Deubel, Schneider, & Bridgeman, 1996). If they could successfully report the direction, then they must have retained information about the first saccade target position. They could only judge the direction of the object's movement if they had stored the initial saccade target position and then compared that representation to the new position. Consistent with previous work, when the target object was present at the conclusion of the saccade, observers failed to notice the displacement. Critically, in this experiment, the reappearance of the target object was sometimes delayed until after the end of the first saccade. This temporal "gap" in which no object was visible helped observers to report the direction of the second displacement. Thus, observers can become aware of their corrective saccades even for displacements that typically are not detected. This finding suggests that the initial position of the target is represented with relatively good precision even after a second saccade and even if observers are unaware of having made the corrective saccade. Thus, some information about the positions of objects in the display is preserved from one fixation to the next, and this position information has sufficient precision to drive subsequent actions.

Given the importance of spatial information in constraining action, it seems reasonable that the visual system might store relatively precise and complete representations of the spatial relations between objects and the observer. Such representations need not be long-lasting, perhaps surviving only for a few fixations, because most actions do not require comparisons over long intervals. These representations might reduce the complexity of navigating and aid in our understanding of the visual world. Although the world may largely serve as its own memory (O'Regan, 1992) with relatively little need to internalize visual details, in order to act in the world we might need to retain some details about spatial layout.

Even if some information were preserved, explicit awareness of every fixation correction would make some simple tasks unbearably complex (e.g., finding a friend at the airport). If our visual system can accomplish the task of preserving just enough information to allow appropriate action, then we would have little need for complete representations of every aspect of our visual world (Hayhoe, 2000). Most of the representations we would need from one instant to the next involve just those properties of the world that can guide actions (such as corrective eye movements and motor adjustments). For example, our motor system can automatically update our own position relative to objects in the world, taking into account the relevant changes (Simons & Wang, 1998). This updating mechanism contributes to the experience of a stable world by integrating just enough information to allow for immediate actions. Furthermore, we rarely need to be aware of the representations underlying our actions. Anecdotally, too much awareness might actually impair performance. For example, thinking too much about a tennis swing can cause a volley to go into the net.

In contrast to positional information needed for motor control, we do not need to maintain a complete and precise internal representation of the features of objects. For the most part, objects and their features do not change from one moment to the next, and we rarely need to integrate textures across a saccade. Provided that the target of an eye movement is roughly in the right place following a saccade and that we can identify it with a minimally precise and somewhat abstract representation, our visual system can simply assume that everything else in the world remains unchanged. This assumption of stability provides an experience of a continuous, unchanging world, thereby explaining change blindness and the failure to integrate patterns across eye movements. Only when the change is particularly significant (i.e., a large displacement or a change to the meaning of the scene) will it affect our immediate actions. Under those conditions, it is more likely to draw attention and be noticed. Thus, important events might well be represented, but for the most part, our representations of the world could be relatively incomplete and only marginally precise without disrupting our experience of a stable visual world.

Conclusion

Much of perception does not require that information be preserved from one view to the next. Our review of the visual-integration and change-detection literatures suggests that precise and complete visual representations may be unnecessary for our experience of a stable, continuous visual world. Instead, our experience of stability is driven by precise representations of the information needed to guide action, accompanied by an assumption that the properties of objects in the world are unlikely to change across views. Of course, more sensitive measures might reveal the existence of complete, precise representations of all aspects of the visual world, but such detailed representations are not needed to explain our experience of an unchanging world from one view to the next.

Notes

Daniel Simons was supported by NSF Grant #BCS-9905578 and by an Alfred P. Sloan research fellowship. Stephen Mitroff was supported by an NSF graduate fellowship. Steven Franconeri was partially supported by a Department of Defense Graduate Fellowship. Please address all correspondence to Daniel Simons (*dsimons@wjh.harvard.edu*).

References

Angelone, B. L., Levin, D. T., & Simons, D. J. (2001). *Representation and comparison failures in change blindness.* Manuscript submitted for publication.

