PERGAMON

Vision Research 40 (2000) 1469–1487

Vision Research

www.elsevier.com/locate/visres

Seeing, sensing, and scrutinizing☆

Ronald A. Rensink *

Cambridge Basic Research, Nissan Research & Development, Inc., 4 Cambridge Center, Cambridge, MA 02142-1494, USA

Received 8 May 1999; received in revised form 23 December 1999

Abstract

Large changes in a scene often become difficult to notice if made during an eye movement, image flicker, movie cut, or other such disturbance. It is argued here that this *change blindness* can serve as a useful tool to explore various aspects of vision. This argument centers around the proposal that focused attention is needed for the explicit perception of change. Given this, the study of change perception can provide a useful way to determine the nature of visual attention, and to cast new light on the way that it is — and is not — involved in visual perception. To illustrate the power of this approach, this paper surveys its use in exploring three different aspects of vision. The first concerns the general nature of *seeing*. To explain why change blindness can be easily induced in experiments but apparently not in everyday life, it is proposed that perception involves a *virtual representation*, where object representations do not accumulate, but are formed as needed. An architecture containing both attentional and nonattentional streams is proposed as a way to implement this scheme. The second aspect concerns the ability of observers to detect change even when they have no visual experience of it. This *sensing* is found to take on at least two forms: detection without visual experience (but still with conscious awareness), and detection without any awareness at all. It is proposed that these are both due to the operation of a nonattentional visual stream. The final aspect considered is the nature of visual attention itself — the mechanisms involved when *scrutinizing* items. Experiments using controlled stimuli show the existence of various limits on visual search for change. It is shown that these limits provide a powerful means to map out the attentional mechanisms involved.

Keywords: Eye movement; Image flicker; Change blindness; Attention; Scene perception; Visual memory

1. Introduction

Once upon a time it was widely believed that human observers built up a complete representation of everything in their visual field. More precisely, it was believed that the stable and richly-detailed world viewed by a normal observer gave rise to a stable and richly-detailed representation within, a 'picture' that could be used for all subsequent visual and visuomotor operations (e.g. Feldman, 1985; Trehub, 1991). This representation was thought to be long-lasting and spatiotopic, its contents built up by the superposition of eye fixations, and the results held in a high-capacity buffer. With such a visual system, presumably, the observer could live happily ever after.

This idea of a stable, general-purpose internal picture accords nicely with subjective experience. It does not, however, accord quite so nicely with objective fact. For example, evidence is accumulating that visual processing is split into two largely independent streams — one concerned with the perception and recognition of objects, the other with visuomotor action (e.g. Milner & Goodale, 1995). And recent studies indicate that perception itself is susceptible to various forms of induced blindness: repetition blindness (Kanwisher, 1987), inattentional blindness (Mack & Rock, 1998), change blindness (Rensink, O'Regan & Clark, 1997), and an attentional blink (Raymond, Shapiro & Arnell, 1992). Taken together, these results show that the failure to experience highly-visible stimuli is not a rare, pathological condition. Rather, it is a much more widespread effect — so widespread that it strongly suggests that observers never form a complete representation of the world around them.

☆ Portions of this manuscript were presented at the 3rd Annual Vision Research Conference, Ft. Lauderdale, FL, 7–8 May 1999. Demos of the basic effect described here are available at http://www.cbr.com/projects.

* Tel.: +1-617-3749650; fax: +1-617-3749697.

E-mail address: rensink@cbr.com (R.A. Rensink)

0042-6989/00/$ - see front matter
PII: S0042-6989(00)00003-1

Throwing out the notion of a stable internal picture[1] has important consequences for vision research, both at the theoretical and the methodological level. To begin with, current research is concerned largely with operations on retinotopic representations, operations that take place within a single fixation of the eye. A common (and often implicit) assumption in vision science is that such operations are of primary interest, with their integration over eye movements being an uninteresting topic safely left for later study. However, if an all-encompassing picture is never formed, it becomes important to determine the extent to which representations *do* capture and hold on to properties of the world. It also becomes important to determine how their formation is co-ordinated with the task at hand, and to establish the particular ways in which they are used (Appendix A). Such issues must become a primary concern of vision science, requiring an approach complementary to the traditional one focused on within-fixation processing (see Churchland, Ramachandran & Sejnowski, 1994; Findlay, 1998).

But if vision is to be studied this way, what kinds of theoretical constructs should be adopted? What kinds of experimental approaches should be used? The purpose of this paper is to show that there already exist theoretical and experimental tools that can help with this task. To this end, reviews will be presented of several recent advances based on *change blindness*, the finding that observers often fail to detect large changes in a display if these are made at the moment that transients appear (e.g. Rensink et al., 1997; Rensink, O'Regan & Clark, 2000). It will be shown that the theoretical and empirical tools originally used to explore this effect, and which helped demolish the idea of a stable internal picture, can also be harnessed to more constructive ends. In particular, it will first be shown that these tools can provide a means of exploring the general nature of *seeing* (i.e. the way that humans use light to pick up information about their surroundings). It will then be shown that these tools can also cast light on the more specialized faculties of *sensing* (i.e. the pickup of information not accompanied by visual experience) and *scrutinizing* (i.e. the pickup of information via focused attention).

[1] In this paper, the term 'stability' is used in two ways. In its spatial aspect, it denotes invariance over position of each particular eye fixation; in its temporal aspect, it denotes invariance over the time of each particular fixation, i.e. a representation that is not volatile. Both these properties are required for a buffer that collects information into a representation that is independent of eye position. Meanwhile, 'complete' refers to a representation that is not only stable, but describes the entire visual field at a resolution equal to that of foveal vision. Given that 'picture' refers here to a high-resolution representation, a stable picture is also a complete representation.

2. Change blindness

...the ordinary person at once responded to the meaning and grouping of familiar scenes...I then sometimes substituted for the original figure a new one, changed in some respect. Here again, it would be worth experimenting in more detail to discover what types of change are most readily reacted to.

Bartlett (1932)

Change blindness can be defined as the failure of observers to detect large, sudden changes in a display (Rensink et al., 1997). Although such changes are readily seen under normal conditions, change blindness is easily induced if the changes are made simultaneously with an eye movement, film cut, image flash, or other transient that masks the motion signals normally accompanying the change. (For a review, see Simons & Levin, 1997.)

When discussing change perception, it is important to keep in mind that a clear distinction should be made between motion and change. As used here, *motion* refers to the instantaneous alteration of a quantity. Although it most commonly refers to shifts in location, it can be generalized to other quantities, such as color (see e.g. Adelson & Bergen, 1991). Motion can be specified — under a fairly broad set of conditions[2] — in terms of derivatives, which describe such things as speed and acceleration (see e.g. Nakayama, 1985; Hildreth & Koch, 1987).

In contrast, *change* refers to the situation where a quantity has a particular state at a time t_0, and a different state at time $t_0 + \Delta t$ (where Δt is some finite increment of time). Change may or may not be accompanied by continuous motion. In many real-world situations, the alteration of a quantity occurs continuously, and so can be picked up by low-level motion detectors. However, change can also take place over a temporal discontinuity, such as when a blank field intervenes between an original and an altered display. In cases where motion is picked up by low-level motion detectors, both motion and change perception may take place concurrently. In order to better isolate the various mechanisms involved, change will be operationally defined here as an alteration that cannot be picked up

[2] In general, simple continuity is not enough: additional conditions are generally required to ensure that the derivatives of that particular order and type can exist (see, e.g. Gelfand & Fomin, 1963). However, the initial stages of visual processing generally involve only simple first- and second-order derivatives. These are generally taken after smoothing by low-pass filters, and so can be defined for most conditions encountered in viewing the natural world (see, e.g. Marr, 1982; Hildreth & Koch, 1987).

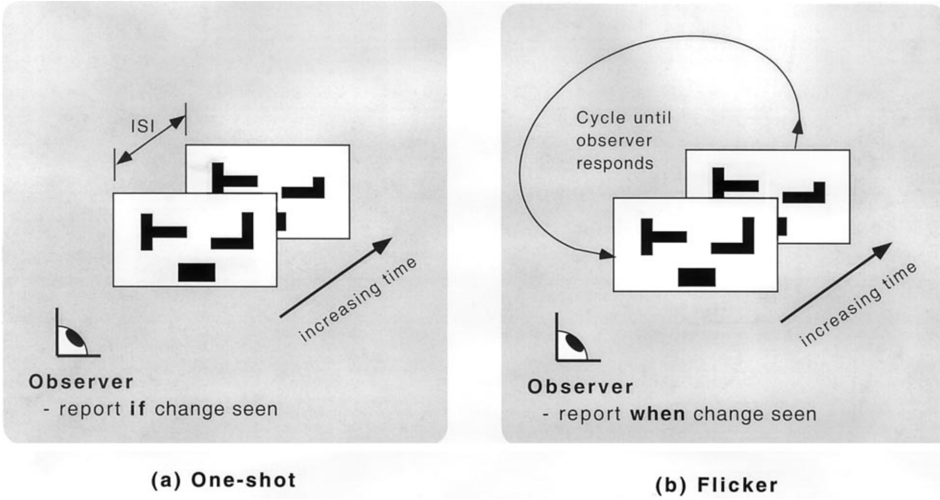

Fig. 1. Experimental paradigms for studying change detection. (a) *One-shot* paradigm. Observer views single alternation of displays, and is asked to determine if a change has occurred. Measure of performance is accuracy of response. (b) *Flicker* paradigm. Observer views continual alternation of displays, and is asked to determine if a change has occurred. Measure of performance is time required for response. Note that these two approaches can also be applied to changes made contingent upon other types of transients, such as eye movements or blinks.

by low-level motion detectors[3]. As such, it is best explored using alterations taking place outside the temporal window of the motion detectors, a window of about 50 ms (e.g. Woodhouse & Barlow, 1982).

2.1. Empirical developments

The study of how visual representations endure over time has its origins in early work on memory recall (Bartlett, 1932). Here, observers were asked to determine a difference between a test figure and a similar figure memorized on some previous occasion. As it turned out, observers could remember the meaning and approximate shape of the figure, but not the visual details. Later studies showed a similar schematic memory for drawings and photographs of real-world scenes (Mandler & Parker, 1976; Intraub, 1997).

This approach, however, was not powerful enough to determine the nature of the perceptual representations involved, for performance could always be explained by limitations of long-term memory. The remedy for this was to switch from the study of *difference* to the study of *change*. Whereas difference involves two instances of a stimulus, change involves only one — a single entity that is perceived to have spatiotemporal continuity. Thus, although memory is still involved, it is a memory tightly bound to perceptual processing[4].

