

Revised manuscript submitted to Visual Cognition, special issue on change blindness and visual memory.

Revision submitted November 1998.

# Some Thoughts on Change Blindness

J.K.O'Regan

Laboratoire de Psychologie Expérimentale, CNRS, EHESS, EPHE, Université René Descartes

28 rue Serpente, 75006 Paris, France

email: [oregan@ext.jussieu.fr](mailto:oregan@ext.jussieu.fr)

<http://nivea.psych.univ-paris5.fr>

## Abstract

*Recent results showing that large changes in a scene are not noticed if they occur at the same time as a global visual disturbance caused by saccades, flicker, "mudsplashes", or film cuts, are generally explained in terms of a theory in which it is assumed that the observer's internal representation of the outside world is very sparse, containing only what the observer is currently processing. The present paper presents some clarifications of the theory, and some new implications and predictions that arise from it.*

## Introduction

Recently a number of studies have shown that under certain circumstances, even very large changes can be made in a picture without observers noticing them (cf reviews by Introub, 1997; Simons & Levin, 1997). What characterises all the experiments showing such blindness to scene changes is the fact that the scene changes are arranged to occur simultaneously with some kind of extraneous, brief, global, disruption in visual continuity, such as the large retinal disturbance produced by an eye saccade (Grimes, 1996; Currie, McConkie, Carlson-Radvansky & Irwin, 1995; McConkie & Currie, 1996; Ballard, Hayhoe & Whitehead, 1992), a shift of the picture (Blackmore, Brelstaff, Nelson, & Troscianko, 1995), a brief flicker (Pashler, 1988; Phillips, 1974; Rensink, O'Regan, & Clark, 1995; 1997) also this issue; Zelinsky, 1997; 1998), a "mud splash" (O'Regan, Rensink & Clark, 1996), an eye blink (O'Regan, Deubel, Clark & Rensink, 1997; see also this special issue), or a film cut in a motion picture sequence (Levin & Simons, in press).

Current explanations of these phenomena are converging to a set of ideas concerning the nature of an observer's internal representation, the role of attention, and the role of visual transients which might be summarized as follows:

- (1) The internal representation of the visual world is much sparser than subjective experience would seem to suggest.
- (2) Attention is required to encode an aspect of the scene into this representation.
- (3) What an observer has the impression of seeing at a given moment is precisely what he or she is attending to at that moment.
- (4) Visual transients caused by scene changes can attract attention to the location of the change.

One example of where such ideas have been applied is in the work we have done recently on the flicker paradigm (Rensink et al., 1997; Rensink, O'Regan & Clark, this issue), but similar views are implicit in the work using other paradigms. As suggested in this work, the reason why changes are often not seen when they are accompanied by flicker or other disruptions is that the disruptions create visual transients that swamp the local transient that would normally be created by the change. Because of this, attention is not attracted to the change location and observers must resort to other strategies to locate the change. Either they must search serially through locations of the scene in search of something that is changing, or they must

rely on their memory of the scene to determine what the change is. In both cases, changes will only tend to be noticed if they occur at locations which in themselves are likely to attract attention because they are somehow "interesting" to the observer.

The purpose of the present paper is to develop a number of points concerning this theory which have not been considered in detail before, and which lead to speculations and predictions for future work. In other words, the present paper is not a new theory, but an attempt to clarify and extend the ideas that are already available in sketchier form in earlier papers.

## Thoughts on normal viewing

The first section will address some questions concerning how the principles of the theory, as listed above, would apply to the situation in which observers *normally* see scene changes, that is, when no simultaneous, experimentally imposed disruption is superimposed at the moment of the change. The section will begin with a disclaimer about philosophical issues, and will then discuss some preliminary details concerning the notions of "central interest" and "marginal interest" which were employed in earlier papers. I shall then come to the main point of this section which concerns the distinction between the "where" component and the "what" component of the mechanism underlying normal change detection. This distinction will reoccur at different points in the rest of the paper. It is important for methodological reasons and also gives rise to a counterintuitive prediction.

### Disclaimer on philosophical issues

#### *Normal "seeing": Visual solipsism*

Consider a situation where an observer is contemplating a scene. Over the course of a few seconds, he or she will attend to and process a number of aspects of the scene, and will encode them into some form of durable memory storage. A troubling consequence seems to arise if we accept the assumptions listed above, that (1) the internal representation is very sparse, and (3) observers only "see" what they are currently attending to in the scene: How can observers have the impression of seeing the *whole* scene simultaneously, and how can the scene appear so rich, if all that is being "seen" is the small bit that is being currently attended to? An explanation for this may be that the world serves as an "outside memory": Even though only a single aspect of the scene is being attended to at any moment, the observer nevertheless has the impression of seeing the whole scene, because if any particular part of the scene should awaken the observer's interest, attention can be swiftly turned to that aspect, and it will immediately come into visual awareness. It is this immediate availability at the slightest flick of attention or of the eye, which gives a kind of "solipsistic illusion"<sup>1</sup> of richness of the visual field, despite the paucity of the internal representation (cf. O'Regan, 1992).

The problem of the apparent richness of the perceived visual world is only one among several related, somewhat philosophical questions brought up by the phenomenon of change blindness. The visual "solipsism" argument just presented, seems to suggest that there is no "seeing" without attention. Under this view, the internal representation is relegated to the status of being just some kind of (short term or long term or "durable") memory store whose content is not "seen". But surely the whole point of the internal representation was for it to be the substrate for what we have the impression of seeing? Obviously clarification is needed on what is meant by "seeing", and on the exact role of the internal representation.

In the present article I will have to refrain from attempting to address these issues. I shall simply take the suppositions (1)-(4) as an axiomatic departure point, and try to examine in more detail where they lead. A justification for proceeding in this way would be that, just as it is said that the proof of the pudding is in the eating, if we can explain a lot of data with these axioms, then they have a good chance of being "correct", even if philosophically they are not quite clear.