Becker, W., & Fuchs, A. F. (1969). Further properties of the human saccadic system: Eye movements and corrective saccades with and without visual fixation points. *Vision Research, 2,* 1247–1258.

Bootsma, R. J., & van Wieringen, P. C. W. (1990). Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 21–29.

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach.* New York: Oxford University Press.

Bridgeman, B. (1992). Conscious vs. unconscious processes: The case of vision. *Theory and Psychology*, *2*(1), 73–88.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research, 15*(6), 719–722.

Bridgeman, B., & Mayer, M. (1983). Failure to integrate visual information from successive fixations. *Bulletin of the Psychonomic Society, 21*(4), 285–286.

Campbell, F. W., & Wurtz, R. H. (1978). Saccadic omission: Why we do not see a grey-out during a saccadic eye movement. *Vision Research, 18,* 1297–1303.

Deubel, H., Schneider, W. X., & Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research, 36*(7), 985–996.

Deubel, H., Wolf, W., & Hauske, G. (1982). Corrective saccades: Effect of shifting the saccade goal. *Vision Research, 22*(3), 353–364.

Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General, 109*(1), 75–97.

Fernandez-Duque, D., & Thornton, I. M. (2000). Change detection without awareness: Do explicit reports underestimate the representation of change in the visual system? *Visual Cognition, 7*(1/2/3), 323–344.

Franconeri, S. L., & Simons, D. J. (2000). The role of abstract representations and motion signals in change detection. *Investigative Ophthalmology & Visual Science, 41*(4), S420.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Perception (Vancouver Studies in Cognitive Science)* (vol. 2, pp. 89–110). New York: Oxford University Press.

Hayhoe, M. (2000). Vision using routines: A functional account of vision. *Visual Cognition, 7,* 43–64.

Hayhoe, M. M., Bensinger, D. G., & Ballard, D. (1998). Task constraints in visual working memory. *Vision Research, 38*(1), 125–137.

Hayhoe, M., Lachter, J., & Feldman, J. (1991). Integration of form across saccadic eye movements. *Perception, 20,* 393–402.

Helmholtz, H. V. (1866/1924). *Treatise on physiological optics* (J. P. C. Southall, Trans.). (Vol. 3). Rochester, NY: The Optical Society of America.

Henderson, J. M. (1992). Object identification in context: The visual processing of natural scenes, *46*(3), 319–341.

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science, 10*(5), 438–443.

Henderson, J. M., Pollatsek, A., & Rayner, K. (1987). Effects of foveal priming and extrafoveal preview on object identification, *Journal of Experimental Psychology: Human Perception and Performance, 13*(3), 449–463.

Henderson, J. M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception and Psychophysics, 45*(3), 196–208.

Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1: Sensory Processes and Perception, pp. 22.21–22.64). New York: Wiley.

Hollingworth, A., Williams, C. C., & Henderson, J. M. (in press). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review.*

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*(6693), 575–577.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 23,* 420–456.

Irwin, D. E., Brown, J. S., & Sun, J.-S. (1988). Visual masking and visual integration across saccadic eye movements. *Journal of Experimental Psychology: General, 117*(3), 276–287.

Irwin, D. E., Yantis, S., & Jonides, J. (1983). Evidence against visual integration across saccadic eye movements. *Perception and Psychophysics, 34*(1), 49–57.

Irwin, D. E., Zacks, J. L., & Brown, J. S. (1990). Visual memory and the perception of a stable visual environment. *Perception and Psychophysics, 47*(1), 35–46.

Jonides, J., Irwin, D. E., & Yantis, S. (1982). Integrating visual information from successive fixations. *Science, 215,* 192–194.

Jonides, J., Irwin, D. E., & Yantis, S. (1983). Failure to integrate information from successive fixations. *Science, 222*(4620), 188.

