VISUAL COGNITION, 2000, *7* (1/2/3), 127–145

On the Failure to Detect Changes in Scenes Across Brief Interruptions

Ronald A. Rensink

Cambridge Basic Research, Nissan Research & Development Inc., Cambridge, MA, USA

J. Kevin O'Regan

Laboratoire de Psychologie Expérimentale, CNRS, Université René Descartes, Paris, France

James J. Clark

Department of Electrical Engineering, McGill University, Montreal, Quebec, Canada

When brief blank fields are placed between alternating displays of an original and a modified scene, a striking failure of perception is induced: The changes become extremely difficult to notice, even when they are large, presented repeatedly, and the observer expects them to occur (Rensink, O'Regan, & Clark, 1997). To determine the mechanisms behind this induced "change blindness", four experiments examine its dependence on initial preview and on the nature of the interruptions used. Results support the proposal that representations at the early stages of visual processing are inherently volatile, and that focused attention is needed to stabilize them sufficiently to support the perception of change.

Over the past few decades, evidence has been accumulating that—contrary to our subjective impressions—we do not have a coherent and detailed representation of the coherent and detailed world that surrounds us. For example,

---

Please address all correspondence to R.A. Rensink, Cambridge Basic Research, Nissan Research & Development, Inc., 4 Cambridge Center, Cambridge, MA 02142, USA. Demos of the basic effect described in this paper are available at the web site: http://www.cbr.com/projects

The authors would like to thank Jack Beusmans, Dan Simons, Ian Thornton, Carol Yin, and two anonymous reviewers for their comments on an earlier draft of this paper. Also, thanks to Vlada Aginsky, Monica Strauss, and Nayantara Santhi for their help in running the experiments. Portions of this paper were presented at the Annual Meeting of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, FL, May 11-16, 1997.

---

observers often find it difficult to detect changes in an image made during a saccade (e.g. Bridgeman, Hendry, & Stark, 1975; Grimes, 1996; McConkie & Zola, 1979). This indicates that the visual details of successive fixations cannot be added, compared, or otherwise combined. Indeed, all that can be carried across a saccade are a few properties of a few previously fixated items (Ballard, Hayhoe, & Whitehead, 1992; Henderson, 1997; Irwin, 1991, 1996). A similar *change blindness* can also be induced when changes are made during the interstimulus interval (ISI) between two displays that are each presented for a limited time (e.g. Pashler, 1988; Phillips, 1974; Simons, 1996). In these experiments, presentation times range from 250 msec to 2 sec; in all cases, observers are poor at detecting change whenever the displays are separated by an ISI of more than 70–100 msec. (For a more complete overview of these and related experiments, see Simons & Levin, 1997.)

The change blindness induced by saccade-contingent techniques might have been caused by disruptions due to eye movements; limited-display techniques might not have given the observer sufficient time to build an adequate representation. Both these concerns are eliminated in the *flicker paradigm* (Rensink et al., 1997). Here, an original image A repeatedly alternates with a modified image A', with brief blank fields placed between successive images. The observer freely views the flickering display until the change is seen (Fig. 1).

But in spite of the extended viewing that this paradigm provides, observers still experience great difficulties in detecting change, even when the changes occupy large parts of the image (up to 30 sq. degrees), are repeatedly made, and the observer knows that they will occur. This suggests that in the absence of focused attention the representations used at early levels of visual processing[1] are not coherent enough to support effortless perception of change (Rensink, this issue; Rensink et al., 1997). If attention is drawn by a local transient generated by the transformation in the image (see e.g. Klein, Kingstone, & Pontefract, 1992) the change will be seen; otherwise, it will be effectively invisible. Because attentional capacity is extremely limited, this view implies that no more than a few objects can be seen to change at any one time. The representation of the other objects is volatile, their descriptions simply being replaced by those of incoming stimuli at the same retinal location (Rensink, 1997, this issue). With this kind of representation, it would appear that no large-scale accumulation of visual detail ever takes place. As such, this rules out the

[1] As used here, "early vision" refers to the stage of visual processing concerned with the initial sorting out of the physical factors responsible for the pattern of illumination on the retina. The knowledge used for this is based on general physical considerations, and is not specific to any particular stimulus. Early representations are generally regarded as retinotopic, and formed rapidly (i.e. within a few hundred msec) and in parallel across the image (see e.g. Marr, 1982; Rensink, 1992; Rensink & Enns, 1998).

