In D. Reisberg (Ed.), *Oxford Handbook of Cognitive Psychology* (2013), pp. 97-116

Perception and Attention

Ronald A. Rensink

Departments of Psychology and Computer Science

University of British Columbia

Vancouver, Canada

Dedicated to the memory of Jon Driver.

Correspondence concerning this article may be addressed to R.A. Rensink, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver BC V6T 1Z4, Canada. Email: rensink@psych.ubc.ca or rensink@cs.ubc.ca.

Introduction

Our visual experience of the world is one of diverse objects and events, each with particular colors, shapes, and motions. This experience is so coherent, so immediate, and so effortless that it seems to result from a single system that lets us experience everything in our field of view. But however appealing, this belief is mistaken: there are severe limits on what can be visually experienced.

For example, in a display for air-traffic control it is important to track all moving items. For a single item, this can be done without problem. Three or four can also be tracked, although some degree of effort may be needed. As the number is increased further, accurate tracking becomes more and more difficult—and eventually, impossible. Performance is evidently affected by a factor within the observer which enables certain kinds of perception to occur, but is limited in some way. This factor is generally referred to as *attention*.

At various times, attention has been associated with clarity of perception, intensity of perception, consciousness, selection, or the allocation of a limited "resource" enabling various operations (see Hatfield, 1998). During the past several decades, considerable progress has been achieved by focusing on the idea of selection (Broadbent, 1982). In particular, attention can be productively viewed as *contingently selective processing*. This can be embodied in various ways by various processes—there need not be a single quantity identified with all forms of attention, or a single site where it operates (Allport, 1993; Tsotsos, 2011). Although "paying attention" is often considered to be a unitary operation, it may simply refer to *the control of one or more selective processes*, ideally in a co-ordinated way. While this view has some cost in terms of conceptual simplicity, it can help make sense of a large set of phenomena.

This article surveys several of the major issues in our understanding of attention and how it relates to perception. It focuses on vision, since many—if not all—considerations are similar for all sensory modalities, and the level of understanding achieved in this domain is currently the most advanced. Although many of the issues discussed here also apply to higher-level cognitive processes, coverage of those is best done separately. [Also see THEORIES OF ATTENTION; EXECUTIVE FUNCTIONS.]

Kinds of Attention

A great deal of work has been carried out on attention over the years, with a wide variety of approaches used on a wide variety of issues (see e.g., Parasuraman, 1998; Pashler, 1999; Wright, 1998). Partly because of such heterogeneity, studies of attention have tended to form "silos" that focus on a particular type of experimental approach or proposed mechanism. To provide a more comprehensive view, this discussion includes all of these issues, but is organized around the possible *functions* that attention might have.

In the view taken here, "attention" is more an adjective than a noun: an attentional process is one that is *contingently selective*, controlled on the basis of global considerations (e.g., selection of a particular item to track, based on perceived importance). Any limited-capacity process is necessarily selective, in that it cannot handle everything; if coupled with appropriate control it becomes *ipso facto* attentional. In the case of a limited-resource process, attention is sometimes considered a kind of "fuel" (see Hirst & Kalmar, 1987; Wickens, 1980). However, this does not really help with understanding—it is essentially just a different term for the same problem (Allport, 1989). It may instead be better to discuss processing resources in terms of more concrete measures such as time or processor connections, which if limited can create a need for selectivity, or may constrain a process, but which are not themselves a form of selection or control. [Also see THEORIES OF ATTENTION.]

Given all this, it is evident that many kinds of selectivity are possible, each corresponding to a particular kind—or at least, aspect—of attention. These are organized here into a taxonomy describing several of the main kinds possible, and the relationships among them. This *function-centered taxonomy* characterizes each kind of attention primarily in terms of its (selective) function. This is followed by the perceptual effects associated with it, the mechanisms that carry it out, and the ways that it enters into our conscious experience of the world.

This taxonomy is incomplete in many ways, and parts of it are provisional. But it encompasses much of what is believed about attention, and provides a natural way to organize a wide variety of results obtained in a wide variety of ways.