The methodology for studying change stems largely from the work of Phillips (1974)[5]. The basic design is shown in Fig. 1a. Here, an initial display is briefly shown (typically for a few hundred ms), followed by an interstimulus interval (ISI), followed by a second display. The observer must determine if a change occurred in the sequence. These limited-display (or *one-shot*) studies showed that when ISIs were less than 80 ms or so, observers were quite good at detecting a change, presumably because it generated motion transients marking its location. At greater ISIs, these transients were no longer effective, and change blindness appeared, with performance corresponding to a memory for only three to five items in any display. Subsequent work extended these results to a wider range of stimuli and conditions (e.g. Avons & Phillips, 1980; Pashler, 1988), forming the foundations for the study of visual short-term memory (vSTM).

Other studies showed that change blindness does not depend on keeping the eyes fixed — it can also appear if changes are made during a saccade (Bridgeman, Hendry & Stark, 1975; McConkie & Zola, 1979). To account for this, consider the sequence projected upon the retina: first the original display (first fixation), followed by a noninformative blank (or smear created by the saccade), followed by a subsequent display (second fixation). Except for the eye movements, this sequence

[3] Loosely speaking, the difference is (at least in the domain of spatial displacement) that between perceived true motion and apparent motion.

[4] The concept of 'primary memory' (James, 1890) is related to this, in that it involves a requirement of temporal continuity.

[5] An earlier use of this procedure appears in Hochberg (1968). Going back further in time, change blindness was also encountered and exploited by movie editors, who early on found that movie cuts could be rendered largely unnoticed if made during an induced eye movement or blink (Dymtryk, 1984). In addition, conjurers have long been aware of the involvement of attention in detecting change, and have devised numerous strategies to misdirect it, resulting in 'magical' appearances and disappearances of objects (see e.g. Sharpe, 1988).

is the same as that used in the one-shot studies. Indeed, the number of items that can be remembered across a saccade is about three to five, the same as the capacity of vSTM (Irwin, 1991, 1996).

A possible objection is that these limitations apply only to the relatively simple characters and line segments used in those experiments. However, the degree of change blindness is not lessened by use of more realistic stimuli — if anything, it increases (Blackmore, Brelstaff, Nelson & Troscianko, 1995; Grimes, 1996; Simons, 1996). For example, when a prominent building in a city skyline was enlarged by 25% during a saccade, none of the observers noticed the change (Grimes, 1996).

Another possible objection is that the one-shot techniques do not give observers sufficient time to form representations durable enough to support the perception of change. Saccade-contingent techniques make changes after several seconds of viewing, and so escape this objection. But the change blindness encountered in those experiments could potentially be due to some specialized mechanism having to do with saccades, and thus may not accurately reflect what is happening in the rest of the visual system.

To investigate this, a *flicker paradigm* was developed in which the original and the modified images continually alternated until the observer saw the change (Rensink, O'Regan & Clark, 1995; Rensink et al., 1997).[6] This allowed images to remain clearly visible during most of the presentation, the only distraction being a flicker caused by their continual cycling (Fig. 1b). In contrast to earlier approaches, this paradigm gave observers sufficient time to develop a representation typical of normal viewing. Furthermore, if eye movements caused disruption, they could be minimized or even eliminated by the observer. But in spite of these advantages, change blindness was still found: large changes in pictures of real-world scenes could often go unnoticed for up to 50 s, even though the change was made repeatedly, the observer knew a change would be made, and the change was easily seen once noticed.

Other studies show that change blindness can be induced in a variety of ways, such as when changes are made during blinks (O'Regan, Deubel, Clark & Rensink, 2000), movie cuts (Levin & Simons, 1997), or brief 'splats' that do not cover the change itself (Rensink et al., 2000). The fact that change blindness can be induced in a large number of ways — together with the strength and robustness of the basic effect — indicates that it is not an experimental artifact, but rather reflects something central about the way we perceive the world.

[6] The 'one-shot' and 'flicker' techniques are complementary approaches, relying on speed and accuracy measures, respectively. Such complementarity is often used in psychophysical techniques, since each approach has weaknesses that can largely be compensated for by the strengths of the other.

2.2. Coherence theory

Given that change blindness touches on something central to vision, what might this be? One suggestion (Scott-Brown, Baker & Orbach, 2000) is that it might be a limitation in the way that multiple items can be compared. In other words, considerable detail may be accumulated in visual memory, but because of constraints on the comparison process it is simply difficult to make comparisons between the contents of this memory and the contents of the current input.

Such a proposal, however, runs into problems. First, it needs to explain why the detailed contents of successive fixations cannot be added together (Irwin, 1991). It likewise needs to explain why visual search for a changing item is difficult (Rensink, 2000b) — if the details of successive displays could be combined in some way, it would be a simple matter to look for the distinctive pattern formed by the combination of the original and modified items. Finally, this proposal also needs to grapple with the fact that the comparison of items that are already attended requires only about 20 ms/item, suggesting that relatively little of the difficulty in detecting change is due a bottleneck in this aspect of processing (Rensink, 1999a).

A more likely explanation for these results may therefore be that there simply is no visual buffer that accumulates an internal picture of the scene (see Irwin, 1991). Without such a buffer, the detailed contents of successive presentations — including successive fixations — can never be added, compared, or otherwise combined on a large scale.[7] Any change in a scene would then be difficult to notice.

But if we are so bad at combining information on a large scale, why don't we notice this? Why do we believe we can always see (i.e. visually experience) change? Part of the answer is the hypothesis that *focused attention is necessary to see change* (Rensink et al., 1997). Under normal circumstances, a change in the world is accompanied by a motion signal in the input, which attracts attention to its location (see, e.g. Klein, Kingstone & Pontefract, 1992) and so is easily seen. But if this signal appears at the same time as other tran-

[7] It is possible that there may be a spatiotopic buffer in which any representation derived from the new input simply replaces the old. (This is essentially the view taken in the following sections, except with a spatiotopic rather than a retinotopic array.) But a spatiotopic array leads to severe problems. To begin with, factors such as retinal inhomogeneity make it difficult to integrate the contents of successive fixations (Yeshurun & Schwartz, 1989). Furthermore, the computational requirements for representing the contents of the visual field with a detail equal to that of foveal vision are overwhelming (Rojer & Schwartz, 1990); as such, it is unlikely that such a scheme is used in human vision. In any event, the main issue at this point in the discussion is not completeness, but the difficulty of combining information from successive displays.

sients (associated with a saccade, flicker, eyeblink, splat, etc.), it is no longer unique, and so cannot draw attention. Change blindness is then induced.

This explanation, however, creates a problem. Attention is thought to 'weld' visual features into long-lasting representations of objects (Kahneman, Treisman & Gibbs, 1992), and to operate at a rate of about 30 items/s (Julesz, 1984; Wolfe, 1994). Why then doesn't it weld all the visible items within the first few seconds of viewing and thereby enable the easy detection of change?

The answer to this goes to the very heart of what it means to be attended. Rather than assuming that the structures formed by attention last indefinitely, their lifetimes may be much more limited. In particular, it may be that attention endows a structure with a coherence[8] that lasts only as long as attention is directed towards it. A more precise specification of this *coherence theory* of attention (Rensink et al., 1997; Rensink, 2000a) is as follows (Fig. 2):

1. Prior to focused attention is a stage of *early* processing, i.e. processing that is low-level (involving only the geometric and photometric properties), rapid (occurring within a few hundred ms), and carried out in parallel across the visual field. The resultant structures (*proto-objects*) provide local descriptions of scene structure, such as three-dimensional orientation, or groupings of related edge fragments (Rensink & Enns, 1995; also see Wolfe, 1995 for a somewhat similar idea.) Proto-objects can be quite complex, but are coherent only over a small region of space (Rensink & Enns, 1998). They likewise have limited temporal coherence — they are volatile, being constantly regenerated, and so are simply replaced by any new stimuli appearing in their retinal location. (Note that such replacement rules out their superposition.)[9]
2. Focused attention acts as a hand that grasps a small number of proto-objects from this constantly-regenerating flux. While held, they form a *coherence field* representing an individuated object with a high degree of coherence over time and space.[10] This field is formed via feedback between the proto-objects and a mid-level *nexus*, a locus where lower-level information is collected and used as the basis of higher-level decisions. This coherence allows the object to retain an identity across brief temporal interruptions; as such, it is transformed rather than replaced by new stimuli arriving at its location.
3. After focused attention is released, the object loses its coherence and dissolves back into its constituent set of proto-objects. Such a position is consistent with results indicating a lack of attentional aftereffect in visual search (see also Wolfe, 1999). Note that in this view, there is little — if any — short-term visual memory apart from what is being attended. This suggests that attentional hold may be largely identified with vSTM.

According to coherence theory, a change in a stimulus can be seen only if it is attended at the time the change occurs. Since only a small number of items can be given focused attention at any time (see e.g. Pashler, 1988; Pylyshyn & Storm, 1988), most of the image will not have a coherent representation. Thus, if attention is not

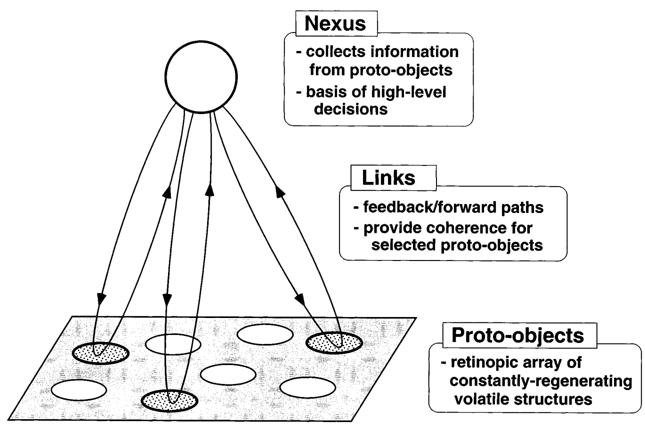

Fig. 2. Schematic of coherence field (Rensink, 2000a). In the absence of focused attention, low-level structures (*proto-objects*) are volatile. Attention acts by establishing feedback *links* between proto-objects and a mid-level *nexus*. The set of interacting proto-objects, links, and nexus is a *coherence field*. The interaction among the various parts of the field allows the establishment of coherence in attended proto-object properties, both in space and in time.