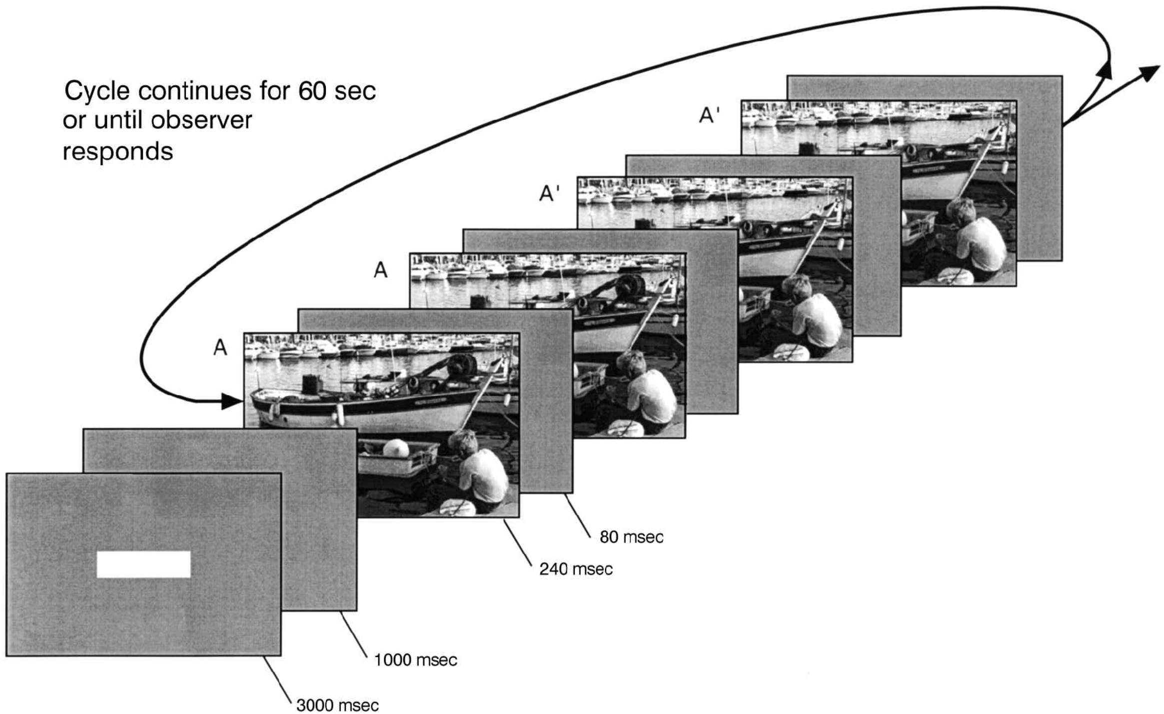

FIG. 1. Standard conditions of the flicker paradigm (i.e. those of Exp. 1 of Rensink et al., 1997). Trials begin with a 3-sec gray field containing a white rectangle to alert the observer that a trial is beginning. This is followed by a 1-sec gray field, followed by an display sequence that continued alternating until the observer responded or 60 sec had elapsed. In the example here, original image A (boat with winch in front) and modified image A' (boat with winch removed) are displayed in the order A, A, A', A', … with gray fields between successive images. For the standard conditions, display durations were 240 msec, blank durations 80 msec, and blank fields were medium gray.

existence of a buffer in which detailed information from successive eye movements is fused into a single, complete representation of the world around us.

The impression that we have a complete, coherent representation of our visible surroundings is a strong one, and belief in it is not easily set aside. To be sure, this impression does have some objective basis: Even under conditions that induce change blindness, the images themselves can be clearly perceived, with observers easily able to find changes in objects described by verbal cues (Rensink et al., 1997). This indicates that at any given moment the early-level representations do contain a considerable amount of detail. But however detailed they may be, these representations are not coherent enough to support the effortless perception of change. The issue examined here is how this lack of coherence comes about.

VOLATILITY VS DISRUPTION

In what follows, the term "coherence" will be used to denote the existence of both interconnection and consistency in a representational structure. The spatial coherence of two adjacent structures implies that they form part of the same object, extended over space; likewise, the temporal coherence of two successive structures implies that they form part of the same object, extended over time.

According to coherence theory (Rensink, 1997, this issue), early-level structures simply do not have much coherence in the absence of focused attention. Some integration of information over space is possible at early levels, allowing for processes such as rapid grouping and rapid completion of occluded objects (see e.g. Rensink & Enns, 1995, 1998). Stimuli at the same retinal location can also be integrated across time, provided they arrive within about 100 msec of each other (see e.g. DiLollo, 1980). But beyond these limits, early-level representations are assumed to be inherently *volatile*: They exist only as long as light continues to enter the eye, with old representations simply being replaced by any new ones formed at their retinal location. In this view, focused attention is required to provide the coherence that knits the early-level structures into larger-scale objects and allows these objects to retain an identity over time (see also Kahneman, Treisman, & Gibbs, 1992).

In contrast to this *volatility hypothesis*, it could be argued that early-level representations generally do attain a high degree of spatiotemporal coherence (or at least temporal coherence) in the absence of focused attention, but that the conditions causing change blindness somehow disrupt consolidation, or disrupt the processes that use the consolidated representation to detect change. According to this *disruption hypothesis*, change blindness is due to a failure that is a side-effect of the experimental manipulations rather than something inherent in the representations themselves. As such, change blindness may be a striking phenomenon, but one that provides relatively little understanding of the way that vision operates under normal conditions.

PREDICTIONS

To determine whether volatility or disruption provides the better account of change blindness, four experiments will examine how this phenomenon depends upon different kinds of viewing conditions:

1. Extended Preview. Both recognition and recall of pictures improve with longer viewing time, at least up to 5 sec (Tversky & Sherman, 1975). This suggests that a relatively long consolidation process might be needed to form a memory durable enough to support the perception of change.[2] Perhaps the interruptions caused by the blank fields interfered with this consolidation process. Or perhaps several seconds of viewing are required for an observer to build up a stable scene representation (possibly by integrating information across several eye movements), so that measuring response times relative to the initial presentation of a scene overestimates the difficulty of detecting change.

Experiment 1 examines these possibilities by providing an 8-sec preview before starting the main flicker sequence. If memory consolidation is disrupted by the blanks, or if several seconds of viewing are needed to build up a scene representation, performance should improve. But if change blindness is due to volatility, extended viewing will not affect the durability of the low-level representations, and so performance should remain the same.

Note that this prediction only requires extending the initial viewing time, and not the viewing time of subsequent stimuli. This is because once a coherent representation has been formed, it should enable changes to be noticed as soon as they occur—there should be no waiting for the consolidation of a new representation that describes the changed situation.