0. Attentional Sampling

0.1. Function

The first step of visual perception is the pickup of information by the eye. This is carried out by the selective sampling of incoming light, via a retina with most of its information pickup in the central few degrees of visual angle (see e.g., Barlow, 1981). To compensate for the relatively poor pickup elsewhere, the eye is continually repositioned via brief jumps, or *saccades*. Augmented by head and body movements, these enable access to any part of the incoming light (see e.g., Henderson & Hollingworth, 1998). A somewhat analogous situation exists in the third dimension (depth), with high-resolution pickup only for the depths that are in focus, a restriction compensated for by continually changing the focus of the lens. When controlled on the basis of global considerations, such *attentional sampling* can enable the right information to be obtained from the environment at the right time, thereby minimizing the computational resources needed for many tasks. In essence, processing over space is traded for processing over time (see Niebur & Koch, 1998). [Also see EYE MOVEMENTS.]

Attentional sampling—sometimes referred to as *overt attention*—is not the only kind of attention possible. Others also exist, occasionally referred to collectively *as covert attention*. Overt and covert systems are at least partly separate: they need not—and often do not—act on the same information at the same time. For example, although covert spatial attention must go to a location before the eye can saccade there (Deubel & Schneider, 1996), it is not necessary to make—or even intend to make—a saccade to generate a covert shift (Juan, Shorter-Jacobi, & Schall, 2004). Indeed, much of the correlation between these two kinds of attention may arise simply because they tend to respond to the same informative parts of the visual input (Tsotos, 2011, ch. 1).

0.2. Perceptual Effects

The limitations of sampling are generally compensated for by appropriate movements of the eye. But they can become evident when such movements are prevented, such as by having the observer fixate a location, or view the stimulus during a brief flash. Results show that acuity and color perception are best in the central few degrees (or *foveal* area) and fall off rapidly with increasing eccentricity in the outer parts (or *periphery*); in contrast, motion perception is poor in the fovea but better in the periphery (e.g., Barlow & Mollon, 1982).

Most kinds of eye movement are rapid, and made several times a second. The particular sequence (or *scan path*) used for viewing an image depends not only on its physical properties, but also on the knowledge and interests of the viewer, and their understanding of the task (Hayhoe & Ballard, 2005; Noton & Stark, 1971).

0.3. Mechanism

Many of the systems underlying sampling are fairly well understood. Incoming light is transformed into neural signals by two interleaved arrays of photoreceptors: *cones* (for high levels of illumination), and *rods* (for low levels). Cones are located mostly in the fovea, rods in the periphery. Because cones alone distinguish color, this property is perceived only in the central few degrees of visual angle (see e.g., Barlow & Mollon, 1982). Meanwhile, an array of *ganglion cells* takes the photoreceptor signals and—after several stages of processing—transmits the results to the brain via the optic nerve. Each ganglion cell receives input from the receptors in a particular area of the retina (its *receptive field*); the size of this increases with eccentricity (see e.g., Kaplan, 1991). Acuity is therefore best in the fovea, where receptive fields are small. Meanwhile, accurate measurement of motion requires the collection of information over a large area, and so motion perception is best in the periphery.

The eye itself is moved via three pairs of muscles, co-ordinated such that one member of each pair contracts while the other relaxes. These are controlled via several areas in the brain, such as the superior colliculus and the frontal eye fields (see e.g., Crowne, 1983). And these in turn are under the control of areas involving higher-level functions (see e.g., Carpenter, 1988). [Also see EYE MOVEMENTS.]

0.4. Relation to Consciousness

Because information pickup is required for visual perception, attentional sampling is necessary for any conscious (*explicit*) as well as nonconscious (*implicit*) process. But although our visual experience is quite compelling, it does not directly correspond to what is picked up by the retina. For example, the limited range of color and motion sampling is never noticed: we experience the same resolution and color information throughout. Likewise, although saccades are usually made several times each second (often over several degrees of visual angle), our impression is that of a single, stable image. The mechanisms that create such uniformity and stability are largely unknown (Bridgeman, van der Heijden, & Velichkovsky, 1994; O'Regan & Noe, 2001). But their involvement means that although attentional sampling may be necessary for visual experience, it cannot be sufficient.