### Comments on the Central/Marginal interest distinction

In our work using flicker, blink and mudsplash techniques, differences in the extent of change blindness were observed depending on whether the changing items corresponded to what we called "central" or "marginal" interest elements in the scene. The observed differences were important because they gave us valuable arguments against the hypothesis that change blindness was caused by some kind of wiping out of the internal representation provoked by the imposed scene disruptions -- presumably a low-level wiping out mechanism would incorrectly predict an equal wiping out of central and marginal interest items. Because of the importance of these arguments, it will be useful here to make the notion of "interest" more precise.

During examination of the scene, those particular aspects of the scene which will be attended to and encoded into the observer's internal representation will be determined by a variety of factors, ranging from low-level visual conspicuity (eg.

---

<sup>1</sup> Although see Noë, Pessoa & Thompson (this issue) who point out that strictly the use of the term "illusion" is inappropriate here.

contrast, curvature, layout and color distribution in the picture), through semantic factors (eg coherence, relevance, prior knowledge), to individual preferences related to the observers' interests of the moment. Suppose that A, B, C,... are aspects of a particular scene. A particular observer might attend to these aspects in the order: A, B, C, A, C, D, E, F, A, G, H,... A different observer, or the same observer on a different occasion, might attend to different items in a different order. But on average certain aspects of the scene will tend to be more often attended to and encoded than others, for example in this case A, B and C. The notions of "central" and "marginal" interest (Rensink et al., 1997) were based on this idea: we defined "central interest" aspects as those aspects that an observer will be most likely to attend to and encode, and "marginal interest" aspects being those that an observer is least likely to attend to and encode<sup>2</sup>.

A point which was perhaps not clearly made in our previous work was the fact that we considered this distinction between central and marginal interest to be merely a statistical, or operational way of defining aspects which observers will or will not generally attend to and encode. Theories of image comprehension might predict which items this might be, although factors of visual salience would additionally have to be taken into account. But for the purposes of understanding the phenomenon of change blindness, it is sufficient to assume that it is possible to classify scene aspects by order of likelihood of encoding.

### ***Locations, objects, or aspects?***

Another point about the central/marginal interest distinction is the following. It is tempting to assume that central and marginal interest aspects of the picture are *locations* in the picture. However evidence in the attention literature suggests that attention is probably better described, not in terms of space, but in terms of objects: that is, not as attaching itself on the basis of spatially defined, contiguous regions in the visual field, but rather to (possibly non-contiguous) collections of attributes that can be grouped together to form objects (Baylis & Driver, 1993; Driver & Baylis, 1989; Duncan, 1984; Duncan & Nimmo-Smith, 1996). The notion of "object file" is also related to this (Gordon & Irwin, 1996; Kahneman, Treisman, & Gibbs, 1992). Despite the literature on this topic however, the term "object" seems unsatisfactory: Presumably an observer can, for example, attend to the sky (which is not really an object), or to a particular aspect of a scene (say its symmetry or darkness) without encoding in detail all the attributes which compose it. Awaiting further clarification on this, it seems safer to say that attention can be directed to scene "aspects" rather than "objects".

### ***Looking without seeing***

An implication of this is as follows. Looking directly at a scene location, an observer may be encoding a variety of aspects. For example, looking at an ambiguous figure, an observer may be encoding a young girl or an old woman. Looking at the middle of a word, an observer may be checking the typography of the middle letter of the word, or alternatively estimating the word's length, or else recognizing it. In each case the observer may be oblivious to the other aspects of the stimulus.

The idea that it should be possible to be looking at a region of the scene, and only attend to a subset of the picture attributes that are there, is confirmed in an experiment we performed (cf. O'Regan, Deubel, Clark & Rensink, this issue) in which display changes were made simultaneously with observers' blinks. We found, by measuring eye movements while the observers searched for changes, that in almost 50% of the cases when an observer looked at a marginal interest location at the moment it changed, the change was not noticed, even though the eye was located less than 1 degree from the change.

Just as this example is an example of "looking without seeing", there is equally the possibility of "seeing without looking". The eye may for example be fixating at the center of a circle, where there is nothing to be seen, in order to check that the circle is round rather than an ellipse.

## **The "where" and "what" components of change detection**

Let us now return to the problem of seeing changes under normal circumstances, and discuss how the points (1)-(4) underlying theories of change blindness would be used to explain what happens when in normal, everyday life an observer sees a change in a scene.

### ***Transients tell "where"***

Consider an observer contemplating a scene, and suppose that suddenly something changes. As is the case in many signal-processing devices, sudden changes in input generally provoke disturbances or "transients"<sup>3</sup> which can be detected by specialized spatiotemporal detectors (cf. e.g. Breitmeyer & Ganz, 1976; Tolhurst, 1975; Klein, Kingstone & Pontefract,

---

<sup>2</sup> By hypothesis (3) we are assuming that the observer encodes only what he or she sees. A possibility will be mentioned later according to which other aspects of the scene might be encoded *implicitly* without the observer being aware of them. However these implicitly encoded aspects would presumably be encoded into a different storage buffer, not available for making conscious comparisons, and changes in such aspects would therefore not be visible.

<sup>3</sup> Analogously to electrical transients, that is, transitory large surges of current or voltage that occur in an electric circuit when it is turned on or off.

1992) that seem to have the special function of causing an "alert" that signals that a change has occurred (cf. e.g. (Yantis, 1993; Yantis & Jonides, 1990), and of exogenously attracting attention to the location where the change was.