2. Different Blank Durations. Many visual mechanisms are tuned to a limited range of temporal parameters. For example, contrast sensitivity at high luminances has a peak at about 20 Hz, i.e. a time scale of 50 msec (see e.g. Woodhouse & Barlow, 1982). Thus, if disruption is achieved by some

---

[2]The advantage obtained by extra viewing time could also be due to a buildup of abstract information in long-term memory, something that may take place when we view a scene repeatedly. But although such a buildup may improve the detail of long-term representations, it does not necessarily improve the ability to detect change. The consolidation considered here, in contrast, is something that *would* improve the ability to detect change, presumably by facilitating the large-scale accumulation of visual detail in a memory of relatively short (i.e. several seconds) duration.

early-level mechanism, the degree of disruption—and therefore change blindness—may differ with different durations of the blank field. In particular, there may be one value at which disruptive effects are at a maximum (or minimum). But if volatility is the cause, performance should only be affected to the extent that the interruptions eliminate information from the local transients; consequently, it should be much the same over a large range of durations.

Experiment 2 tests blank durations of 40 msec, 160 msec, and 320 msec. The results here will be combined with those obtained for durations of 80 msec (Rensink et al., 1997, Exp. 1). This will form a range covering most of the time scales found in early visual processes.

It is important to keep in mind that the mechanisms believed to cause disruption are unknown, and so the disruption hypothesis cannot provide an unequivocal prediction of how performance is affected by the duration of the blanks. Similarly, although volatility predicts similar performance for durations yielding similar levels of attentional distraction, it is not entirely clear which durations these might be, since other factors may also be at play. For both hypotheses, then, a set of outcomes are possible. But fortunately these sets of outcomes are not the same. Thus, if performance is examined at several different blank durations, the pattern of results may be inconsistent with one of these hypotheses, or at least favour one over the other.

3. Different Blank Colours. Most visual mechanisms are sensitive not only to the duration of a stimulus, but also to several other of its properties. Experiment 3 tests the disruption and volatility hypotheses by examining the effect of the luminance and hue of the blank fields. Three colours are used: Black, white, and a red isoluminant with the gray used in Experiment 1. Results here will be combined with those previously obtained using gray blanks (Rensink et al., 1997, Exp. 1).

If change blindness is due to disruption, it should be greatest for the black and white fields, since these provide the largest swings in luminance, and thus send the largest amounts of energy to the early-level mechanisms;[3] meanwhile, hue should have no large effect. But if blindness is due to volatility, performance should be a function of information loss rather than simple energy, and so—assuming that no other mechanisms are involved—roughly the same degree of blindness should be found in all conditions.

---

[3]The average amount of the luminance swing depends in part on the difference between the luminance of the blank field and the average luminance of the images. For example, if most of the images are dark, black fields will cause a relatively small luminance swing; if most of the images are light, they will cause a large one. For the images used here, average luminance was close to that of the gray fields used in the standard conditions. Consequently, in these experiments, both the black and the white fields cause larger luminance swings on average than do the gray fields.

Note that as for the case of blank durations, unequivocal predictions are again impossible. The best that can be done is to test performance for various colours, and then see if the results are incompatible with the set of possible outcomes from either of the two hypotheses.

4. Reduced Coverage of the Interruption. Another possibility is that coherence does exist under conditions of normal viewing, but is so fragile that any kind of global interruption abolishes it. Although interruptions occur each time the eye is moved (3–4 times per second) and so ought to be handled by the visual system, it might be that interruptions are especially disruptive when the eye is stationary.[4] Alternatively, it might be that coherence always exists, but that the global coverage of the interruptions somehow paralyzes the mechanisms that use the coherent representation to detect change. (Note that in both these cases, the key factor is the global nature of the interruption—if local interruptions caused coherence to fail, the perception of change would be virtually impossible in the dynamic world of everyday life.)

To determine if coherence exists in the absence of global interruptions, Experiment 4 replaces the 80 msec blank field by a set of six relatively small (2.0° × 2.5°) but highly salient patches[5] briefly superimposed on the image for 80 msec. The onset of these patches creates brief local interruptions ("splashes" or "splats") that have a fraction of the coverage of the blank fields, but still are transients synchronized to the onset of the image changes. The locations of the splats were selected so that they did not cover the areas that were changing. Because the onset of new items in a visual field draws attention (Yantis, 1993), these splats could deflect attention while avoiding the disruption of structures elsewhere in the image.[6] (See Wolfe, 1994, for a somewhat similar way of deflecting attention.) If disruption were the cause of change blindness, performance should be as good as when no interruption is present. But if volatility were responsible, the distraction created by the splats should impair performance to some degree.

[4]In these experiments, observers are free to move their eyes. Although many of the changes may coincide with an eye movement, at least some of them are likely to take place when the eye is at rest.

[5]Patches were rectangles having a texture consisting of black and white checks, each check being approximately 20 arc minutes in size.

[6]This assumes that different parts of the image are processed independently. Such independence is a natural consequence of the structure of early vision, which is generally believed to be carried out by a retinotopic array of processors operating in parallel. If these processors are not largely independent, a single error or noisy signal could lead to a cascade that would destroy the structures at all other locations (see e.g. Rensink, 1992).

GENERAL METHOD

To allow direct comparison with the results of Rensink et al. (1997), exactly the same experimental procedures and stimuli were used: An original image A and modified image A' were displayed in the sequence A, A, A', A', … with gray blank fields appearing between successive images[7] (Fig. 1). In this "standard" condition, images were displayed for 240 msec and blanks for 80 msec, with blank fields being medium gray.

All experiments used the same set of 48 colour images of real-world scenes. Each image was 27° wide and 18° high. The images and the changes made to them were exactly those used by Rensink et al. (1997). In that study, the presence, colour, or location of a single object or region was changed in each image. All three types of change occurred equally often, with the different types roughly equated for areas and locations in the image. Changes were also divided according to the amount of interest in the items being changed. Interest was defined operationally via an independent experiment in which five observers were asked to view each image and provide a brief verbal description of it. *Central interests* (CIs) were objects or areas mentioned by most observers; *marginal interests* (MIs) were objects or areas mentioned by none. (For more details, see Rensink et al., 1997.)