The involvement of these mechanisms also explains why even though sampling may be necessary for visual experience at a global level, this necessity does not always apply locally. For example, the retina contains an area where the outputs of the ganglion cells exit, rendering it unable to pick up information. But although this *blind spot* may cover as much as 7-10 degrees of angle, it does not enter into our visual experience: even when the eye is fixated, it appears "filled in", so the gap is not noticed (Ramachandran, 1992). Similar effects occur for blind spots created by small lesions elsewhere in the visual system, as well as by stimulus manipulations (Ramachandran & Gregory, 1991). [Also see EYE MOVEMENTS.]

1. Attentional Filtering

1.1. Function

Although a vast amount of information is generally picked up at any moment, most is irrelevant for any particular task. And irrelevant information can be harmful. For example, detection is affected by the noise—or more precisely, the *signal-to-noise ratio*—in the input (e.g., Palmer, Ames, & Lindsey, 1993). Since irrelevant information contains no signal but may contain noise, it can only hurt performance. Likewise, because recognition of a pattern is affected by context (e.g., Oliva & Torralba, 2007), irrelevant background can only interfere. And if not disregarded in some way, irrelevant information can place a load on some processors, causing them to slow down (Tsotsos, 1990). Thus, if perception is to be efficient, it must select the relevant, and only the relevant.

Such *filtering* is in some ways an internal form of sampling. But because it is controlled via changes in neural connections rather than physical forces on the eye, filtering can act more quickly. It can also be more flexible, selecting information from particular regions of space, and with particular ranges of properties (such as color or orientation). This need not be all-or-nothing—a filter can be graded, with inputs weighted as a function of particular values.

Two complementary approaches to filtering are possible: selection *for* and selection *against*. Selection *for* relevant information results in the contents being *expressed*. One way of doing so is *gating*—transmitting information at a particular location, or of a particular color, say. In addition, contents can be *enhanced*, having an effect—or at least, priority—above baseline levels. In a similar way, selection *against* irrelevant information results in the selected contents being *suppressed*. When applied to the control of process or behavior, the two approaches are often referred to respectively as *facilitation* and *inhibition* (Posner & Snyder, 1975). Both are largely equivalent, although practical considerations may favor one over the other.

Another issue—especially important in the case of space—is the extent of gating. When selection is *diffuse* (input from a large area), considerable information is simultaneously accessible, which can enable several kinds of process to be sped up. However, more noise may be introduced, creating greater risk of error. Conversely, when selection is *focused* (input from only a small area), processing is potentially less error-prone but slower. These are sometimes referred to respectively as *diffuse attention* and *focused attention*. There is consequently a trade-off between speed and accuracy depending on the size selected, with the most effective choice depending upon details of the task and environment (see Giordano, McElree, & Carrasco, 2009).

1.2. Perceptual Effects

As in the case of sampling, most observers are unaware of the limitations imposed by filtering. But under the right conditions, they can lead to measurable effects. Although not all of these necessarily involve the same neural systems, they are similar in several ways:

i) Cueing effects

Observers can detect a target item more quickly and accurately if presented with advance information about its location. This *cueing* occurs even when eye movements are prevented, indicating that it is due to attentional (or covert) *orienting*, the adjustment of internal filters to task demands. These filters block information from irrelevant locations or properties, improving performance by reducing noise.

A transient form of enhancement begins within 50 ms of cue onset and peaks 50-100 ms later. It appears to be due to a *spotlight of attention* that expresses whatever is within the "beam" (Posner, Snyder, & Davidson, 1980) and suppresses whatever is just outside (Steinman, Steinman, & Lehmkuhle, 1995; Tsotsos, 2011). The size of this is about 1 degree at fixation, and increases with eccentricity (Sagi & Julesz, 1986). It falls off gradually, rather than having a sharp edge (Laberge & Brown, 1989). Its location can be shifted extremely rapidly, at a rate of c. 10 ms/deg (Tsal, 1983) or faster (see Egeth & Yantis, 1997); intermediate locations do not appear to be affected (see Cave & Bichot, 1999). Only one spotlight usually exists at a time (Jonides, 1983), although under some conditions it may be split among several locations (Bichot, Cave, & Pashler, 1999; Kramer & Hahn, 1995; but see Wright & Ward, 2008, ch. 3).