Note that the occurrence of the transient associated with the change has allowed the observer to detect *where* the change occurred. Turning attention to the change, the observer now sees the new scene element. However, under our assumption (3) in the introduction, unfortunately the observer no longer sees the *old* scene element. There may be some information contained in the nature of the transient that enables the observer to deduce what the change must have involved: the "flavor" of the transient may help in guessing, say, that an object has shifted rather than changing color. But the transient "flavor" itself will usually not be sufficient to determine exactly what the change consisted in. For this, the observer must have some memory of what was at the changed position before: that is, the observer must have at some earlier moment encoded into the internal representation what was previously at the change location.

### ***Memory tells "what"***

In summary, the role of the transient corresponding to the change is to provide information on *where* a change occurs, and the role of the internal representation is to allow the observer to know *what* the change was. This distinction between the "where" component and the "what" component of normal change detection is an important one, and has not been sufficiently stressed in previous papers. One reason it is important is that it relates to the question of what is being measured in a change blindness experiment: failure of the "where" component, failure of the "what" component, or failure of both? If change blindness experiments are used to estimate the richness of observers' internal representations of the visual world, care must be taken to ensure that it is the "what" component that is being probed: observers must really be required to indicate the nature of the change they have detected, and not merely the location of the change.

### ***A counter-intuitive prediction for change blindness in normal viewing***

The distinction between the "where" and "what" components of change detection also leads to the following interesting consideration: previous work on change blindness has been restricted to experiments where some kind of global transient, like flicker etc., is presented simultaneously with the display change. But here we see that an implication of the theory is that, even when no such additional transients are added -- that is, under completely normal viewing conditions -- it should be the case that for what we have called marginal interest changes, though observers know *where* the change occurred, sometimes they should be unable to say exactly *what* the change was. This is because marginal interest aspects will tend not to be encoded. This prediction has not as yet been empirically tested. Indeed, *a priori*, the prediction seems extremely counter-intuitive: we normally have the impression that any change in a scene, even in a marginal interest aspect, is immediately visible.

An explanation of this incompatibility with common sense is perhaps first the fact that the prediction concerns marginal interest aspects: because marginal interest aspects of a scene are by definition not very significant to our conception of the scene, it may actually be the case that we are *not aware that we are not aware* of the exact nature of a change, and satisfy ourselves with imprecise, "where" knowledge. Another, perhaps more likely possibility, is that normally changes in a scene occur slowly enough for attention, or the eye, to reach them before they have completed. Finally, even when a change occurs too quickly for attention, or the eye, to reach it before it has completed, the "flavor" of the transient may often provide sufficient information about the change for the observer to feel satisfied he or she knows what it was.

On the other hand cases should nonetheless exist when attention or the eye cannot quickly reach the change, and where the "flavor" of the transient is not sufficient to guess what the change was. In these cases, observers, when interrogated, will have to admit that though they had the impression that something had changed in the scene, they cannot say exactly what the change was. This prediction, expected mainly for marginal interest aspects of a scene, would be worth testing.

## **Thoughts on disruptions**

Having considered what happens in normal everyday detection of changes in scenes, let us now turn to the situation generally studied in the literature on change blindness, in which, at the moment of the change, some kind of additional, brief disruption of visual continuity is imposed by say, the concurrent occurrence of flicker, blinks, eye saccades, etc.. The generally accepted theory explaining how such disruptions cause change blindness is to suppose that they primarily interfere with the "where" component of change detection: they create additional transients in the visual field which interfere with the attention-grabbing action of the local transient caused by the change: an observer's attention is no longer drawn immediately to the change location, and the observer must perhaps use a serial search strategy to locate the change (Rensink et al., 1997; this issue).

### ***Transients as masks and transients as distractors***

But note that the interference with the "where" component of change detection can occur in two ways: First, the additional transients can act locally by combining with the transient corresponding to the change, rendering it less "attention-grabbing".

This local combination can either involve luminance interactions or metacontrast-type masking effects, but in either case it makes the transient associated with the local change less salient and less effective in attracting attention to itself. Let us call *local masking* this way of impeding the "where" component of change detection.

A second way the brief disruption imposed by the experimental manipulation may interfere with the "where" component of change detection is by *diversion*. The local transient corresponding to the true change location is only one of a flood of additional, extraneous transients all over the scene. The "where" component of change detection is impeded because attention has no more reason to go to the "true" transient than to the many other, extraneous ones.

In most of the experiments performed in recent years on change blindness, both the local masking and the global diversion mechanisms may be active in interfering with the "where" component of change detection. This applies for example to the eye saccade, flicker, blink and film cut experiments, which involve large disturbances in visual continuity all over the visual field. Though in the past we have given more attention to the diversion mechanism, it is clear that for a good understanding of the change blindness phenomenon, it would be worth studying in greater detail the separate influence of both the local masking and the diversion mechanisms. A start in this direction has been made recently with the use of the "mudsplash" and "masking rectangle" manipulations.

### ***Measuring diversion with the mudsplash experiment***

Instead of using a global disturbance at the moment of the change, in the "mudsplash" paradigm (O'Regan, Rensink & Clark, in press; O'Regan et al. 1996) we briefly superimposed on the visual scene six small, high contrast shapes, somewhat like mudsplashes on a car windshield. The position of the mudsplashes was carefully chosen so that they did not cover the change location itself. In this experiment therefore, it was only through the mechanism of diversion that the imposed visual interruption interfered with the detection of the location of the transient associated with the change.

The results of the experiment confirmed that the diversion mechanism was sufficient to obtain a change blindness effect: a local masking of the transient associated with the change was not necessary.

This result is of course expected from the standard explanation of change blindness: the observer, not having a rich internal representation of the scene, attempts to rely on local transients to direct attention to the change location. But because there are now many, rather than just a single local transient, the "where"-mechanism of change detection is impeded, and observer's performance suffers.