A Macintosh computer presented the displays, controlled the experiments, and collected the data. Ten naïve observers participated in each experimental condition. Observers were informed of the types of changes possible. They were asked to press a key when they saw the change, and then to verbally describe it. In the standard conditions, each trial consisted of a 3-sec gray field containing a white rectangle, followed by a 1-sec delay, followed by a display sequence that continued until the observer responded or until 60 sec had passed (Fig. 1). A set of six practice trials familiarized observers with the protocol. The order of the images was randomized for each observer. As might be expected from the use of large changes, virtually all the changes were easily seen once noticed: Identification error rates were low, averaging less than 2.0% across all experiments.

Data analysis primarily consisted of two-tailed *t*-tests, with the 48 measurements generating six scores for each observer, each score being the average response time for each type of change. ANOVAs were used to determine any systematic variations caused by particular stimulus factors. To compensate for

---

[7] In these experiments, the ratio of image presentations per image alternation (the *Clark ratio* κ), is always 2. As shown in Experiment 2 of Rensink et al. (1997), there is no significant difference between this and a Clark ratio of 1 (i.e. where original image A and modified image A' are displayed in the sequence A, A', A, A' ...), provided that alternation rates are kept the same. To keep the experimental methods identical to those of Rensink et al. (1997), a κ of 2 is used in all the experiments here.

the different alternation rates, all response times were normalized by dividing through by the alternation time, resulting in measures that were always in terms of number of alternations.

EXPERIMENT 1: EXTENDED PREVIEW

Experiment 1 gave observers an 8-sec uninterrupted preview of the original image before the flicker sequence began. Observers were asked to remember as much of the image as possible in order to improve their ability to notice change. If an extended period of uninterrupted viewing can enable the construction or consolidation of a scene representation, change detection should be improved. But if early-level representations are inherently volatile and no large-scale accumulation of information occurs, preview should have no effect.

The results (Fig. 2) are clear—there is no effect of preview. For no type of change was there any significant speedup of detection (all Ps > .2), and no overall effect of extended preview was found, $F(1,10) = 0.6$, $P > .4$. Thus, change blindness is not caused by an insufficient time to construct a coherent representation, or by some flicker-induced interference with its consolidation in memory.

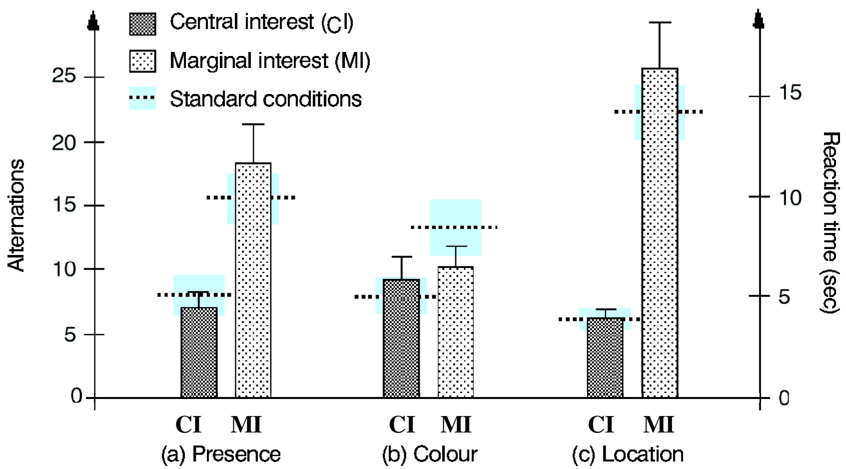

FIG. 2. Effect of extended preview (Experiment 1). Error bars indicate one standard error. Dashed lines indicate results under standard conditions; horizontal gray bars indicate their standard error. (These values taken from Rensink et al., 1997.) As is evident, no significant differences were found between the preview and standard conditions.

EXPERIMENT 2: DIFFERENT BLANK DURATIONS

Experiment 2 examined change detection using blank field durations of 40 msec, 160 msec, and 320 msec. Results here were combined with those obtained for the standard durations of 80 msec (Rensink et al., 1997, Exp. 1). If a disruption mechanism does exist, there may be one particular time scale at which its effects are especially pronounced. If one of the durations corresponds to this value, the degree of blindness will be at a maximum, and its strength for other durations will fall off accordingly; if the durations tested fall on one side of this value, performance will monotonically increase or decrease with the duration of the blanks. In contrast, the volatility hypothesis predicts that performance should be much the same when durations are large enough to create global transients that swamp the local transients generated by the transformations in the image.

The results (Fig. 3) show that detection of change was relatively poor for all conditions. No significant interactions were found between duration and interest, $F(3,30) = 2.8$, $p > .05$, duration and change type, $F(6,60) = 0.87$, $p > .5$, or all three together, $F(6,60) = 0.98$, $p > .4$. However, a significant effect of duration was detected, $F(3,30) = 3.66$, $p < .03$.