Cueing is also possible for properties such as color and orientation (Humphreys, 1981; Rossi & Paradiso, 1995; Maunsell & Treue, 2006), as well as spatial frequency (Davis, Kramer,

& Graham, 1983). It can also affect purely subjective aspects of perception, such as apparent contrast (Anton-Erxleben, Abrams, & Carrasco, 2010; Carrasco, Ling, & Read, 2004).

In general, cueing involves simple properties such as orientation or location. But there is also an influence of larger-scale "preattentive" structures—or *segments*—in the background[1], these apparently formed on the basis of luminance or texture boundaries at early levels of processing, before attentional filtering can act (e.g., Driver, Davis, Russell, Turrato, & Freeman, 2001; Lamme, 1995). For example, cueing is facilitated if cue and target are both in a region interpreted as the surface of the same segment (Egly, Driver & Rafal, 1994), but not if this region is interpreted as a hole (Albrecht, List, and Robertson, 2008). Consistent with this, cueing appears to operate over perceived rather than retinal space (Robertson & Kim, 1999).

ii) Search asymmetry

In *visual search*, observers must detect as quickly as possible the presence of a *target* item among a set of *distractors*—e.g., a vertical line among a set of obliques. For many kinds of search, the time needed is proportional to the number of items in the display. This has been explained via a spotlight of attention that accesses information from each item in turn, at a rate of 30-50 ms per item (see Wolfe et al., 2000; but see Thornton & Gilden, 2007). Search slows with increasing eccentricity (Carrasco, Evert, Chang, & Katz, 1995), possibly due to the sparser representation of peripheral areas in the cortex (Carrasco & Frieder, 1997).

When the target and distractors differ only by the amount of a simple property (e.g., length or luminance), an interesting *asymmetry* emerges: search is faster for a longer item among short ones, say, than the other way around. This has been explained by a pooling of the signals in the spotlight. When a target has more of a property, it creates a higher signal. Since accuracy is determined by the ratio of signal to noise, a larger spotlight—which collects more noise—can be used while keeping accuracy the same. Since a larger spotlight requires fewer shifts to cover an area, the result is faster search (Treisman & Gormican, 1988).

Although the properties involved in search may be simple (e.g., length or orientation), they are not directly obtained from simple structures such as line segments. Instead, they appear to be measurements of complex structures—*proto-objects*—of limited extent, formed at early levels prior to attentional filtering (Rensink, 2000a; Rensink & Enns, 1995). Search speed is also affected by background segmentation, being faster for items on the same surface (He & Nakayama, 1995).

iii) Interference effects

Performance in brief displays can be degraded by nearby items, or flankers (Eriksen & Eriksen, 1974). This *flanker effect* is believed to be caused by a pooling of the information in the spotlight, with signals from the flankers being mixed in with those from the target. Such effects point to a spotlight with a minimum area of 1 degree of angle in the fovea (Eriksen & Eriksen, 1974). It can "zoom out" to a larger area if needed, changing size within 100 ms (Eriksen & St. James, 1986), and automatically becomes smaller when task difficulty or load is increased (Lavie, 1995). Only one such spotlight is believed to operate at a time (Eriksen & Yeh, 1985).

A related effect is *crowding,* where items in the periphery that are easy to recognize when alone become less so when other items are placed nearby; subjectively, their components become “scrambled”. The size of the affected area increases with eccentricity (Bouma, 1970); it can be reduced somewhat if target location is cued (Yeshurun & Rashal, 2010). While part of crowding may be due to poor information pickup in the periphery, much of it appears due to the loss of positional information about the components; indeed, crowding may be much the same as the flanker effect (Strasburger, 2005).

As in the case of cueing and search, interference is influenced by structural factors. For example, it is greater for items grouped by common motion (Driver & Baylis, 1989) or on the same background segment (Kramer & Jacobson, 1991). And the expression of a face can be still perceived even when crowding destroys conscious awareness of the detailed configuration that conveys this (Fischer & Whitney, 2011).

iv) Inattentional blindness

Consider an observer carrying out a demanding primary task, such as determining which of two similar lines is longest. If an unexpected irrelevant item is shown at the same time, it will often not be seen, especially if less than 1 degree in size. This *inattentional blindness (IB)* has been explained by a failure to "attend" the irrelevant item—i.e., a failure to transmit the information from its location (Mack & Rock, 1998). This can last for several seconds, provided that the irrelevant stimulus remains unexpected (Most et al., 2001; Simons & Chabris, 1999). Otherwise, IB virtually disappears, possibly because the filter(s) involved can be reconfigured quickly. IB appears to involve the suppression of items with visual properties similar to those that are ignored (Most et al., 2001).