### ***The (critical?) number of diversions***

An interesting question concerning the mudsplash experiment is the question of how many mudsplashes are necessary in order to prevent identification of the change. A first model would be to suppose that the effectiveness of the mudsplashes is determined merely by how many there are. The model would suppose that attention is attracted with equal probability to all the transients. The probability of moving to the transient corresponding to the true change would then just be  $1/(N+1)$ , where  $N$  is the number of diverting mudsplashes. A better model might weight the probabilities by the relative salience of the transients, perhaps determined by their brightness, color, contrast, or other discriminating quality.

An alternative intuition one might have concerning the mudsplash paradigm would be related to the claim that there may be a critical number of events (say four or five) that attention can be monitoring simultaneously (cf Pylyshyn & Storm, 1988; Wolfe, Cave, & Franzel, 1989). It might therefore be that when fewer than four or five transients occur in the scene, verification of whether they correspond to a true scene change can occur easily, but when the number exceeds this critical value, change detection would suddenly break down. On the other hand, this view probably requires the notion of a visual buffer in which "what is currently being seen" is held: in that case it would not be favored by the approach suggested here. Indeed, in pilot experiments presented in O'Regan (1998), I have found preliminary evidence against the view. Observers attempted to detect a letter that changed within an "alphabet soup" of scattered letters. The change occurred at the same time as a number of mudsplashes were scattered on the scene. The results showed that few large mudsplashes had an effect similar to many small mudsplashes, suggesting that number of mudsplashes in itself is not the determining factor.

### ***Proximity of the transient***

An additional question concerning the diversion mechanism is the proximity of the diverting transients in relation to the true change position. It might be expected that a change would actually be more likely to be detected if an irrelevant transient occurs at a location quite near the true change location, since now the transient is attracting rather than diverting attention from the area of the change. On the other hand, as we have seen from the idea that attention may not necessarily enclose all aspects of a region, and as suggested by the results of our "blink" experiment, one might equally well postulate the opposite, namely that spatial proximity would not be a direct determiner of probability of detection. This question would merit further investigation.

### ***More questions on diversions***

Many further interesting questions remain to be answered concerning how diversion can interfere with the "where" component of change detection. For example, can an observer learn to ignore diversions that have a known attribute or attend to

transients that have a known attribute. If mudsplashes are always in the same locations, or if they are always of the same distinctive color, shape, or size, then an observer might be able to resist the tendency to orient attention to the diverting location. Similarly, if the local transient corresponding to the true change has some feature that makes it stand out within the flood of irrelevant local transients caused by the diversions, then change detection might become easier. On the other hand it might be the case that the sudden onsets caused by transients cause an irrepressible attention-grabbing effect, within which no selective mechanisms can operate. If the attention-grabbing action of the transients is an automatic, irrepressible, low-level mechanism, then it might be expected that the force of attraction might not depend on a computation determined by the other transients in the scene.

Another question concerns what kind of search strategy observers perform among the locations which attract attention. Is there an inhibition of return mechanism which prevents attention returning to a previously examined location? If so, how many previously visited locations can the inhibition of return mechanism keep in memory? Can this mechanism operate even though, unlike what happens in normal search tasks, the previously visited locations are repetitively creating attention-grabbing transients?

### ***A transient pop-out task?***

It must be stressed that although the mudsplash experiment used a purely diversional method to interfere with the "where" component of change detection, the measurements of change detection probability obtained from the experiment are still not a pure indication of the "where" component. The reason is that the "what" or encoding component of change detection presumably also played a role: depending upon whether observers construed the task they were accomplishing as consisting of merely indicating *where* the change occurred, or as consisting of actually determining exactly *what* the change involved, the role of observers' internal memory representations of the scene will be more or less great.

Indeed, probably a purer way of investigating how diversion contributes to change blindness would be to use a task where the "what" or encoding component was not solicited at all. A possibility would be to use a kind of "pop-out" task in which observers have to judge if, among a set of many transients provoked for example by mudsplashes, one in particular appears as being different from the others and pops out.

### ***Does local masking interfere with the "what" component?***

The mudsplash experiment constituted an approach to the study of the diversion mechanism underlying the change blindness effect. In the experiments in which there are global disturbances like eye saccades, flicker, blinks, and film cuts, there is also a local masking transient which, by combining with the transient at the change location, may decrease its salience. It would be interesting to study the relative weight, in interfering with the "where" component, of this local masking transient compared to the diversion mechanism.

In particular, an important point concerns the possibility that the local masking transient actually somehow does more than just interfere with the "where" component of change detection: There is also the possibility that the local, masking transient actually interferes with the "what" or encoding component, by wiping out the internal representation itself. For example it could be postulated that the transient provokes a kind of "reset" of the internal representation, in preparation for reception of new incoming information.

If this were true then the basic tenet of our explanations of change blindness, namely that the internal representation is very sparse, could be discarded: we could claim that the internal representation is actually very rich, but that it is wiped out by the local masking transient.

However this hypothesis has been addressed and rejected by several arguments in (Rensink et al., 1997), and in Rensink, O'Regan & Clark (this issue), and O'Regan, Rensink & Clark (in press). One of the main points is that it is incompatible with the finding that central and marginal interest scene aspects suffer different amounts of change blindness: a wiping out process would presumably wipe out both kinds of change equally. The mudsplash result, where there is no local masking at all, also argues against this claim.

### ***Prediction for very slow changes.***

If we accept that the local masking effect imposed by flicker, etc., only acts on the "where" component of change detection, reducing the salience of the transient caused by the change, then it should be possible to manipulate the extent of this interference. One intriguing possibility would be if it were possible to arrange a situation in which aspects of a visual scene change so slowly that they do not generate visual transients at all. We would predict that changes created in this way would not be noticed if they were marginal interest aspects. Supporting this claim, D. Levin (personal communication) has constructed film sequences where the setting in which the action takes place gradually changes without the observers noticing it.