Kahneman, D. (1968). Method, findings, and theory in studies of visual masking. *Psychological Bulletin, 70,* 404–425.

Levin, D. T., Simons, D. J., Angelone, B. L., & Chabris, C. F. (in press). Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology.*

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review, 4*(4), 501–506.

Lindberg, D. C. (1976). *Theories of vision from Al-Kindi to Kepler.* Chicago: University of Chicago.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281.

McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 563–581.

McConkie, G. W., & Rayner, K. (1976). Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed., pp. 137–162). Newark, DE: International Reading Association.

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception and Psychophysics, 25*(3), 221–224.

Mitroff, S. R., & Simons, D. J. (in press). Changes are not localized before they are explicitly detected. *Visual Cognition.*

Mitroff, S. R., Simons, D. J., & Franconeri, S. L. (2001). *The siren song of implicit change detection.* Manuscript in revision.

O'Regan, J. K. (1992). Solving the 'real' mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46*(3), 461–488.

O'Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition, 7,* 191–212.

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of "mudsplashes." *Nature, 398*(6722), 34.

Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics, 44*(4), 369–378.

Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General, 113*(3), 426–442.

Pollatsek, A., Rayner, K., & Henderson, J. M. (1990). Role of spatial location in integration of pictorial information across saccades. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 199–210.

Proffitt, D. R., Bhalla, M., Gossweiler, R., & Midgett, J. (1995). Perceiving geographical slant. *Psychonomic Bulletin and Review, 2*(4), 409–428.

Rayner, K., McConkie, G. W., & Ehrlich, S. F. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 529–544.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12,* 206–226.

Rayner, K., & Pollatsek, A. (1983). Is visual information integrated across saccades? *Perception and Psychophysics, 34*(1), 39–48.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics, 44,* 563–575.

Rensink, R. A. (1998). Mindsight: Visual sensing without seeing. *Investigative Opthalmology and Visual Science, 39,* 631.

Rensink, R. A. (2000a). The dynamic representation of scenes. *Visual Cognition, 7,* 17–42.

Rensink, R. A. (2000b). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition, 7,* 345–376.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*(5), 368–373.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual Cognition, 7,* 127–146.

Scott-Brown, K. C., Baker, M. R., & Orbach, H. S. (2000). Comparison blindness. *Visual Cognition, 7,* 253–267.

Scott-Brown, K. C., & Orbach, H. S. (in press). Contrast discrimination, non-uniform patterns and change blindness. *Proceedings of the Royal Society, B.*

Shebilske, W. L. (1976). Extraretinal information in corrective saccades and inflow vs. outflow theories of visual direction constancy. *Vision Research, 16*(6), 621–628.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science, 7*(5), 301–305.

Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition, 7,* 1–15.

Simons, D. J., Chabris, C. F., Schnur, T., & Levin, D. T. (in press). Preserved representations in change blindness. *Consciousness and Cognition.*

Simons, D. J., Franconeri, S. L., & Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception, 29,* 1143–1154.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*(7), 261–267.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people in a real-world interaction. *Psychonomic Bulletin and Review, 5*(4), 644–649.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science, 9*(4), 315–320.

Smilek, D., Eastwood, J. D., & Merikle, P. M. (2000). Does unattended information facilitate change detection? *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 480–487.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*(11), 1–29.

Thornton, I. M., & Fernandez-Duque, D. (2000). An implicit measure of undetected change. *Spatial Vision, 14*(1), 21–44.

Trehub, A. (1991). *The cognitive brain.* Cambridge, MA: MIT Press.

Trehub, A. (1994). What does calibration solve? *Behavioral and Brain Sciences, 17,* 279–280.

Van Hoorn, W. (1972). *As images unwind: Ancient and modern theories of visual perception.* Amsterdam: University Press Amsterdam.

Williams, P., & Simons, D. J. (2000). Detecting changes in novel, complex three-dimensional objects. *Visual Cognition, 7,* 297–322.