An important issue is whether the failure to report an irrelevant item results from a failure to *perceive* it, or to *remember* that it appeared (Wolfe, 1999). If the primary task is sufficiently demanding, a degree of IB can be induced for repeated stimuli (Cartwright-Finch & Lavie, 2007; Rensink, 2005). Assuming that advance knowledge about the appearance of an item (obtained via the repetition) can prepare the viewer to place a record of it into durable memory at the moment it appears, this indicates that IB is a failure of perception.

v) Interocular suppression

In *binocular rivalry (BR)*, two stimuli are simultaneously presented to an observer, one to each eye. But instead of the observer experiencing a combination of the two stimuli, the percept alternates, with only one stimulus experienced at any time (Blake, 2001). This occurs both for simple stimuli and complex events (Neisser & Becklen, 1975), and likely results from the suppression of information on the basis of eye of origin.

A related effect is *continuous flash suppression (CFS)*, where a set of random images flashed into one eye at a rate of about 10 Hz completely suppresses the experience of an image shown to the other (Tsuchiya & Koch, 2005). This effect can be sustained for several minutes. The degree of suppression is more severe than that found in BR, although some of the same neural systems may be involved (Tsuchiya, Koch, Gilroy, & Blake, 2006). The afterimage of a suppressed stimulus is weakened, consistent with a suppression of information even to mechanisms not directly involved with conscious perception (Kaunitz, Fracasso, & Melcher, 2011; Tsuchiya & Koch, 2005).

vi) Neglect (Hemineglect; Unilateral neglect)

Damage to the right posterior parietal cortex can result in *neglect*—the absence of visual experience in some part of space (Bartolomeo & Chokron, 2002; Bisiach, 1993). Two variants can be distinguished. In *egocentric* neglect, the absence is of one side—usually the left—of viewer-centered space (Posner, Walker, Friedrich, & Rafal, 1984); in *allocentric* neglect, absence is of one side—usually the left—of each item in the entire visual field (Driver & Halligan, 1991). A related condition is *extinction*, where an item vanishes when a competing item appears on the opposite side of the visual field. These deficits appear to result from a failure to attentionally gate—i.e., transmit—information from the appropriate area (or structure), causing either an outright failure to access the information it contains, or at least a slowdown in the speed of processing (Posner et al., 1984). These deficits can be influenced by the structure of unseen objects, as well as by background segmentation (Rafal, 1994).

1.3. Mechanism

The effects described above appear to involve a mechanism with a characteristic mode of operation: it acts within a very brief time (c. 50 ms), involves simple variables (such as spatial location, orientation, and color), and for spatial variables operates like a reconfigurable spotlight. It is modulated by a retinotopic (i.e., eye-centered) organization, with performance deteriorating as eccentricity increases, likely due to the sparser representation of peripheral areas in the cortex. The filtering common to all of these is sometimes referred to as *spatial attention* or *space-based attention* (see Cave & Bichot, 1999). The information transmitted does not seem to keep track of its precise position in the visual field; when it is pooled, confusions often arise.

Although the variables that control filtering are simple, the structures it operates upon are not, having at least some degree of organization (also see Duncan, 1984; Egeth & Yantis, 1997). Background structures—*segments*—appear to facilitate information access (also see Franconeri, Bemis, & Alvarez, 2009; Kimchi, 2009). Filtering in turn can influence these, possibly via feedback connections (Driver et al., 2001). Likewise, although the variables transmitted (such as length) are simple, they are derived from relatively complex proto-objects in which various properties have already been bound together to some extent, presumably at low levels (Rensink, 2000a; Rensink & Enns, 1995). [Also see ORGANIZING THE SENSORY INPUT.]