[8] In this paper, 'coherence' refers to the conjunction of consistency and logical interconnection, i.e. agreement that the structures refer to parts of the same system. Like 'stability', this term is used in two ways. In its spatial aspect, it denotes a set of representations at different locations that refer to the same object; these representations are essentially grouped together, with any interpretation assigned to one representation influencing the others. In its temporal aspect it likewise denotes a set of representations at different times that refer to the same object (see Rensink, 2000a). Note that spatial stability does not imply spatial coherence: the statistics of local features in a scene, for example, could be largely invariant over eye movements, even though never given a coherent representation. Conversely, spatial coherence does not imply spatial stability: a coherent retinotopic representation could be rebuilt with each eye movement.

[9] It is possible for the fusion of two successive stimuli to occur preattentively, presumably due to summation of their signals by filters operating at the initial stages of visual processing. However, such fusion requires that the onsets of the two presentations be within 100 ms of each other (e.g. DiLollo, 1980), a condition that does not generally occur here.

[10] This is somewhat similar to the idea of object files (Kahneman et al., 1992). However, the emphasis here is on coherence in time as well as space. Furthermore, coherence theory states that the representation of the object (i.e. the coherence field) is in existence only as long as attention is being directed to the corresponding items in the display — there is little or no memory of object structure once attention has been withdrawn. This contrasts sharply with object files, which can be accumulated over time.

drawn to the location of a change, the change will likely not be seen.

Coherence theory also explains results from several related lines of research. For example, the volatility of the proto-objects is sufficient to account for finding that — apart from the limited amount held in transsaccadic memory — the details of successive fixations cannot be combined (see Irwin, 1996). Furthermore, if transsaccadic memory is identified with vSTM (Irwin, 1991) and vSTM with the attentional grasp, it explains why only a few saccade-contingent changes are perceived at any one time. And given the close connection between attention and saccade programming, it explains why detection of change is best for objects that are saccade targets (Irwin, 1996).

The failure to notice saccade-contingent displacements in attended targets (e.g. Bridgeman et al., 1975) might appear to argue against coherence theory. However, these experiments generally ensure that no background markings are visible. If the perceptual system never codes position in absolute terms — relying instead on relative locations — this lack of background information would then cause a failure to perceive change in location. Note that a change in the relative location of a set of attended items *is* easy to perceive, provided the number of items is small (Lachter & Hayhoe, 1995).

It is also worth pointing out that coherence theory helps elucidate the nature of mid-level processing[11] — in particular, the processing required to form a coherent object representation. Previous models (e.g. the 'object files' of Kahneman et al., 1992) tended to echo the earlier belief that vision accumulated detailed information over time — once attention formed an object file, it continued to exist for a time, even if unattended (Kahneman et al., 1992). In contrast, coherence theory posits that coherent representations (i.e. coherence fields) collapse as soon as attention is withdrawn. And whereas earlier models such as object files embodied the view that attention is *sufficient* to see change, coherence theory makes the explicit claim that it is *necessary*. (For a more detailed comparison of these theories, see Rensink, 2000a.)

2.3. Inattentional blindness

Coherence theory asserts that without attention, there is no visual experience of change. But what about the perception of static stimuli? Several studies (e.g. Neisser & Becklen, 1975; Becklen & Cervone, 1983; Mack & Rock, 1998) have found that observers attending to a particular object or event often fail to report the appearance of irrelevant, unexpected items. This has been interpreted as *inattentional blindness*, a failure to represent unattended items (Mack & Rock, 1998). Is this the same thing as change blindness? If not, what is the relation between the two phenomena?

To begin with, it is important to keep in mind that the phenomenon of change blindness shows only that we are blind to changes made in the scene — it says nothing about whether we are also blind to the scene itself. Coherence theory goes a step further, arguing that unattended representations are not coherent enough to support the perception of change. But it takes no position on whether any visual experience can be supported by unattended representations.

Conversely, the phenomenon of inattentional blindness concerns itself with perception of the scene itself — it says nothing about the perception of change. It might be, for example, that we could see attended items but not be able to see them change. Thus, at a purely descriptive level inattentional blindness and change blindness are largely separate phenomena: one pertains to the static aspects of the scene, the other to its dynamic aspects.

A relationship does emerge, however, when the mechanisms explaining the two phenomena are examined. Consider the status of unexpected, unattended stimuli in an inattentional-blindness experiment. It has been shown that these items provide a context that can influence the perception of attended structures (Moore & Egeth, 1997). This suggests that representations of considerable detail and sophistication are present, even though they are not reported. Similarly, coherence theory posits that low-level representations with some degree of detail — proto-objects — are continually generated in the absence of focused attention (Section 2.2). As such, explanations for both phenomena rely on an assumption that unattended stimuli can give rise to fairly sophisticated representations at early levels.

But what is that role played by these early-level representations? It has been suggested that unattended stimuli are indeed *seen*, but that in the absence of attention are simply not *remembered*; consequently, inattentional blindness should be renamed *inattentional amnesia* (Wolfe, 1999). Support for this position comes from studies of iconic memory, which indicate that much more can be scanned in a briefly-presented display than can be directly reported (Sperling, 1960). A related finding is that change blindness can be substan-

[11] There is at the moment some non-uniformity in terminology. Some authors (e.g. Ullman, 1996) use 'high-level' to refer to object perception, leaving 'mid-level' to refer to the perception of relatively unstructured surface properties. Other authors (e.g. Henderson & Hollingworth, 1999) use 'high-level' to refer to the perception of meaning (i.e. identifying an object or a scene *as* something), with 'mid-level' vision referring to the formation of coherent, structured representations such as objects. This latter usage is the one followed here. Fortunately, the underlying conceptual structure is much the same in both cases — it is only a question of what label is assigned to what.

tially reduced if a verbal cue is presented that describes (in one or two words) the item that is changing (Rensink et al., 1997). These results suggest that there may exist some kind of scanning process that acts faster than the process involved in the perception of change. But although unattended items appear to support such scanning, and so are seen in some sense, it is not clear that this scanning process involves conscious visual experience.

In summary, then, change blindness and inattentional blindness (or inattentional amnesia) pertain to different aspects of the world: change blindness to its dynamic aspects, and inattentional blindness to its static ones. As such, they are different phenomena. The issue of visual experience — whether we experience unattended stimuli and then forget, or whether we never experience them at all — remains unresolved. But even though little can be said about the particular issue of visual experience, much *can* be said about many other issues concerning how we see.

3. Seeing

What does it mean, to see? The plain man's answer (and Aristotle's too) would be, to know what is where by looking.

Marr (1982)

Coherence theory can explain why change blindness occurs under a wide variety of conditions. But more importantly, the consequences of this theory spill over into several other aspects of vision, where they force reconsideration of many basic issues.

For example, consider the general question of how we see[12]. Coherence theory posits that there is no visual buffer containing a complete description of the visible scene. Even worse, it posits that only a small number of objects have a coherent representation at any one time. If this is so, why does it seem that we simultaneously experience all the objects in our surroundings? And if we do not build up a detailed description of a scene, how does our visual system know what is where by looking?

3.1. Virtual representation

To reconcile the large number of objects we experience with the much smaller number that can simultaneously be given a coherent representation, an appeal can be made to the idea of a *virtual representation*: instead of simultaneously representing in detail all the objects in our surroundings, represent only those objects — and only those particular properties of those objects — needed for the task at hand (Rensink, 2000a). If attention can form a coherent representation of an object whenever requested, the representation of a scene will appear to higher levels as if it is 'real', i.e. as if all objects simultaneously have a coherent representation.

At its heart, the use of virtual representation is an information-processing strategy that reduces complexity in space by increasing complexity in time. The factors critical to the success of such a strategy are:

1. only a few objects need to have a coherent representation at any instant.
2. detailed information about any object must be made available whenever requested.

The first requirement is easily met for most (if not all) visual tasks, since there is usually only one object in play at any one time; tasks involving more than a few objects can be handled by rapidly switching attention between the objects (Ballard, 1991; Rensink, 2000a). The second requirement is also met under most conditions: given that there is sufficient information to guide attentional shifts and eye movements to the location of a requested object, detailed information can always be obtained from the incoming light and entered into a stable representation. A high-capacity memory visual buffer is therefore not needed — the information can be obtained from the world itself. As pointed out long ago by Stroud (1955):

Since our illumination is typically continuous sunlight and most of the scenery stays put, the physical object can serve as its own short-term memory…

Virtual representation is a central element of computer and network design (e.g. Tanenbaum, 1976). Although this form of representation was long neglected as a possibility for biological systems, Stroud's insight has recently received increased recognition, and virtual representation is now proposed as a way to explain several aspects of perception and cognition (Brooks, 1991; Dennett, 1991; O'Regan, 1992; Grimes, 1996). Indeed, it is becoming apparent that the environment is a critical part of many perceptual processes, capable of acting not only as an external *memory* but also as an external *processor* (see Clark, 1997).

But how might such a dynamic representation be implemented in the human visual system? One possible

[12] As used here, 'seeing' is defined in a purely behavioral way, referring to the use of light to carry out various tasks (Appendix A). This section is therefore primarily concerned with the problem of how information is managed. There is some discussion of visual experience, viz., how observers get the impression that many objects are represented simultaneously. However, given that attention appears to be sufficient for visual experience (Merikle & Joordens, 1997; Braun & Julesz, 1998), that discussion pertains mostly to the way that attention can be managed so that it is effectively allocated to many objects simultaneously.