A comparison of response times against those of the standard conditions shows that detection was significantly faster for durations of 40 msec; CI: $p <$

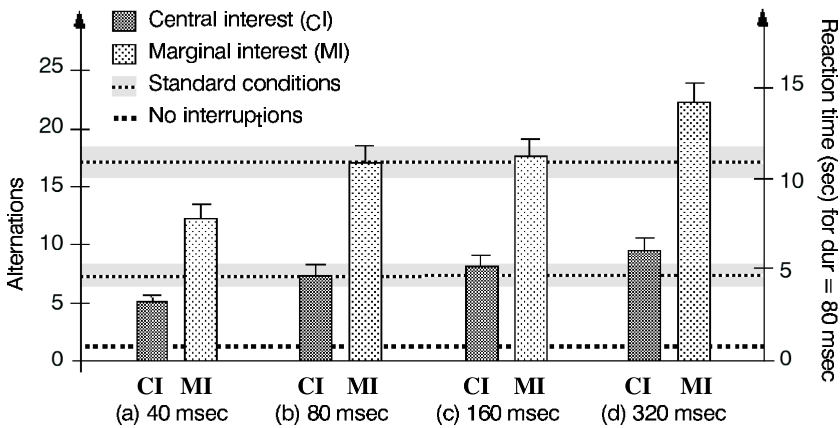

FIG. 3. Effect of blank field duration (Experiment 2). Error bars indicate one standard error; dotted line indicates baseline performance when no interruption is present (taken from Rensink et al., 1997). Dashed lines indicate results under standard conditions; horizontal gray bars their standard error. (a) 40 msec blank. For all types of change, a similar pattern was found: Changes were easier than in the standard condition, although not as easy as when no blanks were present. (b) 80 msec blank. This is the standard condition; data from Rensink et al. (1997). (c) 160 msec blank. When analysed in terms of number of alterations required to see the change, no significant differences were found between this and the standard condition. (d) 320 msec blank. Although not large in magnitude, slowdown was significant.

.05; MI: $p < .02$. This could be due to a decreased disruption by early-level processes. Alternatively, this could be due to an increased involvement of early-level motion detectors, which have temporal windows of about 100 msec or less.[8] As the duration of the blank decreases, temporal integration begins to take place between the original and modified images, so that the transients become weaker in areas of no change, and stronger in areas of change. At 40 msec, the strength of the local signals generated by the change in the image is presumably at the point where it begins to rise above the global signals generated by the blanks. But note that in this case change blindness still exists—although response times are faster, they are still significantly slower than when interruptions are absent.

In contrast to the speedup that occurred with shorter durations, longer durations of 160 msec (Fig. 3c) led to no significant differences in response times; CI: $p > .4$; MI: $p > .7$. Such indifference suggests that disruption may not be the cause of change blindness, since the degree of disruption is unlikely to be exactly the same for two different temporal scales. This indifference is, however, compatible with an attentional mechanism that examines a constant number of items at each alternation.

Blanks of 320 msec (Fig. 3d) gave rise to detection times somewhat slower than for 80 msec blanks, although this slowdown was significant only for marginal interests; CI: $p > .1$; MI: $p < .05$. Although not large, this slowdown is not predicted by a simple volatility hypothesis. Perhaps a degradation of sorts is occurring here, such as the decay of some kind of memory. A likely candidate in this regard would be iconic memory, which begins to decay within about 250 msec of visual offset (Sperling, 1960). Note that this slowdown is also compatible with the disruption hypothesis, assuming a monotonic decrease in performance with duration.

In summary, then, these results were unable to conclusively disprove either the volatility or the disruption hypothesis. Performance for the 40 msec

[8] The exact extent of the temporal window for early-level motion detection has proven rather difficult to determine. Although many estimates place it at around 100 msec or less (e.g. Bischof, Seiffert, & DiLollo, 1996; Braddick, 1973), under some conditions it appears to be as large as 1–2 seconds (van de Grind, Koenderink, & van Doorn, 1986). However, the effect of these more sluggish detectors does not seem to be large: Motion phenomena can usually be modelled quite well by detectors with temporal windows of less than 100 msec (e.g. Bischof & DiLollo, 1995).

It is also important to keep in mind that the important factor here is how well motion *attracts* attention. As such, the motion signal does not need to be eliminated—all that is required is that the signals from the changing region be no stronger than the signals from the rest of the image. This will happen when the temporal window cannot extend over the ISI to effectively encompass both the original and the changed image; in this case, two strong motion signals—corresponding to the onset and offset of the blank field—are generated across the image during each ISI. For the conditions used here, this seems to occur when ISIs are about 80 msec or greater, a limit similar to that found in the early "one-shot" experiments on change detection (e.g. Phillips, 1974).

durations was such that it could be made compatible with either hypothesis. The slowdown at 320 msec was compatible with both hypotheses as well. The similarity in performance for durations of 80 msec and 160 msec was more indicative of volatility. But this was only a single data point, a data point also compatible with the disruption hypothesis.

EXPERIMENT 3: DIFFERENT BLANK COLOURS

Experiment 3 examined three different colours of the blank field: Black, white, and red. The red field was isoluminant with the standard gray. If change blindness were due to disruption, it should be greatest for the black and white fields; if due to volatility, it should be the same in all conditions.

Figure 4 shows performance for the black and white fields; for the sake of comparison, the results for the standard gray (taken from Exp. 1 of Rensink et al., 1997) are also included. As can be readily seen, no reliable differences were found. No interactions were found between luminance and interest, $F(2,20) = 0.55$, $p > .5$, luminance and change type, $F(4,36) = 0.44$, $p > .7$, or between all three factors taken together, $F(4,36) = 0.047$, $p > .9$. Indeed, no overall effect of luminance was found, $F(2,20) = 1.03$, $p > .3$. Thus, the increased magnitude of the transients did not cause performance to deteriorate; if anything, there was a slight (although non-significant) tendency to detect changes more quickly.