This situation is perhaps similar to what happens when one looks at the hands of a watch: seeing a change in position can be done only by attributing a new code or classification to the current position: e.g. the minute hand was exactly on the 30-

minute mark, but now has passed it. The same could be said of flowers: suddenly one is aware that they have wilted and need water, even though they are presumably slowly wilting all the time. In both cases, in order to "see" such a slow change, one has to classify (encode) the new situation and judge it to be different from the previously encoded situation.

### *Estimating the "what" component of change detection using the masking rectangle experiment*

Whereas the mudsplash experiment was (mainly) a way of studying the "where" component of change detection, the masking rectangle experiment (cf. O'Regan, Rensink & Clark, in press; O'Regan, Rensink & Clark, 1996) was a way of looking at the "what" or encoding component.

Instead of diverting attention from the change location by means of extraneous mudsplashes, a black-and-white checkered rectangle was flashed over the area of the change while the change occurred. A large transient was therefore generated at the change location, presumably causing observers' attention to move to that region.

In this experiment, the location of the change is perfectly obvious, since it is indicated by the rectangle. The observers' task was thus clearly defined, and was to say what was present before the masking rectangle appeared. Since the masking rectangle was of very high contrast and completely covered the change location, the "flavor" of the transient provided no information about what was previously at the change, and subjects had to rely entirely on what information they had encoded prior to the appearance of the masking rectangle.

Coherent with the theory of sparse internal representation, we found that, particularly for the case of marginal interest changes, observers were often unable to report what the change was. Note that since in this experiment there are no diverting extraneous transients, the effects observed must be due wholly to the "what" (or encoding) component of the change blindness phenomenon, and not to any part caused by a "where" related diversion component. The masking rectangle experiment therefore has the advantage over most other techniques used in change blindness research, of representing a "pure" measure of the content of the internal representation<sup>4</sup>.

Note that in fact it would have also been possible to do the experiment without a masking rectangle at all: simply make the change directly. This corresponds to the situation discussed above where a scene change occurs in normal viewing without any additional visual disruption. However a disadvantage of using this technique in order to estimate the content of the internal representation is that since the transient caused by the change is not masked in any way, it will contain some information about the change that has taken place. Observers may make use of this "flavor" of the transient to deduce what the change might have been -- for example a shift in an object may produce motion energy, which will be a different kind of transient from that produced by a color change, say.

## **Other issues concerning the theory**

In the following sections I will discuss some additional points about the explanation of change blindness based on the notion that the internal representation of the visual scene is sparse.

### *A prediction for the moment of change detection*

An interesting, so far unexplored consequence of this way of understanding change detection in the flicker, saccade, blink and film cut experiments concerns the time course of exploration of a scene. Let us recall what we assume happens when an observer examines a scene. Attention is directed sequentially to different aspects of the scene, and these aspects are encoded into a categorical, more durable memory store. What determines the order with which different scene aspects will be encoded is presumably low-level factors such as their visual salience, as well as high-level factors related to the process of object and scene identification. Suppose that the sequence of exploration for a particular scene involves attention being directed sequentially to aspects A, B, C, D, E, F, G, ..., labelled in descending order of "interest". The basic hypothesis of a sparse internal representation that we have adopted postulates that once a scene aspect A has been encoded, and once attention has moved onwards to aspect, B, the original scene aspect A is no longer being attended to and processed. If a change should now occur in A, and if conditions are arranged so that the visual transient created by the change is masked or camouflaged by one or other of the techniques currently being used in the change blindness literature, then the change should not be noticed. On the other hand, if the change had occurred while A was being encoded, the change would be noticed.

Since central interest aspects of a scene are presumably precisely those which will tend to be attended to at the beginning of picture contemplation, we can therefore make the curious prediction that even a change in a central interest aspect of a scene may be missed if the change occurs after the scene has already been inspected for a while. In other words, we would expect that changes in the most significant aspects of a scene would be easier to detect *early* in the period of contemplation of the

---

<sup>4</sup> Note that it must be assumed that the local masking caused by the masking rectangle does not wipe out the internal representation. The arguments against this were mentioned above.

scene, and harder to detect *later* in the period of contemplation. On the other hand, the opposite would be true of the less significant, marginal interest aspects of a scene.

### *A prediction: seeing illusory appearances*

A related prediction from this explanation of the scene change experiments is that under some circumstances it should happen that an observer claims that a change has occurred in a picture when actually no change occurred at all. Suppose again that a picture contains aspects A, B, C, D, E, F, G, ..., in descending order of "interest". Suppose that before the transient the observer has encoded only the most "interesting" aspects, namely A, B, C and D. Suppose that now a transient occurs, but that there is no change in the picture. It could now happen that on scanning the scene after the transient, the observer notices aspect E which had not been previously encoded. The observer might well incorrectly deduce that this aspect had appeared during the transient. It is noteworthy that this kind of error would tend only to involve illusory appearances and not illusory disappearances.

### *A special role for layout?*

Recently some studies have addressed the question of the particular role of layout in scene change experiments (Simons, 1996; Simons, in preparation). Under the conditions of these experiments, changes in layout are quite easy to detect, whereas changes in objects are harder to detect. This particular finding may perhaps be due to increased visual salience of the layout change. But additional findings, showing that detection of layout changes is not affected by verbal interference, whereas detection of object changes is; as well as other results showing a differential effect of changing the observer's viewpoint, all suggest that in fact layout plays a special role in perception.

On the other hand, one might make the following argument. It is clear that objects are never recognized in isolation. Modern theories of vision often suppose that visual analysis proceeds simultaneously at several spatial scales, so that information about an object always has associated with it information about the surroundings within which the object is situated. An analogy with music is appropriate: a melody played by a violin is perceived quite differently if the violin is playing as part of an orchestra.