Models of these mechanisms include those of Posner et al. (1980) and Treisman & Gelade (1980). Neural implementation may involve the dynamic restriction of receptive fields of neurons in visual cortex (Desimone & Duncan, 1995; Moran & Desimone, 1985). Control has been modeled by a system with three components: disengagement, shifting, and reengagement (Posner et al., 1984). A more general framework for filtering is that of Nakayama (1990). Detailed computational models include those of Cave (1999), Niebur & Koch (1998), and Tsostos (2011). [Also see SPATIAL ATTENTION.]

1.4. Relation to Consciousness

Much of consciousness appears to involve the consolidation of information over a global scale (Baars, 2005; Cohen & Dennett, 2011). This potentially connects to attention in at least two ways. First, conscious experience might be involved in transmitting the *output* of an attentional process elsewhere. Second, given that an attentional process can be responsive to global considerations, experience may enter into the control of its *operation*—e.g., the override of automatic settings (Libet, 1985). As such, there may well be an association between conscious experience and at least some aspects of attentional filtering.

i) Conscious perception

Attentional filtering—or at least, gating—appears necessary for conscious visual experience. In binocular rivalry, for example, observers simply fail to experience anything from the unmonitored eye (see Rensink, 2009). A similar failure occurs in inattentional blindness when the irrelevant stimulus is less than a few degrees (Mack & Rock, 1998), and in neglect, where gating has presumably failed entirely (Posner et al., 1984).

An even deeper relation may exist. When given a primary task of identifying a pattern at the center of a brief display, observers can still detect simple *features* elsewhere, such as localized colors or orientations (Braun & Sagi, 1990). They can also rapidly determine the meaning—or *gist*—of an image (Li, VanRullen, Koch, & Perona, 2002; Oliva & Torralba, 2006), and summary statistics such as average size of items (Ariely, 2001; Chong & Treisman, 2003). Here, gating may occur over a relatively large expanse, possibly in a sequential way (VanRullen, Reddy, & Koch, 2004). As the primary task becomes more demanding, gist perception begins to fail (Cohen, Alvarez, & Nakayama, 2011). If this is due to reduced gating (Lavie, 1995), it would suggest that gating—but possibly no other kind of attention—is necessary for any process concerned with the fragmentary aspects of an image.

The perception of these fragmentary aspects appears to correlate with a distinct kind of subjective experience. When an irrelevant IB stimulus is sufficiently large and the primary task not too demanding, observers often report a "fragmented" experience, being able to see "something" but not being able to say what it is (Neisser & Becklen, 1975). Such a base-level—or *ambient*—experience may be similar to the *background consciousness* of briefly-presented displays (Iwasaki, 1993), where contents are fleeting, fragmentary, and seem to contain more than can be reported. It may also be related to the *phenomenal-* or *P-consciousness* taken to describe raw experience (Block, 1995; also see Lamme, 2003).

ii) Nonconscious perception

Visual processing in the complete absence of visual experience—and presumably, spatial gating—is evident in many situations. During binocular rivalry, for example, faces and familiar stimuli in the unmonitored eye are not experienced, but can enter conscious awareness more quickly (Jiang, Costello, & He, 2007). And during inattentional blindness, an unseen item can cause *priming*—speeding up the identification of a related item that is presented soon afterwards (Mack & Rock, 1998). It can even draw attention to itself (Mack & Rock, 1998; but see Harris & Pashler, 2004).

Such effects are likely related to those found in *subliminal perception*, where observers generally fail to consciously experience a target that is presented extremely briefly (c. 20 ms), presumably because attentional processes are not given sufficient time to operate. But although not consciously experienced, stimuli presented under these conditions can still cause priming; interestingly, this occurs only when the time of their appearance is predictable, indicating the involvement of (possibly automatic) temporal gating (Naccache, Blandin, & Dehaene, 2002).

More generally, effects such as these suggest the existence of what might be called "*dark structure*"—structure that never can be part of visual experience, yet still affects perception. Whether the structures that influence filtering are also dark, or just the initial stages of structures that are later experienced, is an open question.