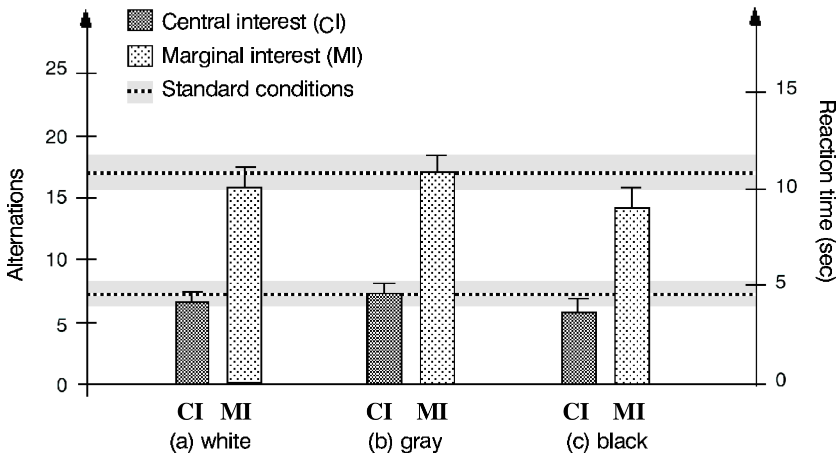

FIG. 4. Effect of blank field luminance (Experiments 3a–3b). Error bars indicate one standard error. Dashed lines indicate results under standard conditions; horizontal gray bars their standard error. (a) White fields. For all types of change, a similar pattern was found: no significant difference between this and the standard condition. (b) Gray fields. This is the standard condition; data from Rensink et al. (1997). (c) Black fields. Although changes here were detected slightly faster, this speedup was not significant.

This result points toward two rather different possibilities, corresponding to the extremes of the disruption and volatility hypotheses. The first is that the gray blanks may have already disrupted things as much as possible. This would then create a floor effect: The higher-energy changes due to the black and white fields were simply unable to cause any further deterioration in performance. Alternatively, the blank fields may have caused no disruption at all, since there may have been no coherence to begin with. Instead, the effect of the blanks may have been only to impede the automatic drawing of attention to the location of the change, with the degree of luminance change not affecting this in the least.

Results for the red fields are shown in Fig. 5. Performance here was not significantly affected for changes in presence or location, or for colour changes in central interests (all *p*s > .2). However, colour changes in marginal interests were significantly more difficult to detect ($p < .05$), with response times comparable to those for the other types of MI change. This dependence on colour is difficult to account for in terms of disruption, for it is not clear why a disruption mechanism should be so sensitive to hue. And the fact that performance *worsened* is even more problematic: In order to explain the indifference to luminance, the disruption hypothesis required the contrary conclusion that the standard conditions already caused maximal disruption.

The greater degree of blindness generated by the red fields can, however, be reconciled with the volatility hypothesis. This can be done by assuming that detection of colour change under standard conditions is *assisted* to some extent by early-level mechanisms. A good candidate here is a motion mechanism sensitive to hue (see e.g. Cropper & Derrington, 1996). If the achromatic transients

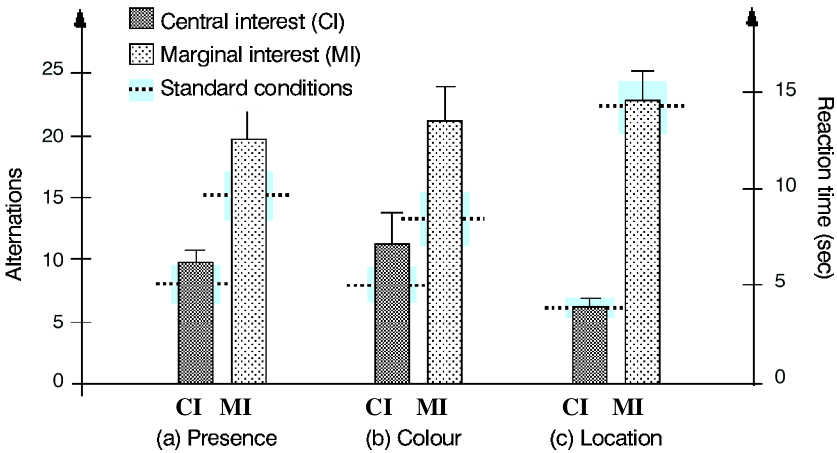

FIG. 5. Effect of blank field hue (Experiment 3c). Error bars indicate one standard error. Dashed lines indicate results under standard conditions; horizontal gray bars their standard error. Response patterns remained largely unaffected by field hue. However, a significant slowdown occurred for colour changes in marginal interests.

generated by the gray fields do not swamp the chromatic transients generated by the colour changes, remnant signals could help draw attention to the change, and so improve performance. But when the interruption itself contains a chromatic transient, the local chromatic signals are swamped, and detection becomes just as difficult as for the other kinds of change. Note that selective attention to the chromatic signal may also explain why observers do better when they know that the change will be one of colour (Aginsky & Tarr, this issue; Aginsky, Tarr, & Rensink, 1997).

EXPERIMENT 4: REDUCED COVERAGE OF THE INTERRUPTION

The previous experiments show that if early-level coherence exists under normal viewing conditions, then either it must be relatively fragile or else the mechanisms that use it to detect change must be easy to paralyse. To provide a test that disturbs early visual processing as little as possible, Experiment 4 used brief, localized achromatic "splats" that did not cover the areas of the image that were changed. According to the disruption hypothesis, performance should now be as good as when no interruption is present; according to the volatility hypothesis, the splats should continue to deflect attention, and so continue to impair performance to some degree.

Results (Fig. 6) show that responses were considerably faster than when entire blank fields were used. But more importantly, they were still slower than for the corresponding no-interruption conditions[9] (all *p*s $< .01$). Although there was no overall effect of change type, $F(2,18) = 2.70$, $p > .05$, there was an interaction between change type and interest, $F(2,18) = 9.83$, $p < .002$. This was due largely to presence, which alone showed no significant difference between CIs and MIs ($p > .6$). Pairwise comparison of MI change types showed no significant differences (all *p*s $> .4$), whereas pairwise comparisons of CI change types showed that presence was slower than the others ($p < .05$ for location). This suggests that the interaction was due to a slowdown in responding to changes in CI presence, something that can be understood by considering that these items were visible only half the time, and so could not be examined quite as readily as the items for the other kinds of CI change.