If this is true, then the distinction between layout and object becomes less clear. In a task where the observer is looking at a scene consisting of objects A, B, C, D, E, within a layout L, it could be that observers attending, say, first to object A and then to object C, are in fact coding A+L, followed by C+L. Since the layout is always being encoded and put into memory storage, no matter what part of the scene the observer is processing (let us say it is element X) when a change subsequently occurs, then what is encoded after the change also involves the layout: X+L. If the layout has changed to L', then the observer will encode X+L' and the change in layout will always be noticed.

Despite this objection, it is nevertheless tempting to attribute a particular role to layout. Intuitively it seems reasonable to suppose that the layout of a scene represents a sort of framework within which a picture is perceived. The situation may be similar to what is postulated in linguistics, where a distinction is sometimes made between "given" and "new" information. The "given" information is often the subject of a sentence, that is, what is known in advance and to be commented upon, and the new information or "topic" constitutes what is to be added to the "given" information (cf. eg. Chafe, 1970; Haviland & Clark, 1974).

Another point is the following. There are reasons to believe that layout, because it constitutes an aspect of the environment which may be used in locomotion, navigation, and sensorimotor coordination, may be processed by different mechanisms, and may have a special status in determining perception (eg. Jeannerod, 1997; Milner & Goodale, 1995). It is possible that the information-selection processes needed for processing layout may therefore be distinct from those that are used to make (verbal) decisions, judgments and commentaries about objects. Indeed, if we say that what we mean by seeing is merely what is currently being processed, usually with a view to making decisions, judgments or commentaries about a visual stimulus, and if a separate, mainly motor-control oriented process is usually dealing with layout, then why should layout changes be seen at all?

A possibility is that the separate, layout-processing mechanism, when it detects a change, creates some kind of attentional "alert" signal which is registered by the whole organism, including the processes that underlie what we usually call seeing. If this were so, then we would make the interesting prediction that the *fact* that a layout change has occurred might be registered, but not the exact nature of the layout change. Furthermore it might be expected that such an alert would only occur if the sensorimotor actions that would potentially be affected by the layout change were substantially altered. It might be possible to arrange situations in which two visually equally salient layout changes had different significance for sensorimotor coordination and thereby differed in detectability.

It is of course possible to arrange conditions where an observer must make (conscious) decisions and commentaries on scene layout, and in that case it must be the case that the observer really is "seeing" the layout. It would be interesting to see if the same layout change would be detected differently depending on whether the layout was being visually attended to (ie being "seen") or was being used implicitly in a sensorimotor task. It might be possible to arrange conditions where, when the observers are asked to visually attend to layout (ie "see" it), they do *less* well in detecting *that* a change has occurred than when layout is being made use of only implicitly in a sensorimotor task.



### ***Implicit knowledge of changes?***

An intriguing question about the scene change experiments is the question of whether, despite the fact that a change might not have been consciously noticed, it might nevertheless have been implicitly recorded so that it affects subsequent behavior.

One possibility concerns the low-level modules which analyse the incoming information in the first stages of visual processing. Since these modules will have automatically registered and processed all the information in the scene, some adaptation or modification of their functioning may occur through the mere presence of the information, and this may modify subsequent conscious or unconscious processing of the scene.

Another possibility relates to the fact that the visual system is not a unitary system. It is quite possible that, independently of the process which underlies conscious seeing, other processes (for example concerned with maintaining posture, adapting the grasp, controlling eye movements etc.) will have made use of the information (cf. Jeannerod, 1997; Milner & Goodale, 1995). Within these processes, some memory of the information, or at least some adaptation to its presence, may therefore have occurred, and this might affect the behavior of certain subsystems at a later time. For example, it is possible that eye movement scanning of the scene will be modified by the presence of the unseen elements. Some results of (Hayhoe, Bensinger, & Ballard, 1998) are in support of this prediction, since they show that even though observers did not notice changes in blocks that they had to assemble in a block-copying task on a computer screen, their eye fixation durations were nevertheless modified.

### ***Other frameworks for explaining change blindness***

In some recent papers (e.g. Rensink et al., 1997) an explanation of change blindness has sometimes been phrased in terms of the idea that attention is needed to see *changes*. Although this would be a possibility, in fact the theory which we have presented actually makes a more drastic claim, which is that attention is not only needed to see *changes*, but to see *anything at all*. Is there a real difference between the two views? As mentioned by (Pashler, 1995) it is not clear whether it actually makes sense to postulate a model of vision in which observers see everything in a scene, but, where, when a change occurs, they cannot see it unless they are attending to it. Such a view seems rather strained, and furthermore it runs the risk of espousing the "philosophically incorrect" position according to which there is an internal picture-like representation of the world which corresponds to what observers are currently seeing. It seems more theoretically coherent to take the extreme view that nothing is seen unless it is attended to. This is also the view taken by Mack & Rock (1998) in the context of their studies of "inattention blindness".

Another question is raised by the proposition of Wolfe's (1997a,b; Wolfe, Klempe & Dahlen, submitted) according to which change blindness phenomena, inattention blindness, as well as phenomena that he has described using a "repeated visual search" paradigm, can all be understood in terms of the idea that everything is seen, but only what is attended to is remembered. Wolfe calls this "amnesic vision" or "inattentional amnesia". Again, as pointed out by Wolfe, Klempe & Dahlen (submitted), the distinction between the two approaches is largely a question of philosophical preference. Nevertheless Wolfe argues in favor of his view by claiming that, first, it accounts better for the subjective impression of visual presence that we have, and that, second, it accounts for the fact that in change blindness experiments, changes can be missed even when they are being directly attended to. However these arguments can be countered: first, as mentioned above (cf. also O'Regan, 1992) the impression of complete visual presence that our subjective experience provides us with may be a sort of "solipsistic illusion" created by the fact that the slightest desire to see any part of the scene is immediately satisfied by a flick of the eye or of attention. The second counterargument to Wolfe is that in change blindness experiments where apparently an attended change is missed, the observer, though looking directly at the change location (as in Levin & Simons' (in press) film cuts and O'Regan et al.'s (this issue) blink experiment), he or she might have been attending to an *aspect* of the scene at that location which was not the aspect that actually changed.