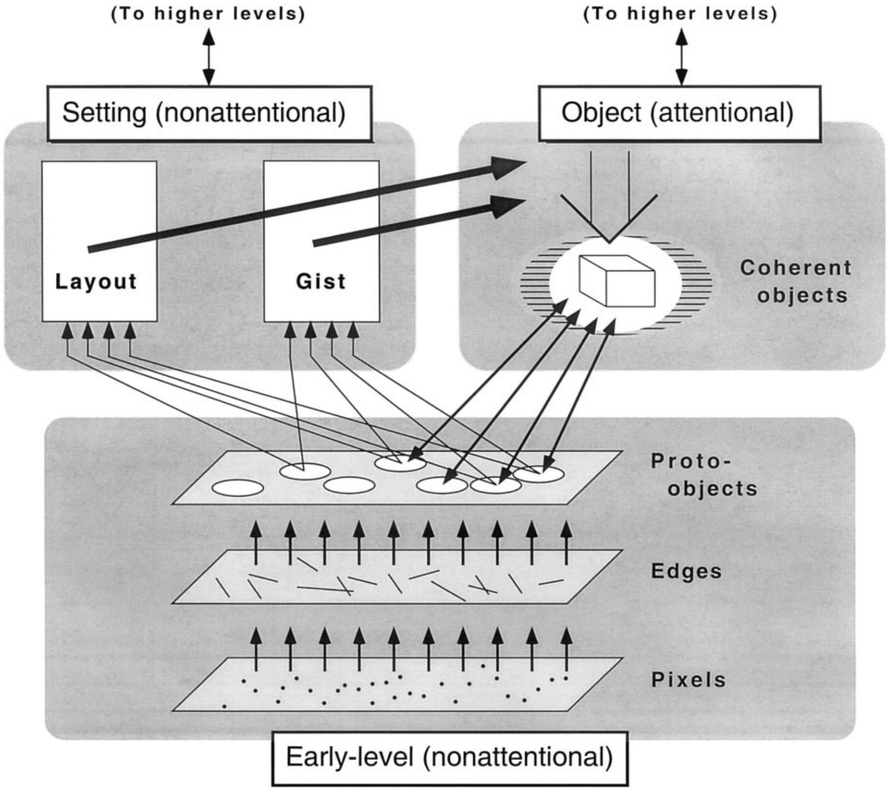

Fig. 3. Schematic of triadic architecture (Rensink, 2000a). In this architecture, visual processing is split into three largely independent systems: (i) early vision — a high-capacity system concerned with the rapid formation of sophisticated (although volatile) proto-objects, (ii) attentional vision — a low-capacity system concerned with the formation of coherent objects, and (iii) nonattentional vision — a low-capacity system concerned with establishing a context to guide focused attention. Note that according to coherence theory, the attentional system is needed for the visual experience of change; nonattentional systems provide no such experience. As such, the unattended structures at early levels do not support the visual experience of change. However, their contribution to the visual experience of static stimuli is unknown.

solution (Rensink, 2000a) is a *triadic architecture* of the type shown in Fig. 3. This architecture comprises three largely independent systems. The first is a low-level system that rapidly creates highly-detailed, volatile structures. The second is a limited-capacity attentional system that forms these structures into coherent object representations. These two systems are already part of coherence theory. What is now added is a limited-capacity nonattentional system that provides a setting to guide attention. This *setting system* involves at least two stable aspects of scene structure:

1. The abstract meaning (or *gist)* of the scene, e.g. whether the scene is a harbor, city, picnic, barnyard, etc. Gist appears to be extracted without attention (Oliva & Schyns, 1997), possibly on the basis of the statistics of low-level structures such as proto-objects (e.g. Swain & Ballard, 1991). Since gist is largely invariant with eye position, it can provide a stable context that would constrain the kinds of objects expected, and perhaps indicate their importance for the task at hand (Friedman, 1979).
2. The spatial arrangement (or *layout)* of the objects in the scene. This appears to be described via a stable representation of the location of each item (Simons, 1996; Sanocki & Epstein, 1997); a few of the properties of each item could also be included. Some such representation is vital if the (limited) information obtained from individual eye fixations is to be integrated into a structure capable of directing subsequent eye movements and attentional shifts (Hochberg, 1968).

The central assumption here is that the information in this system is sufficient to direct attention and the eyes to whatever object is required, which then allows the attentional stream to temporarily form a coherent representation of the object of interest. If this *attentional management* is adequate, knowledge of the detailed structure and location of any object will then be readily available whenever requested.

3.2. Dependence on task

Given a dynamic representation of the type sketched above, it is clear that attention should be managed so that the limited amount of information stabilized is used as effectively as possible. Such management is presumably done via the use of strategies that depend on the nature of the task.

Early studies on picture perception (Yarbus, 1967) showed a clear influence of task type on eye movement patterns; given the close association between eye movements and attentional shifts (e.g. Henderson, 1996), they also provide an indirect demonstration that task type can influence the pattern of attentional shifts. Change-blindness studies, however, can provide a more direct means of determining how the allocation of attention depends on the task at hand.

An example of this is the finding (Hayhoe, Bensinger & Ballard, 1998) that detection of change depends on the demands of the task at the moment the change is made. Here, observers were asked to form a copy of a pattern of blocks on a computer monitor; this was to be done by using a mouse to drag blocks from a set of resource blocks elsewhere. During the course of construction, a block in the target pattern would sometimes change color when an eye movement was made toward it. Results showed that the ability to detect this change — measured by the length of time the block was fixated — depended on the particular point in the task sequence. When the change was made before the start of a resource-block move (presumably to obtain the color of the pattern block), detection of the change was poor; when the change was made after the move (when the pattern block was again fixated, presumably to check its color), detection of change was much better. As such, these results show that attending to the item served only for the immediate needs of the task at hand, and not for the construction of some general-purpose representation.

Change-blindness experiments also provide evidence that the nature of the task not only affects *when* the item is attended, but also the particular set of *properties* selected. An example of this is a study which looked at search for change in an array of simple figures (Rensink, 1999b). When observers searched for a change in *orientation*, search speed was considerably influenced by the shape of the items, being almost twice as fast for simple lines as for compound figures made of two or more lines. However, when the same figures were used in a search for change in *contrast polarity*, no significant effect of shape was found. Evidently, at least some aspects of geometric information did not enter into the representations formed when only nongeometric information needed to be stabilized.

Similar effects have also been found using more naturalistic stimuli (Levin & Simons, 1997). Here, observers watched a film sequence in which the main character — presumably the focus of attention — changed during a movie cut. Almost all changes went unnoticed. When observers were then asked to watch for changes in a second presentation of the film, performance improved considerably (although most changes were still missed). This again suggests that the only aspects of a stimulus put into coherent form are those that serve the immediate functional needs of the observer (Rensink, 2000a; Simons & Levin, 1998).[13]

3.3. Dependence on high-level knowledge

Another consequence of a dynamic representation is that the high-level knowledge of an observer greatly affects how attention is managed, which in turn affects how they see. Knowledge can have at least two effects in this regard:

1. a 'chunking' of data into a unitary structure that can be attended and held in visual memory in much the same way as a simpler structure (e.g. Chase & Simon, 1973).
2. a more effective guidance of attention, via the knowledge of which aspects are important as well as anticipation of what might happen next (e.g. Neisser, 1976; Becklen & Cervone, 1983).

Given that attention is needed to see change, the perception of a change in a realistic situation can be a powerful tool to map out both of these factors. This in turn can help map out what an observer knows, and thus how they see their world.

As an example of using change perception to investigate general knowledge, Rensink et al. (1997) showed that changes in real-world scenes were noticed most quickly for *central interests*, i.e. objects or regions that were most often mentioned in brief verbal descriptions of the scene. Evidently, attention is sent to these before being sent to less interesting (and presumably less important) parts of the scene. Although part of the reason that the central interests are attended relatively quickly might be the salience of their visual features (e.g. color or contrast), experiments on line drawings show that higher-level factors such as semantic consistency also play a part (Hollingworth & Henderson, 2000).

It has also been shown that knowledge can affect the perceptual representations themselves. One study (Archambault, O'Donnell & Schyns, 1999) had observers learn a set of objects either at a general level of categorization (e.g. 'a mug'), or at a specific level (e.g. 'the mug owned by Bill'). Changes were then made to these objects in images of real-world scenes. As before, detection of change was poor when the changes were made during an interruption. However, performance was better when the objects being changed had been learned at a specific level than at a general level,

[13] Care must be taken in distinguishing between 'object' defined as a structure in the external world (i.e. a concrete spatiotemporal entity) and as a structure internal to the observer (i.e. a coherence field). If 'object' is taken to be an external structure, then attention to it will encode only some of its properties. As such, attention is necessary but not sufficient to perceive change in an object. If 'object' is defined as the contents of the coherence field, however, then attention is both necessary and sufficient to form the field, and thus, to perceive change in it.

indicating that the particular type of learning influenced the perceptual representations of the stimulus that were formed.

In another study, Simons and Levin (1998) had an experimenter ask people on the street for directions. A few seconds after conversation began, workers carrying a large board passed between the (unwitting) subject and the experimenter, permitting the experimenter to be switched with another person. Observers were much more likely to notice the change when the experimenter was considered to be a member of their social class, indicating a high-level influence on the representations used during the interaction.

Change perception has also been used to explore the effects of specialized expertise. Werner and Thies (2000) examined the ability of observers to notice changes in scenes of American football games. Comparing the performance of experts against that of non-experts, they found that experts were quicker to spot changes, regardless of whether or not these changes affected the meaning of the scene. This is consistent with the proposal that experts code the scene in a way that allows them to spot changes in meaning relatively quickly, and that their familiarity with the domain also allows them to attentionally scan meaningful scenes more efficiently.

4. Sensing

We begin not with a sensory stimulus... In other words, the real beginning [*of seeing*] *is… not a sensation of light*.

Dewey (1896)

Often implicit in the definition of seeing is the requirement that the 'knowing by looking' be accompanied by a sensory experience — more precisely, by a conscious visual experience. But it is not necessary that all aspects of visual processing must lead to visual experience. Indeed, a large body of experimental evidence (e.g. Schacter, 1987; Milner & Goodale, 1995; Weiskrantz, 1996; Merikle & Joordens, 1997) points towards the existence of *sensing*, i.e. the processing of visual information without accompaniment by visual experience, or by any conscious awareness at all. (For definition of terms, see Appendix A.)

Historically, investigation of sensing has proceeded rather slowly (see e.g. Merikle & Reingold, 1992). Much of this has been due to the problem of producing stimuli that are not visually experienced but yet still affect the observer in some way. This has most often been accomplished by weakening stimuli (e.g. by brief exposure or added noise) to the point where they are below the threshold of conscious perception. However, such stimuli have also proven to exert only a weak effect on other visual systems, making empirical investigation difficult.

In contrast, it is relatively easy to generate large suprathreshold changes that are — until attended — effectively invisible. Given the magnitude of the changes that can be made, and the duration over which change blindness can last, it would appear that there is a high likelihood of engaging other (nonattentional) systems. As such, change perception may have considerable potential for the empirical study of those aspects of visual processing not accompanied by visual experience.