As in the case of global interruptions, there was a strong overall effect of interest, $F(1,9) = 154.06$, $p < 10^{-6}$, with detection slower for MIs (5.2 alternations) than for CIs (3.2 alternations). Given that it takes about 1.4 alternations to respond to a change in the no-interruption condition (Fig. 6), the additional viewing time needed for MIs is about 3.8 alternations. This corresponds to the time required to check about seven transient items, assuming one item per

[9]Data for the no-interruption condition were taken from Experiment 1 of Rensink et al., 1997.

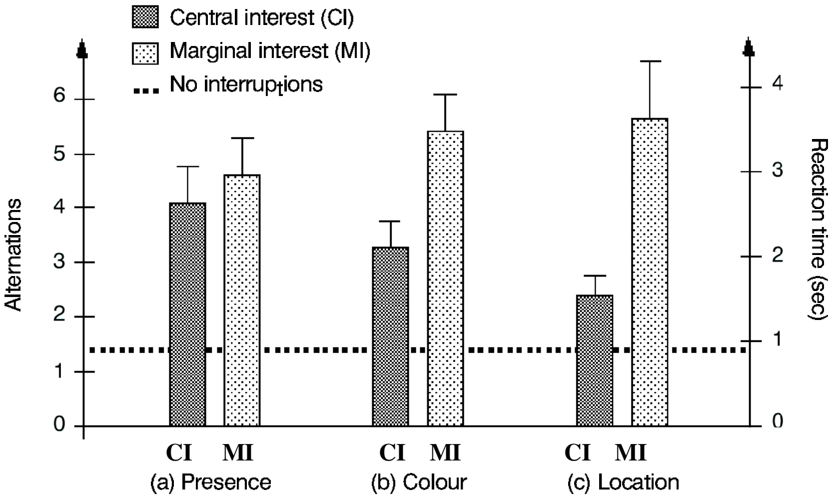

FIG. 6. Effect of splats (Experiment 4). Error bars indicate one standard error; dotted line indicates baseline performance when no interruption is present (taken from Rensink et al., 1997). Responses for all types of change were not significantly different from each other. However, all responses were significantly slower than for the no-interruption conditions, and the amount of this slowdown was roughly constant under these conditions.

display presentation.[10] This suggests a serial search of the image, with each splat checked in turn, and the MI being checked after all six splats have been examined. Such behaviour is consistent with the highly salient splats drawing attention more effectively than the MIs. The faster detection of CI changes would then be expected if the more interesting CIs can draw at least some attention before all the splats have been checked.

In any event, the main point here is that even though the image at the location of the change is not disrupted, the change still cannot be detected effortlessly. Some kind of limited process—presumably involving focused attention—is needed to see it.

DISCUSSION

Four sets of experiments investigated the following issue: Is change blindness due to an inherent volatility of early visual representations, or to the disruption of a coherent representation that would ordinarily support effortless detection

[10] The presentation of the splats was designed to be the same as the presentation of the blank fields. Since the Clark ratio κ is 2 (i.e. two presentations of an image for each alternation), there are also two presentations of the splats for each alternation. Each time a set of splats appears, one of the splats could easily be checked: Identifying the transient as a splat can be done by visual appearance alone, and so does not require waiting until an image alternation has occurred.

of change? Experiment 1 examined whether the interruptions caused by the blank fields used in the flicker paradigm interfered with the formation of a coherent representation. An 8 sec preview was given to observers before the interruptions began. No effect of preview was found. Such a result is a natural outcome of volatility. In contrast, it is difficult to explain if a coherent representation is formed in normal viewing: 8 sec ought to be enough to consolidate things to the point where at least some changes are detected more easily. This lack of preview effect indicates that if disruption is the cause of change blindness, it could only come about by the destruction of a coherence present almost from the initial appearance of the stimulus.

To determine if the interruptions used in the flicker paradigm could cause such destruction, Experiment 2 looked at the effect of the duration of the blank fields. If change blindness were caused by disruption, performance might deteriorate (or improve) with increasing duration, or be maximal at some particular value. In contrast, the volatility hypothesis would predict that performance should be similar for all durations, provided that the information from early-level detectors had been swamped.

The results followed a rather complex pattern: performance deteriorated as durations increased from 40 ms to 80 msec, remained constant up to 160 msec, and then deteriorated further at 320 msec. Both hypotheses can be made consistent with this pattern. The diminished blindness with the 40-msec durations can be explained either by a reduced ability to disrupt the early-level representations, or by a reduced ability of the global transients to swamp the local signals. Similarly, the increased blindness with the 320-msec durations can be explained either by an increased ability to disrupt the early-level representations, or by a degradation of iconic memory. The only evidence that helps to decide things is the finding of identical performance for the 80-msec and 160-msec durations. This result is somewhat unlikely if disruption is maximal at some particular time scale, but it is a natural outcome of the view that once attentional guidance has been disrupted nothing more will affect performance. But it is important to note that although this result favours the volatility hypothesis, it does not completely discredit disruption. To put things on a firmer basis, it would be useful to know more about the way that performance depends on the duration of the blanks.