### ***Relation to early literature on partial report***

It is interesting to consider the relation between the recent data on change blindness and the considerable literature that has accumulated on "iconic memory", since (Sperling, 1960) and (Averbach & Coriell, 1961) performed their classic experiments on partial report. Pashler (1998) in a review of this literature concludes that there is agreement about the fact that there probably exist two distinct forms of memory in the visual system: a high-capacity sensory memory which has a life-time of the order of 100 ms, but which is sensitive to masking, and a low-capacity, more durable memory (usually called short term memory), not sensitive to masking. A typical partial-report experiment involves extracting information from the sensory memory and transferring it into the more durable storage, and the observed phenomena can be adequately modelled by assuming that the choice of what is transferred is determined by attentional set, cues, and task instructions (Gegenfurtner & Sperling, 1993).

From such a point of view, what would be predicted concerning the present experiments? When a scene is presented to an observer, information from the scene impinges on the high-capacity sensory memory, and those parts of the scene that the observer wishes to process start being transferred into the more durable storage. After a while this durable storage becomes full, and nothing more can be transferred. Now a change occurs in the scene. Because of the transients produced by the different experimental manipulations (saccades, flicker, mudsplashes, etc.), the observer has no cue indicating which part of the picture should be processed. Only if the observer, for whatever reason, happens to decide to process the part of the scene that actually

changed, and (2) he/she happened to have encoded that part of the scene into durable storage before the scene changed, will a change be detected.

This is exactly the same analysis of change blindness as has been proposed in the change blindness literature. There is just one difference in the approaches, which is that in the iconic memory literature it is often implicitly supposed that the icon is what observers have the impression of seeing (this is similar to Wolfe's "amnesic vision" view), whereas here we prefer to suppose that observers only see what they are processing. Essentially, if we replace the notion of "more durable storage", or "short term memory" with "what the observer is currently processing", and if we assume that the icon is not what is seen, then the two approaches are identical.

Pashler (1995) has also considered these questions and alluded to the fact that these distinctions may be too philosophical to be tested. However he suggests a possible line of empirical investigation based on experiments in which he studied the effect of allowing the partial report stimulus array to be previewed prior to the appearance of the report cue (Pashler, 1984).

## Conclusion

The present paper has examined in detail the implications of the four main assumptions made in current explanations of change blindness. Some distinctions which seem important for future work have been pinpointed, and some new experimental predictions have been made.

Perhaps the most important distinction that was made in the paper, and which permeated the reasoning throughout, concerned not just the change blindness situation, but also the situation where in normal, everyday life, a change occurs in the visual scene. The idea, which though quite clear and well-known, had not previously been sufficiently stressed and developed, is that the process of detecting a change in the visual field involves two components: a "where" component, and a "what" component. The "where" component provides information about the location of the change in the visual field, and is signalled by the local transient caused by the change. The "what" component allows the exact nature of a change to be ascertained, and involves the use of information previously encoded in a durable memory store of some kind. The phenomenon of change blindness occurs because one or both of these components of normal change detection is interfered with.

Making precise the distinction between the "where" and "what" components of change detection led to the realization that even under normal viewing conditions, that is, when there are no experimentally imposed disruptions of the visual field, change blindness might still be found for certain marginal interest changes. Change blindness should also be found in the cases where changes are so slow that they do not create salient transients.

The where/what distinction also raised a methodological issue: To the extent that the precise task demanded of subjects in change blindness studies is sometimes not sufficiently clearly specified, these studies may not provide a pure measure of the content of the observer's internal representation. Future work should carefully control whether subjects are simply asked to locate a change, or whether they are also required to make precise judgments about the nature of the change.

Another series of points raised in the paper concerned the precise mode of action of the experimentally imposed scene disruptions (eye saccades, flicker, blinks, etc.) in the change blindness experiments. These seem primarily to interfere with the "where" component of change detection, but do they do so by local masking or by diversion? Within the diversion mechanism a number of questions for future work were raised.

Another consideration in the paper was to make more precise the definition of "central" and "marginal" interest aspects of a scene which we had used in prior work. This led to a prediction about how change blindness might depend on the moment at which the change occurs, and to a prediction about the possibility of illusory appearances of changes.

Other points developed in the paper concerned the question of the role of layout, and the possibility of implicit perception of changes. Finally some comments were made about other theories of change blindness and their relation to the early literature on partial report.

Acknowledgements: I thank R. Rensink, J. Clark, for teamwork during the experimental stages of this research, and Nissan Cambridge Basic Research for supporting it. I thank H. Ben Salah, S. Chokron, H. Deubel, V. Gautier, A. Gorea, T. Nazir, A. Noë, and J. Wolfe for encouragement and comments, and two anonymous reviewers and, especially, D. Simons for their patience and help in improving the manuscript.