4.1. Visuomotor response

Evidence is accumulating (e.g. Goodale, 1993; Milner & Goodale, 1995) that vision may be composed of two largely independent systems: an *on-line* stream concerned with immediate visuomotor action (e.g. movement of the eyes, maintenance of balance), and an *off-line* stream concerned with more time-demanding processes (e.g. the conscious perception and recognition of objects in the immediate surround). When the off-line stream is damaged (e.g. by a lesion), an observer may not have a conscious experience of part of their environment, e.g. an object in some part of their visual field. However, if their on-line stream still functions, they will still be able to interact with it in various ways, such as grasping or pointing to it. This phenomenon is known as *blindsight* (Weiskrantz, Warrington, Sanders & Marshall, 1974; Weiskrantz, 1996).[14]

A similar separation of visuomotor response and conscious perception has been found in normal observers, reflecting the predominance of the visuomotor system under certain conditions (Goodale, 1993). Much of the evidence for this has been based on change-detection experiments. For example, if the location of a rod is changed the moment the observer attempts to grasp it, an adjustment to the trajectory of the hand will take place several hundred milliseconds before conscious report of the change (Castiello, Paulignan & Jeannerod, 1991). Motor adjustment — although taking about a hundred ms longer — also precedes conscious report when the size of the rod is changed (Castiello & Jeannerod, 1991). In both cases, then, there existed a brief period of time during which the visuomotor system responded to the change, even though there was no conscious perception of it.

[14] The term 'blindsight' also extends to those situations where a particular aspect of conscious perception is affected, such as loss of color vision (see e.g. Weiskrantz, 1996; Stoerig & Cowey, 1997). However, consideration in this section is limited to the case where the visuomotor system responds to stimuli that are not consciously perceived. The case of responding to unseen perceptual quantities is essentially a form of implicit perception, which is considered in the next section.

A similar dissociation of perception and oculomotor response also exists: even when a displacement in location made during a saccade is not consciously perceived, the eye still makes a corrective saccade to the new location (Bridgeman et al., 1975). Other studies (Bridgeman, Lewis, Heit & Nagle, 1979; Goodale, Pelisson & Prablanc, 1986) found this to be true of manual pointing as well — if a target was displaced during a saccade, the hand of an observer reaching toward it would correct its trajectory, even though the observer never consciously noticed the displacement.

Likewise, Hayhoe et al. (1998) found that a saccade-contingent change to the color of a block in a display caused the eye to fixate on it longer than usual, even though observers had little conscious experience of the change. Interestingly, this effect increased in strength when other, irrelevant blocks changed color at the same time, indicating that some more global factor was also at play. In any event, these findings indicate that various visuomotor systems are able to support the detection of changes to the targets of their actions, even though these may not be attended (or at least attended in a way that enables the conscious perception of the change).

4.2. Implicit perception

The studies described above show that the occurrence of change can affect on-line action without being accompanied by any visual experience. But what about the perceptual stream itself, where the processing of information is presumably for off-line control?

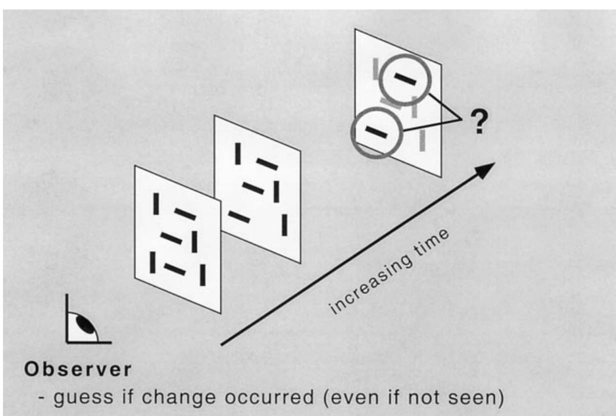

Fig. 4. Perception of change without awareness (Fernandez-Duque & Thornton, 2000). Here, observers view a one-shot display sequence, with displays composed of arrays of rectangles. They are then shown a test display containing a highlighted subset of the original display: the item that changed, and the item diagonally across from it. Observers are asked which of the two items changed. If they did not notice a change in the one-shot sequence, they are asked to guess. For cases where observers had no awareness of change, performance was still significantly above chance, indicating some degree of implicit perception.

An interesting phenomenon in this regard is *implicit perception*. When a stimulus is masked almost immediately after being presented, it can cause priming (i.e. an increased sensitivity to the subsequent occurrence of the same stimulus), even though the observer has no subjective experience of it (see e.g. Marcel, 1983). It has been suggested that such implicit perception is exactly perception without attention (Merikle & Joordens, 1997).

Fernandez-Duque and Thornton (2000) examined whether change could be detected implicitly. They presented two brief displays, the first a simple array of rectangles, the second a similar array in which one of the rectangles changed its orientation. These displays were followed by a test display highlighting two of the display items — the changed item, and the item diametrically opposite it in the display (Fig. 4). Observers were asked to guess — regardless of whether they had seen the change — which of the two highlighted test items had changed. Results indicated that when observers did not visually experience the change, they were able to correctly guess the correct test rectangle more often than not. The level of performance was not high, ranging from 55% when the display contained 16 items to 63% when it contained eight. However, in all cases performance was significantly above chance.

Importantly, control experiments showed that observers were faster to discriminate an item appearing at the location of a change, but only when the observer was aware of the change. The finding that explicit perception is accompanied by (presumably attentional) priming is consistent with the assertion of coherence theory that attention is needed to visually experience change (Section 2.1). Meanwhile, the finding that implicit perception is not accompanied by such localized priming is in accord with the proposal of Merikle and Joordens (1997) that implicit perception never involves attention. As such, this result supports the idea of a clear separation of processing: the explicit perception of change is mediated by focused attention, whereas the implicit perception of change is not.

In another set of studies, Thornton and Fernandez-Duque (1999) showed that non-localized priming could be found for changes that were not explicitly perceived. For example, when an item changed from horizontal to vertical, a subsequent horizontal probe tended to be perceived as vertical, but only if the observer had no awareness that a change had taken place. Response times also tended to slow down, but only if the observer *did* have an awareness of the change. Interestingly, this form of priming does not require the probe to be at the location of the change, indicating that the system involved is concerned with the general *type* of change.

Taken together, these results suggest that the implicit perception of change involves a mechanism operating independently of focused attention. Given that perfor-

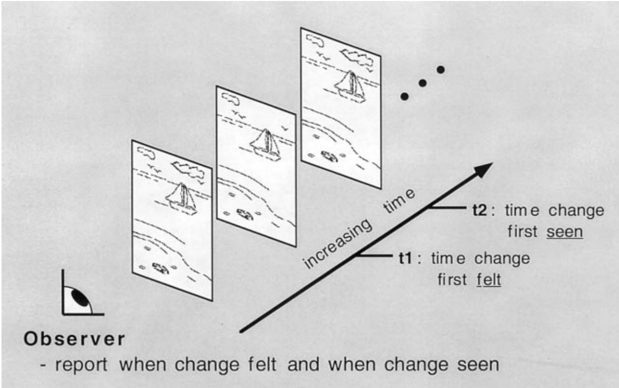

Fig. 5. Mindsight (Rensink, 1998a). Observers view a flicker display sequence containing a change in a real-world scene. Observers are asked to press a button twice: first when they *feel* a change occurring (t_1), and then when they *consciously see* it (t_2). For many observers, t_2 exceeds t_1 by more than one second on roughly a quarter of all trials. This phenomenon of 'feeling of seeing' with no accompanying visual experience is termed *mindsight*.

mance is better when fewer items are in the display, this mechanism is likely to have a limited capacity. A possible candidate in this regard may be one of the nonattentional streams posited for the triadic architecture sketched above (Section 3.1). Since these are limited-capacity and nonvolatile, such a stream could support the implicit detection of a limited number of changes at any one time.

4.3. Mindsight

Can an observer consciously sense that a change is occurring but still have no visual experience of it? In one study (Rensink, 1998a), 40 observers were presented with a flicker sequence in which an image of a real-world scene alternated with a similar image altered in some way (Fig. 5). Observers pressed a button first (t_1) when they were *aware* that something was changing, and then again (t_2) when they *visually experienced* the change, i.e. when they believed they could describe the item that was changing. Catch trials (where no change occurred) made up 12.5% of the presentations, allowing an estimate of what happened when observers simply guessed that a change was occurring.

The results of this study were clear: even when a conservative criterion was used ($t_2 - t_1 \geq 1$ s), sensing was found to occur in a large number of trials. Interestingly, most of these were from a small subset (14/40) of observers; half the observers (20/40) experienced little or no sensing. In both groups, guessing occurred on about 15% of the catch trials. However, invalid t_1 responses occurred several seconds later than valid t_1 responses, indicating that whenever an observer simply guessed, there was a significant hesitation in their response. Only in a small subset (6/40) of observers was there no significant difference between valid and invalid t_1 responses, indicating a strategy of pure guessing unaffected by any sensing of change.

Evidently, then, a considerable fraction of observers can have an abstract mental experience without sensory experience. It is worth mentioning that this phenomenon — *mindsight* (Rensink, 1998a) — may allow the exploration of visually-mediated processes that are difficult to study using more conventional techniques. An interesting possibility in this regard is that mindsight may correspond to the 'sixth sense' sometimes believed to provide a warning to an observer in a dangerous situation.

The mechanisms underlying mindsight are not well understood. Trials where $t_2 - t_1 \geq 1$ s (the criterion for mindsight) did not have smaller t_1 values — i.e. feelings were not faster than visual experience. Rather, t_2 values were larger, indicating that there was a delay in experiencing the change. This delay was limited to only about 25% of the trials, showing that mindsight is not due to a general slowdown of attentional processing. Rather, it appears to be due to a nonattentional mechanism[15], possibly an alert of some sort. The reason for the delay in visual experience is unclear — perhaps the observer has difficulty disengaging from the nonattentional mechanism, thereby slowing down the processes underlying the visual experience of change. However, given that only a small amount data has been gathered, such an explanation is necessarily speculative. Future investigations will hopefully cast more light on the mechanisms underlying this phenomenon.

5. Scrutinizing

For many reasons, the question that once appeared fundamental and clear-cut, about the locus of attentional selection, now looks confusingly ill-defined or (at best) fragmented into a variety of subquestions.

Allport (1992)

Just as change-blindness experiments can shed light on sensing, so can they shed light on other, more familiar aspects of visual perception. One aspect of particular relevance in this regard is *scrutinizing*, i.e. the application of focused visual attention. According to coherence theory (Section 2.2), focused attention is necessary for the perception of change. If this view is correct, not only can focused attention be used to

[15] Such an explanation would be consistent with the finding of Bonnel and Hafter (1998) that detection of luminance change does not have a capacity limit, and so presumably does not require attention.

explain change blindness, but change blindness can be used to investigate the nature of focused attention itself.