Experiment 3 followed an alternate route—investigating how performance depends on the colour of the blank fields. The degree of change blindness was found to be independent of luminance level. The volatility hypothesis easily explains this independence: The global transients of the standard conditions are already sufficient to swamp the local signals, so that increasing their magnitude will do nothing more. The disruption hypothesis can also explain this pattern, but only by assuming that coherence is relatively fragile, so that performance with gray fields is already as poor as it could get. This assumption, however, was contradicted by the finding that red fields caused a further deterioration in detecting colour changes in marginal interests. This result is difficult (if not impossible) to account for in terms of disruption alone. But it can be easily explained by appealing to an early-level chromatic motion mechanism that helps draw attention to colour changes in the standard condition.

Yet another approach was used in Experiment 4, which replaced the solid fields with patches that did not cover the areas that were changing. As such, attention could be distracted from the change without disrupting the changing parts of the image. Results showed that although detection of change was faster than before, it was still impaired relative to the no-interruption condition. This behaviour is rather difficult to account for in terms of disruption. But it is easy to account for in terms of attentional distraction.

Taken together, these results show that change blindness is not an artifact of experimental manipulation, but instead is a phenomenon supporting the view that the visual system never forms a coherent, detailed representation of the world around us (Dennett, 1991; Grimes, 1996; Marr, 1982; O'Regan, 1992; Stroud, 1955). It could be argued that only the comparison of successive images was investigated, and so the conclusions drawn here apply only to comparison operations (see e.g. Scott-Brown, Baker, & Orbach, this issue). This would leave open the possibility of coherent representations that could support other kinds of operations. But changes can be detected by means other than direct comparison. For example, if successive images can be added together (superimposed), observers could find changes in position by looking for instances of "doubled" structure, where the original and the translated object existed side by side. The failure to easily detect translation indicates that operations such as superposition are not supported either. As such, it would appear that these results rule out the existence of a complete, detailed representation that could be described as coherent in any meaningful sense. (See also Rensink, this issue).

It is important, however, to keep in mind that a detailed representation does exist at each moment that our eyes are open, for we are obviously able to see that part of the scene being fixated. But although detailed, this representation is not temporally coherent—we do not have sufficient memory for the effortless combination of detailed information from successive moments in time. This is not to say that we have no visual memory. If we attend to something, we certainly can see it change. And if we close our eyes we can remember several things about the world in front of us, such as the location of a few of the objects present, their shape and size, and perhaps their colour. But the amount of visual detail contained in these representations is limited. Thus, although we have representations that contain large amounts of detailed information, and representations that provide coherence over time and space, we do not have representations that do both. Our impression of a world that is both coherent and detailed is therefore based only on the properties of the world itself, and not on the properties of the representations that underlie our visual experience.

REFERENCES

Aginsky, V., & Tarr, M.J. (this issue). How are different properties of a scene encoded in visual memory? *Visual Cognition, 7,* 147–162.

Aginsky, V., Tarr, M.J., & Rensink, R.A. (1997). The stability of color, location, and object presence in mental representations of natural scenes. *Investigative Ophthalmology and Visual Science*, *38*, S1009.

Ballard, D.H., Hayhoe, M.M., & Whitehead, S.D. (1992). Hand–eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society of London B*, *337*, 331–339.

Bischof, W.F., & DiLollo, V. (1995). Motion and metacontrast with simultaneous onset of stimuli. *Journal of the Optical Society of America, A*, *12*, 1623–1636.

Bischof, W.F., Seiffert, A.E., & DiLollo, V. (1996). Transient-sustained input to directionally selective motion mechanisms. *Perception, 25*, 1263–1280.

Braddick, O.J. (1973). The masking of apparent motion in random-dot patterns. *Vision Research*, *13*, 355–369.

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

Cropper, S.J., & Derrington, A.M. (1996). Rapid colour-specific detection of motion in human vision. *Nature*, *379*, 72–74.

Dennett, D.C. (1991). *Consciousness explained*. Boston: Little, Brown & Co.

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

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Vancouver studies in cognitive science: Vol. 5. Perception* (pp. 89–109). New York: Oxford University Press.

Henderson, J.M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, *8*, 51–55.

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

Irwin, D.E. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science*, *5*, 94–100.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219.

Klein, R., Kingstone, A., & Pontefract, A. (1992). Orienting of visual attention. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 46–65). New York: Springer.

Levin, D.T., Momen, N., Drivdahl, S.B., & Simons, D.J. (this issue). Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition, 7,* 397–412.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

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

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

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

Phillips, W.A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics*, *16*, 283–290.

Rensink, R.A. (1992). *The rapid recovery of three-dimensional orientation from line drawings*. PhD thesis (also Tech. Rep. No. 92–25), Department of Computer Science, University of British Columbia, Vancouver, BC, Canada.

Rensink, R.A. (1997). How much of a scene is seen? The role of attention in scene perception. *Investigative Ophthalmology and Visual Science*, *38*, S707.

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

Rensink, R.A., & Enns, J.T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review*, *102*,101–130.

Rensink, R.A., & Enns, J.T. (1998). Early completion of occluded objects. *Vision Research, 38*, 2489–2505.

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*, 368–373.

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

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

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

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

Stroud, J.M. (1955). The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology: Problems and methods*. (pp. 174–207). Glencoe, IL: Free Press.

Tversky, B., & Sherman, T. (1975). Picture memory improves with longer on time and off time. *Journal of Experimental Psychology: Human Learning and Memory, 104*, 114–118.

van de Grind, W.A., Koenderink, J.J., & van Doorn, A.J. (1986). The distribution of human motion detector properties in the monocular visual field. *Vision Research*, *26*, 797–810.

Wolfe, J.M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*, 202–238.

Woodhouse, J.M., & Barlow, H.B. (1982). Spatial and temporal resolution and analysis. In H.B. Barlow (Ed.), *The senses* (pp. 152–157). Cambridge, UK: Cambridge University Press.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science*, *5*, 156–161.