## REFERENCES

- Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. Bell Systems Technical Journal, 40, 309-328.
- Baylis, G. C., & Driver, J. (1993). Visual attention and objects: evidence for hierarchical coding of location. J Exp Psychol Hum Percept Perform, 19(3), 451-470.
- 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.
- Blackmore, S. J., Brelstaff, G., Nelson, K., & Troscianko, T. (1995). Is the richness of our visual world an illusion? Transsaccadic memory for complex scenes. Perception, 24(9), 1075-1081.
- Breitmeyer, B.G. & Ganz, L. (1976) Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. Psychological Review, 83, 1-36.
- Chafe, W. (1970) Meaning and the structure of language. Chicago: University of Chicago Press.
- Currie, C., McConkie, G.W., Carlson-Radvansky, L.A., & Irwin, D. E. (1995) Maintaining visual stability across saccades: role of the saccade target object. Technical Report No. UIUC-BI-HPPP-95-01. Champaign: Beckman Institute, University of Illinois.
- Driver, J., & Baylis, G. C. (1989). Movement and visual attention: the spotlight metaphor breaks down [published erratum appears in J Exp Psychol Hum Percept Perform 1989 Nov;15(4):840]. J Exp Psychol Hum Percept Perform, 15(3), 448-456.
- Duncan, J. (1984). Selective attention and the organization of visual information. J Exp Psychol Gen, 113(4), 501-517.
- Duncan, J., & Nimmo-Smith, I. (1996). Objects and attributes in divided attention: surface and boundary systems. Percept Psychophys, 58(7), 1076-1084.
- Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. Journal of Experimental Psychology: Human Perception & Performance, 19(4), 845-866.
- Gordon, R. D., & Irwin, D. E. (1996). What's in an object file? Evidence from priming studies. Percept Psychophys, 58(8), 1260-1277.
- Grimes, J. (1996) On the failure to detect changes in scenes across saccades, in Perception: Vancouver Studies in Cognitive Science, Vol 2 (Akins, K. ed.), pp. 89-110, OUP.
- Haviland, S.E., & Clark, H.H. (1974) What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior, 13, 512-21.
- Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. Vision Research, 38(1), 125-137.
- Intraub, H. (1997). The representation of visual scenes. Trends in Cognitive Sciences, 1, 217-221.
- Jeannerod, M. (1997). The cognitive neuroscience of action. Oxford, England UK: Blackwell Publishers, Inc.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. Cognit Psychol, 24(2), 175-219.
- Klein, R., Kingstone, A. & Pontefract, A. (1992) In K. Rayner (ed.) Eye movements and visual cognition: Scene perception and reading New York: Springer. pp. 46- 65
- Mack, A., & Rock, I. (1998). Inattention blindness. Cambridge, MA, USA: The Mit Press.
- McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. Journal of Experimental Psychology: Human Perception & Performance, 22(3), 563-581.
- Milner, A. D., & Goodale, M. A. (1995). The visual brain in action. Oxford, England UK: Oxford University Press.

Noë, A., Pessoa, L., & Thompson, E. (in this issue) Beyond the grand illusion hypothesis: What change blindness really teaches us about vision. Visual Cognition, ???

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. (1998) Detecting scene changes: an overview and a framework for recent findings" (Abstract) Perception, 27(suppl.): 36. [ECVP 1998; Oxford, England.]

O'Regan, J.K., Rensink, R.A. & Clark, J.J. (1996) "Mud splashes" render picture changes invisible. Invest. Ophthalmol. Vis. Sci. 37 S213.

O'Regan, J.K., Deubel, H., Clark, J.J. & Rensink, R.A. (1997) Picture changes during blinks: Not seeing where you look and seeing where you don't look. Invest. Ophthalmol. Vis. Sci. 38, S707.

O'Regan, J.K., Deubel, H., Clark, J.J. & Rensink, R.A. (this issue) Picture changes during blinks: looking without seeing and seeing without looking. Visual Cognition, ?????

O'Regan, J.K., Rensink, R.A. & Clark, J.J. Blindness to scene changes caused by "mudsplashes". Nature, in press.

Pashler, H. (1984). Evidence against late selection: Stimulus quality effects in previewed displays. Journal of Experimental Psychology: Human Perception & Performance, 10(3), 429-448.

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

Pashler, H. (1995). Attention and visual perception: Analyzing divided attention. In S. M. Kosslyn, Osherson, Daniel N. (Ed.), Visual cognition: An invitation to cognitive science, Vol. 2 (2nd ed.). An invitation to cognitive science (pp. 71-100). Cambridge, MA, USA: Mit Press.

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

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. Spatial Vision, 3(3), 179-197.

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., O'Regan, J.K. & Clark, J.J. (1995) Image flicker is as good as saccades in making large scene changes invisible. Perception, 24 (suppl.) 26-27.

Rensink, R.A., O'Regan, J.K. & Clark, J.J. (this issue) On the failure to detect changes in scenes across brief interruptions. Visual Cognition, ?????

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

Simons, D. J. (in preparation). Representing spatial layout across viewpoint changes. Manuscript in revision.

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

Sperling, G. (1960). The information available in brief visual presentations. Psychological Monographs, 74((11, Whole No. 498)).

Tolhurst, D. J. (1975). Sustained and transient channels in human vision. Vision Research, 15, 1151-1155.

Wolfe, J. M. (1997a). Visual experience: Less than you think, more than you know. In C. Taddei-Ferretti (Ed.), Neuronal basis and psychological aspects of consciousness, (pp. ???). Singapore: World Scientific.

Wolfe, J.M. (1997b) Inattentional amnesia, In V. Coltheart (Ed.) Fleeting Memories. Cambridge, MA: MIT Press, pp. ???-???

Wolfe, J.M., Klempe, N. & Dahlen K. (submitted) Post-attentive vision.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception & Performance, 15(3), 419-433.

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. Journal of Experimental Psychology: Human Perception & Performance, 19(3), 676-681.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. Journal of Experimental Psychology: Human Perception & Performance, 16(1), 121-134.

Zelinsky, G.J. (1997) Eye movements during a change detection search task. Invest. Ophthalmol. Vis. Sci. 38/4, S373.

Zelinsky, G.J. (1998) Detecting changes between scenes: a similarity-based theory using iconic representations. Beckman Institute for Advanced Science & Technology, Technical Report No. CNS-98-01.