Much has been learned of attention using other approaches, such as visual search on static displays, and priming (see e.g. Pashler, 1998). However, the perception of change involves a more dynamic aspect of visual perception than these, and so is likely to be a largely independent source of information. Furthermore, the magnitude of change-blindness effects is usually quite large. From a methodological point of view, then, change blindness has the potential to yield experimental results with a high signal-to-noise ratio. And this in turn may allow us to explore the mechanisms of focused attention in great detail.

5.1. General methodology

A powerful way to investigate the mechanisms of focused attention is by using one-shot or flicker paradigms (Section 2.1) with arrays of simple figures (Phillips, 1974; Pashler, 1988; Rensink, 2000b). To see how this works, consider an array of rectangles, where one item (the target) changes its orientation on half the trials while the other items (the distractors) do not; the observer must then report for each trial whether or not a change was present (Fig. 6). If roughly half the items are horizontal and half vertical in both images, the target cannot be detected from any single display — both displays must be compared. If the interstimulus interval (ISI) between displays is sufficiently long, the transients due to the changing target will be swamped by the transients produced by the flickering distractors. Detecting the target will then require an attentional scan of the items in the display.

This approach extends the 'classical' visual search paradigm on static displays (e.g. Treisman & Gormican, 1988) into the temporal realm. All the power of the static techniques (e.g. using different shapes, different features, search asymmetries) is retained; in addition, two more degrees of freedom can be manipulated: display time (*on-time*) and ISI (*off-time*). Note that the one-shot and flicker techniques are natural extensions of the two different ways that have historically been used to measure performance on static displays, viz., accuracy on briefly-presented displays (e.g. Sagi & Julesz, 1985), and reaction time on displays that are shown until the observer responds (e.g. Treisman & Gormican, 1988).

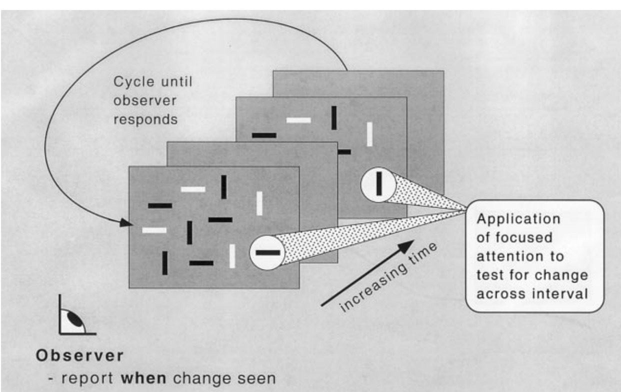

Fig. 6. Flicker paradigm with controlled displays. Displays are arrays of rectangles. In half the displays, one item (the target) changes orientation while the other items (the distractors) remain constant. In the other half of the displays, all items remain constant. According to coherence theory, the only way for an observer to determine if a change is occurring is to carry out an attentional search of all items in the display.

Visual search for change allows various aspects of attentional processing to be investigated. These include:

5.1.1. Speed

As in the case of static visual search, this can be determined by examining how performance varies as a function of the number of items in the display; for the flicker variant, reaction time often increases linearly with the number of items, indicating a constant search speed (Rensink, 2000b). In contrast to static search, the geometric properties of the target do not have to be specified in advance — it can be defined simply as the item that changes. Furthermore, the influence of salience can often be eliminated by proper choice of items in the display, such as when half the items have one value of a property (e.g. vertical) and the other half the other value (e.g. horizontal).

5.1.2. Capacity

This can be determined by measuring performance as a function of on-time (i.e. the display time) — as on-time increases, more items can be 'grabbed' by attention, until saturation is reached (Pashler, 1988; Rensink, 2000b). The value of this asymptote has been interpreted as a measure of vSTM; however, given the close identification of vSTM with attention posited by coherence theory (Section 2.2), it can also be taken as a measure of attentional capacity. Note that this estimate is an upper bound on the number of attentional links involved, since grouping factors may lead to chunking, causing more than one stimulus item to be assigned to each link.

5.1.3. Selectivity

This can be determined by comparing the speed when all items must be examined against the speed when the change occurs in a selected subset. For example, the speed for orientation change can be measured for a set of black and white items; it can then be measured for the same items, but with the change occurring only in the black ones. The ratio of these two speeds is a measure of the selectivity for black (Rensink, 1998b).

5.1.4. Task

Different aspects of change perception can be explored using different types of task, such as detection (reporting if there is a change in the display), identification (reporting *what* the change is), and localization (reporting *where* is it). Although it might be thought that these tasks should lead to similar estimates of attentional ability, such is not the case — for example, capacity estimates for identification are always below those for detection (Wilken, Mattingley, Korb, Webster & Conway, 1999).

Furthermore, the stimuli in the experiments can have different levels of complexity. This allows experiments to progress in a smooth way from simple tasks on highly-controlled arrays to more natural tasks on images of complex, real-world scenes.

5.2. Empirical results

Relatively few experiments have yet been carried out on the perception of change under controlled conditions. However, the results already obtained show some interesting patterns. For example, flicker experiments indicate that as long as memory limitations do not intrude, search for the *presence* of change takes place at approximately 100 ms/item, the exact speed depending on the shape of the items (Rensink, 1999b) and particular property being changed (Rensink, 2000b). Meanwhile, search for the *absence* of change (i.e. search for a constant target among a set of changing distractors) takes place much more slowly — typically, about 300–400 ms/item (Rensink, 1999a). And search for a *conjunction* of change (i.e. search for a target changing both dimensions among a set of distractors that each change in one) also takes place much more slowly — again, about 300–400 ms/item (Rensink, 1999a).

Interestingly, a similar pattern is obtained for estimates of attentional capacity. Search for a *presence* of change leads to estimates of about four to six items, i.e. only about four to six items can be attended at a time (Pashler, 1988; Rensink, 2000b). This number is similar to the capacity of vSTM, and supports the conjecture that vSTM is largely (if not entirely) identical to what can be held by the attentional 'hand' (Rensink, 2000a). However, search for an *absence* of change leads to a much lower number — a value of only 1.4 items (Rensink, 1999a). And search for a *conjunction* of change leads to an estimate of only 1.0 items (Rensink, 1999a). Note that this pattern is not only the same as that for the speed estimates — it also has a striking similarity to the pattern found in static visual search, where presence of a distinct feature can be detected quite quickly, while detecting an absence or a conjunction of features is slower and more effortful (Treisman & Gormican, 1988).

The limit of one also appeared in tests of attentional selectivity (Rensink, 1998b). Here, it was found that selection of items based on contrast polarity was essentially perfect, as might be expected. However, selection based on orientation (e.g. looking for polarity changes among vertical items) was poor, corresponding to the selection of just one item at a time.

5.3. Attentional mechanisms

Although the pattern of results described above is not sufficient to provide an unequivocal specification of the attentional mechanisms involved in change perception, it is sufficient to provide several constraints on its operation. In what follows, discussion will often be couched in terms of how these constraints affect the structure of coherence fields (Section 2.2); however, it should be kept in mind that many of these constraints can also be applied to other models of attentional representation (e.g. the object files of Kahneman et al., 1992).

The first of these constraints is *aggregation*, the constraint that the information from each attended item is not kept separate, but is — at least to some degree — pooled into an aggregate description at a mid-level nexus (Fig. 7). This constraint follows from the finding that capacity is much higher for presence of change than it is for absence of change (Rensink, 1999a). In essence, it posits that the cause of this asymmetry is the same as the cause of the asymmetry for simple features (Treisman & Gormican, 1988): the pooling of signals (across the attended items for change; across the visual field for features) causes detection of presence to be easier than detection of absence. More concretely, imagine a situation where information is pooled into one nexus, and where a sufficiently large change gives rise to a signal of 1 (Fig. 7). When looking for a changing item among constant distractors, the signal is 1 for target present and 0 for target absent, leading to relatively easy search. In contrast, when looking for a constant among changing items, the difference is much less: assuming that information can be collected from five links, the signal is 4 for target present and 5 for target absent. A more effective strategy in this case would then be to simply collect information from one link — and thus one nexus — at a time (Fig. 7).

Note that this constraint also explains why search is so much slower for detecting the absence of change. Given that only one to two items can be 'grabbed' at a time, and that search speed is about 300–400 ms (Rensink, 1999a), it follows that the formation of a coherence field takes about 400–500 ms[16]. Search for the presence of change can be done more quickly,

[16] This value is close to the attentional dwell time, which is sometimes thought to reflect the minimum amount of time that attention can be given to an item (Ward, Duncan & Shapiro, 1996). Whether this similarity a deeper connection between the two values is an interesting question for future work.

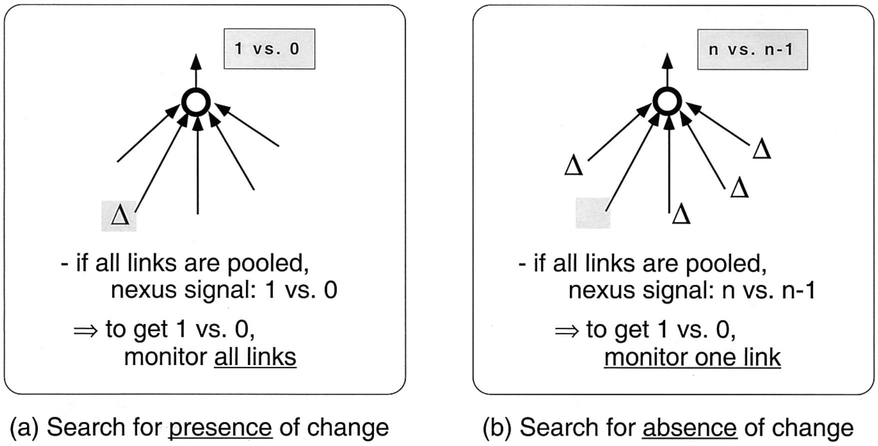

Fig. 7. Explanation for search asymmetry (Rensink, 1999a). (a) When searching for a changing target among constant distractors, there will be 1 change (target present) versus 0 changes (target absent). As such, there is a relatively strong signal differentiating target-present from target-absent trials, even when information from several links is collected. (b) When searching for a constant target among changing distractors, there will be $n-1$ changes (target present) versus n changes (target absent). Given that n is about 5 or 6, only a weak signal differentiates target-present from target-absent trials. To obtain a 1:0 signal ratio, the nexus must collect information from only one link at a time.

since more items can be incorporated into the field, thus reducing the number of times it needs to be set up.

Another constraint is that the number of items that can be independently attended at a time[17] can be no more than two, and possibly no more than one. This follows from the capacity limit of 1.4 items on search for absence of change, and the limit of 1.0 item on search for conjunctions — there appears to be no real sense in which two items can be independently checked to see if they still have the same orientation as before. In addition, a limit of 1 (i.e. a *singularity* constraint) could help explain why only one item at a time can be selected on the basis of orientation (Rensink, 1998b).

The difficulty of detecting conjunction changes gives rise to a third constraint: little (if any) intrafield binding occurs. If the orientation and polarity of each item could have been bound separately within the coherence field, capacity for conjunction changes would have been five to six, just as for simple property changes. A similar lack of intrafield binding may also explain the results of Saiki (1999) and Scholl, Pylyshyn and Franconeri (1999), who discovered a large amount of change blindness existed for remappings of colors, even when all the items involved were being tracked, and thus presumably attended. However, more work needs to be done before this matter can be completely resolved.

[17] It is of course possible to attend to more than one *item* in the image. However, if the set of attended items feeds into the same nexus, it is not possible to treat them as completely separate *objects*. For example, when threading a needle both the needle and the thread must be attended. However, they are not independent objects, but parts of a single needle-thread system. Note that attending to items in this way is not necessarily counterproductive. In the case of the needle and thread, the fact that their locations are sent to the nexus allows accurate calculation of the relative distance between them, thereby facilitating the threading of the needle.

Note that the possibility that attended information is collected into only one nexus — that in effect, only one object is attended at a time—has the potential to resolve one of the more vexing issues in the study of attention: the binding problem. Although this problem takes on many forms, at heart is the issue of how a set of properties (shape, location, etc.) assigned to one object can be prevented from being erroneously assigned to others (see e.g. Ashby, Prinzmetal, Ivry & Maddox, 1996; Treisman, 1996). Although many solutions have been proposed, none has won widespread acceptance (see Strong & Whitehead, 1989). However, if only one object is attended to at a time, there may simply be no problem — whatever is in the coherence field is assigned to that object. Different objects in the world can be assigned different properties by having a field first form for one object, then for the next object, and so on for the others. If objects can be given coherent representations when requested (Section 3.1) it will appear as if multiple objects are present, each with its own set of properties. Seen in this way, the binding problem is similar to the problem of combining the contents of successive fixations (see e.g. Bridgeman, van der Heijden & Velichkovsky, 1994): it is considered relevant only because of the way things appear to us as observers. In reality, however, it does not need to be solved. Rather, the problem to be addressed is one of *gating*, i.e. how to select the various features that are to enter into the coherence field at any moment in time.

6. Summary

The work surveyed here shows how change perception can be used as a tool to cast light on the operation of our visual system, both in regards to the general nature of seeing, and in regards to the operation of its more specialized aspects, such as sensing and scrutinizing. It also shows how coherence theory — originally developed to account for change blindness — can provide a rough theoretical framework that can help us understand several aspects of visual processing.

In regards to the issue of how we see, coherence theory indicates that our visual perception is based on a representation with qualities quite different from the world it represents — in particular, it is not simultaneously detailed and stable. Consequently, a durable representation such as an intrinsic image or primal sketch (e.g. Marr, 1982) cannot be the basis for visual perception. Instead of such a static structure, it appears that a dynamic *virtual representation* is used, which provides only a limited amount of detailed, coherent structure, but provides it whenever requested, making it appear as if all the detailed, coherent structure is present simultaneously. In this view, then, vision is an inherently dynamic process, a 'just in time' system whose detailed representations are in constant flux.

Empirical advances in the areas of sensing and scrutinizing have also begun to provide new insights into particular aspects of visual perception. Change-detection studies show that observers have a limited ability to detect a change in a scene even when they have no conscious experience of it. They also show that some observers can have a conscious experience of change without an accompanying visual experience. This suggests that nonattentional processing streams are able to detect change to some extent, and to influence the observer in ways not always open to conscious visual experience. Highly controlled studies on change detection have also provided considerable insight into focused attention, both about the underlying mechanisms and on the way attentional allocation depends on task structure and observer knowledge. An interesting possibility in this regard is that the binding problem may be illusory — it may be that the properties of only one object at a time are ever bound together.

Change perception will not provide insight into every aspect of visual processing. But change-perception experiments can range over a wide variety of stimuli (simple arrays to complex natural scenes), physical parameters (different durations of displays and interruptions), task demands (looking for various combinations of changes and non-changes), and levels of knowledge (naive observers to experts). They can therefore cover many of the interesting questions in vision, while remaining linked to each other. Consequently, such experiments have the potential to show how low-, mid-, and high-level processes are knit together, and how they allow us to interact with our environment.

Acknowledgements

I would like to thank Adriane Seiffert, Dan Simons, Ian Thornton, Carol Yin, and two anonymous reviewers for their comments on an earlier draft of this paper.

Appendix A. Taxonomy and terminology of visual operations

As long as vision was believed to be a unitary faculty, it was possible to use just a few (largely interchangeable) terms to refer to its operation — for example, 'seeing' and 'perceiving'. However, recent results suggest that vision involves a set of quasi-independent systems, with these being used in a number of different ways. Much of this paper is devoted to delineating the various systems involved. To help keep the resultant description in mind, this appendix provides an overview in the form of a brief taxonomy of visual operations. This taxonomy (including the choice of terms) is somewhat provisional. However, it has been based as much as possible on existing characterizations, and so should inherit a degree of robustness.

A simple sketch of the taxonomy is shown in Fig. A1. The first level is simply that of 'vision', i.e. the use of light to carry out actions of the system, regardless of what these actions are or how they are carried out. The term 'sight' will be used interchangeably with this, and 'seeing' will refer to the associated activity. Such usage is a purely behavioral one — it allows us to say things like 'The fly saw the wall' without concerning ourselves about whether or not it had an accompanying subjective experience. Note that the use of 'sight' in this nonrestrictive way allows consistent use of the terms 'blindsight' and 'mindsight'. It can also allow for more restrictive use (e.g. when referring to subjective experience) when the restrictions are clear from the context.

The first division in the taxonomy is that proposed by Milner and Goodale (1995): the partitioning of operations into *on-line* processes that use light for the immediate control of visuomotor actions (visual praxis) and *off-line* processes that form representations serving as the basis for more deliberate activities (visual perception). The next division is that of the perceptual operations, which are split into implicit processes (involving no conscious awareness of their output), and explicit processes (for which there is at least some awareness). Again, this division is a fairly well-established one (see e.g. Merikle & Joordens, 1997).

Finally, explicit perception is separated into visual awareness (in which there are no phenomenal impressions, or *qualia*), and visual experience (in which there

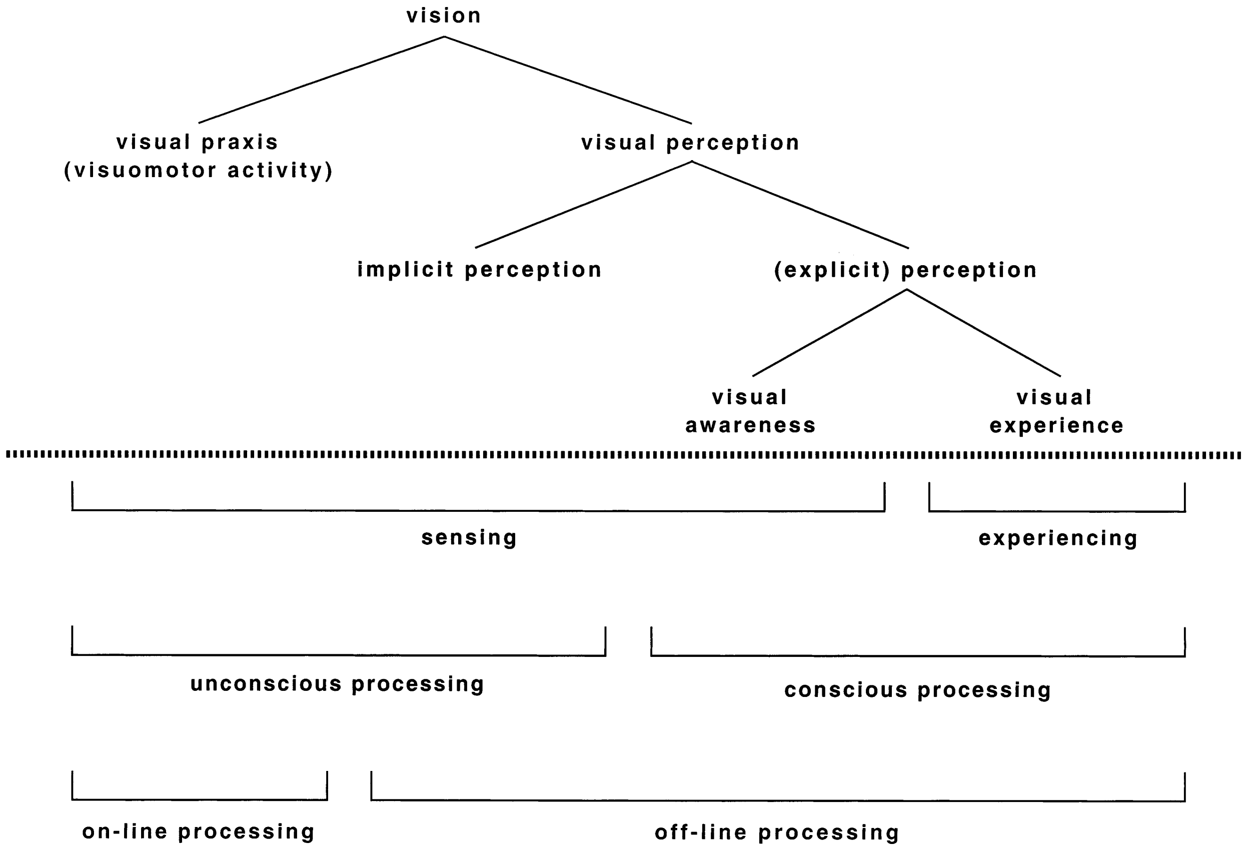

Fig. A1. Taxonomy of visual operations. Upper half shows relations of individual operations. Lower half shows groupings of these operations on the basis of visuomotor involvement, conscious awareness, and visual experience.

are). Note that in both cases, there is an accompanying conscious experience of some kind — the division depends on whether or not the experience is a *visual* one, i.e. whether qualia are an essential part of it). Also note that the term 'sensing' refers to all aspects of vision that do not involve visual experience, and so includes visual awareness as well as implicit vision.

Defined this way, the border between sensing and experiencing does not depend — at least in principle — on the involvement of focused attention (or *scrutiny*). Given that the relation between focused attention and visual experience has yet to be established (Section 3.3), it remains to be seen how attention fits into this taxonomy. If focused attention is necessary and sufficient for conscious visual experience, the current divisions will suffice; otherwise, a refinement of some kind will be needed.

References

Adelson, E. H., & Bergen, J. R. (1991). The plenoptic function and the elements of early vision. In M. Landy, & J. A. Movshon, *Computational models of visual processing* (pp. 3–20). Cambridge, MA: MIT Press.

Allport, A. (1992). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In D. E. Meyer, & S. Kornblum, *Attention and performance XIV* (pp. 183–218). Cambridge, MA: MIT Press.

Archambault, A., O'Donnell, C., & Schyns, P. G. (1999). Blind to object changes: when learning the same object at different levels of categorization modifies its perception. *Psychological Science*, *10*, 249–255.

Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, W. T. (1996). A formal theory of feature binding in object perception. *Psychological Review*, *103*, 165–192.

Avons, S. E., & Phillips, W. A. (1980). Visualization and memorization as a function of display time and poststimulus processing time. *Journal of Experimental Psychology*: *Human Learning and Memory*, *6*, 407–420.

Ballard, D. H. (1991). Animate vision. *Artificial Intelligence*, *48*, 57–86.

Bartlett, F. C. (1932). *Remembering: a study in experimental and social psychology*. Cambridge: Cambridge University Press (Reissued 1995.).

Becklen, R., & Cervone, D. (1983). Selective looking and the noticing of unexpected events. *Memory & Cognition*, *11*, 601–608.

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*, 1075–1081.

Bonnel, A-M., & Hafter, E. R. (1998). Divided attention between simultaneous auditory and visual signals. *Perception & Psychophysics*, *60*, 179–190.

Braun, J., & Julesz, B. (1998). Withdrawing attention at little cost: detection and discrimination tasks. *Perception & Psychophysics*, *60*, 1–23.

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

Bridgeman, B., Lewis, S., Heit, G., & Nagle, M. (1979). Relation between cognitive and motor-oriented systems of visual position perception. *Journal of Experimental Psychology*: *Human Perception and Performance*, *5*, 692–700.

Bridgeman, B., van der Heijden, A. H., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Brain and Behavioral Sciences*, *17*, 247–292.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, *47*, 139–159.

Castiello, U., & Jeannerod, M. (1991). Measuring time to awareness. *Neuroreport*, *2*, 797–800.

Castiello, U., Paulignan, Y., & Jeannerod, M. (1991). Temporal dissociation of motor responses and subjective awareness. *Brain*, *114*, 2639–2655.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Churchland, P. S., Ramachandran, V. S., & Sejnowski, T. J. (1994). A critique of pure vision. In C. Koch, & J. L Davis, *Large scale neuronal theories of the brain* (pp. 23–60). Cambridge, MA: MIT Press.

Clark, A. (1997). *Being there: putting brain*, *body*, *and world together again*. Cambridge, MA: MIT Press.

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

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, *3*, 357–370.

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

Dymtryk, E. (1984). *On film editing*. Boston, MA: Focal Press.

Feldman, J. A. (1985). Four frames suffice: a provisional model of vision and space. *Behavioral and Brain Sciences*, *8*, 265–289.

Fernandez-Duque, D., & Thornton, I.M. (2000). Change detection without awareness: do explicit reports underestimate the representation of change in the visual system. *Visual Cognition*, *7*, 324–344.

Findlay, J. (1998). Active vision: Visual activity in everyday life. *Current Biology*, *8*, R640–R642.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology*: *General*, *108*, 316–355.

Gelfand, I. M., & Fomin, S. V. (1963). *Calculus of variations*. Englewood Cliffs, NJ: Prentice-Hall (R. A. Silverman, translator).

Goodale, M. A. (1993). Visual routes to knowledge and action. *Biomedical Research*, *14*, 113–123.

Goodale, M. A., Pelisson, D., & Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature*, *320*, 748–750.

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

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

Henderson, J. M. (1996). Visual attention and the attention-action interface. In: K. Akins, *Perception*. *Vancouver Studies in Cognitive Science*, *vol. 5* (pp. 290–316). New York: Oxford University Press.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271.

Hildreth, E. C., & Koch, C. (1987). The analysis of visual motion: from computational theories to neuronal mechanisms. *Annual Review of Neuroscience*, *10*, 477–533.

Hochberg, J. (1968). In the mind's eye. In R. N. Haber, *Contemporary theory and research in visual perception* (p. 329). New York: Holt, Rinehart & Winston.

Hollingworth, A., & Henderson, J.M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, *7*, 213–235.

Intraub, H. (1997). The representation of visual scenes. *Trends in Cognitive Sciences*, *1*, 217–222.

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.

James, W. (1890). *The principles of psychology*. New York: Dover.

Julesz, B. (1984). A brief outline of the texton theory of human vision. *Trends in Neuroscience*, *7*, 41–45.

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

Kanwisher, N. G. (1987). Repetition blindness: type recognition without token individuation. *Cognition*, *27*, 117–143.

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

Lachter, J., & Hayhoe, M. M. (1995). Capacity limitations in memory for visual locations. *Perception*, *24*, 1427–1441.

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

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge MA: MIT Press.

Mandler, J., & Parker, R. E. (1976). Memory for descriptive and spatial information in complex pictures. *Journal of Experimental Psychology*: *Human Learning and Memory*, *2*, 38–48.

Marcel, A. J. (1983). Conscious and unconscious perception: an approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology*, *15*, 238–300.

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

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

Merikle, P. M., & Joordens, S. (1997). Parallels between perception without attention and perception without awareness. *Consciousness and Cognition*, *6*, 219–236.

Merikle, P. M., & Reingold, E. M. (1992). Measuring unconscious perceptual processes. In R. F. Bernstein, & T. S. Pittman, *Perception without awareness* (pp. 55–80). New York: Guilford Press.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Moore, C. M., & Egeth, H. (1997). Perception without attention: evidence of grouping under conditions of inattention. *Journal of Experimental Psychology*: *Human Perception and Performance*, *23*, 339–352.

Nakayama, K. (1985). Biological motion processing: a review. *Vision Research*, *25*, 625–660.

Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.

Neisser, U., & Becklen, R. (1975). Selective looking: attending to visually significant events. *Cognitive Psychology*, *7*, 480–494.

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, *34*, 72–107.

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.

O'Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, *7*, 191–211.

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

Pashler, H. E. (1998). *The psychology of attention*. Cambridge, MA: 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*, 179–197.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP: an attentional blink? *Journal of Experimental Psychology*: *Human Perception and Performance*, *18*, 849–860.

Rensink, R. A. (1998a). Mindsight: visual sensing without seeing. *Investigative Ophthalmology & Visual Science*, *39*, 631.

Rensink, R. A. (1998b). Limits to attentional selection for orientation. *Perception* (*Suppl*), *27*, 36.

Rensink, R. A. (1999a). The magical number one, plus or minus zero. *Investigative Ophthalmology & Visual Science*, *40*, 52.

Rensink, R. A. (1999b). Attentional processing of geometric figures. *Perception*, *28* (Suppl), 55–56.

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

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

Rensink, R. A., & 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. (1995). Image flicker is as good as saccades in making large scene changes invisible. *Perception* (*Suppl*), *24*, 26–28.

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.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual Cognition*, *7*, 127–145.

Rojer, A. S., & Schwartz, E. L. (1990). Design considerations for a space-variant visual sensor with complex-logarithmic geometry. *Proceedings of the 10th international conference of pattern recognition* (pp. 278-285) 16–21 June 1990, Atlantic City, NJ.

Sagi, D., & Julesz, B. (1985). 'Where' and 'what' in vision. *Science*, *228*, 1217–1219.

Saiki, J. (1999). Relation blindness: difficulty in detecting a violation of spatial relations in regular rotations of a triangular pattern. *Investigative Ophthalmology & Visual Science*, *40*, 49.

Sanocki, T. A., & Epstein, W. (1997). Priming spatial layout of scenes. *Psychological Science*, *8*, 374–378.

Schacter, D. (1987). Implicit memory: history and current status. *Journal of Experimental Psychology*: *Learning*, *Memory and Cognition*, *13*, 501–518.

Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. L. (1999). When are featural and spatiotemporal properties encoded as a result of attentional allocation? *Investigative Ophthalmology & Visual Science*, *40*, 797.

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

Sharpe, S. H. (1988). *Conjurers' psychological secrets*. Calgary, AB: Hades Publications.

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.

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

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

Stoerig, P., & Cowey, A. (1997). Blindsight in man and monkey. *Brain*, *120*, 535–559.

Strong, G. W., & Whitehead, B. A. (1989). A solution to the tag-assignment problem for neural networks. *Behavioral and Brain Sciences*, *12*, 381–433.

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

Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision*, *7*, 11–32.

Tanenbaum, A. S. (1976). *Structured computer organization*. Englewood Cliffs, NJ: Prentice-Hall.

Thornton, I. M., & Fernandez-Duque, D. (1999). Representing the 'what' and the 'where' of undetected change. *Abstracts of the Psychonomic Society*, *4*, 1.

Trehub, A. (1991). *The cognitive brain* (pp. 55–78). Cambridge, MA: MIT Press.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, *6*, 171–178.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological Review*, *95*, 15–48.

Ullman, S. (1996). *High-level vision*. (pp. xi–xii). Cambridge, MA: MIT Press.

Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology*, *30*, 79–109.

Weiskrantz, L. (1996). Blindsight revisited. *Current Opinion in Experimental Psychology*, *6*, 215–220.

Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain*, *97*, 709–728.

Werner, S., & Thies, B. (2000). Is 'change blindness' attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*, *6*, 163–173.

Wilken, P., Mattingley, J.B., Korb, K.B., Webster, W.R., & Conway, D. (1999). Capacity limits for detection versus reportability of change in visual scenes. Paper presented at the 26th Annual Australian Experimental Psychology Conference. 9–11 April, 1999, Sydney, Australia.

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

Wolfe, J. M. (1995). Preattentive object files: shapeless bundles of basic features. *Vision Research*, *37*, 25–43.

Wolfe, J. M. (1999). Inattentional amnesia. In V. Coltheart, *Fleeting memories* (pp. 71–94). Cambridge, MA: MIT Press.

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

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.

Yeshurun, Y, & Schwartz, E. L. (1989). Shape description with a space-variant sensor: algorithms for scan-path, fusion and convergence over multiple scans. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*, 1217–1222.