2. Attentional Binding

2.1. Function

At some point, the pieces of information obtained in the initial stages of vision must be *integrated*. This already occurs to an extent at early stages of processing, e.g., the creation of background segments and proto-objects (Section 1.3). But to better capture the structure of the world, a more sophisticated structure is needed, one in which *all* the properties relevant at a moment in time are appropriately linked, and connected to a particular position in space.

One aspect of this is the linking of properties at the same location. This can occur at early levels of visual processing, with some neurons responding to particular combinations of simple properties (see e.g., Sincich & Horton, 2005); it also shows up at higher levels (e.g., van Dam & Hommel, 2010). But other aspects also exist, which may involve greater distances, greater precision, greater numbers of properties—and therefore, greater challenges (see Tsotsos, 2011, ch. 6). And all of these may require that structures be created—or at least modified—on the basis of global considerations. In other words, integration may require *attentional binding*.

Attentional binding is often associated with attentional filtering. Indeed, the two are sometimes considered to be much the same, with binding referred to as "focused attention". But binding involves rather different functions. For example, if a line segment cannot be assigned to a particular group on the basis of purely local factors, its assignment must be based on more global considerations. Similar concerns exist for other aspects of structure, e.g., high-resolution estimates of orientation. More generally, the determination of integrated structure requires solving the *binding problem*: how to correctly link all relevant components (Robertson, 2003; Treisman, 1999). As such, binding is considered here to be a separate attentional process.

2.2. Perceptual Effects

The limitations of attentional binding show up in various ways. As in the case of filtering, not all necessarily involve the same neural systems. But they do appear to have several elements in common:

i) Conjunction search

Experiments on visual search (Section 1.2) show that the time needed to detect a unique combination (or *conjunction*) of simple features is proportional to the number of items present. This is believed due to a spotlight of attention that binds the features at each location, at a rate of 30-50 ms per item (Treisman & Gelade, 1980; Wolfe, 1998; but see Eckstein, 1998). Search is slower at greater eccentricity, perhaps due to sparser representation in the corresponding cortical areas (Carrasco et al., 1995; Carrasco & Frieder, 1997). Comparable rates are found when spatial form needs to be resolved (Duncan & Humphreys, 1989; Kleiss & Lane, 1986).

Not all combinations need attentional binding. Relatively fast search is possible for some scene properties and groups (Rensink & Enns, 1995, 1998), suggesting that it can access proto-objects that already contain a considerable degree of binding (but see VanRullen et al., 2004). In addition, some kinds of conjunction search can be sped up if the target is known ahead of time and the items are relatively far apart. Such *guided search* may occur because items with one of the target properties are picked out via filtering, resulting in a simpler task involving the remainder (Treisman & Sato, 1990; Wolfe, 1994). If the selection of one property in a structure can result in the immediate selection of the others, a possible explanation is that some degree of binding is in place prior to attentional processing.

ii) Illusory conjunctions

When nearby items with different properties are displayed briefly (e.g., 200 ms), viewers often report seeing the properties combined incorrectly (Prinzmetal, 1995; Treisman & Schmidt, 1982). These *illusory conjunctions* have been explained in terms of a process that requires 100-200 ms to link the features within the spotlight of attention and determine a definite position for each. If insufficient time is given, positional information remains poorly established, creating the possibility of erroneous bindings (Cohen & Ivry, 1989).

In foveal vision, illusory conjunctions occur primarily when items are within 1 degree of each other (Cohen & Ivry, 1989). In peripheral vision, they can occur for larger separations and slower speeds (Prinzmetal, Henderson, & Ivry, 1995), possibly because crowding (Section 1.2) increases the difficulty of access to individual features (also see Pelli, Cavanagh, Desimone, Tjan, & Treisman, 2007). Illusory conjunctions occur more frequently within perceptual groups (Prinzmetal, 1995), showing the influence of background structures (such as segments) that contain a degree of organization. They are unlikely if the binding would create a semantic association not in the original items—e.g., there is little illusory conjunction of circle plus black if the result would be a tire not present originally (Treisman, 1986).

iii) Repetition blindness

If an observer looks for a target item in a stream of briefly presented-stimuli (100-200 ms each), a repeated target is much less likely to be seen than a different item at the same position in the sequence. Such *repetition blindness (RB)* is thought to occur because the second item is perceived as an instance of a category (or *type*) encountered recently, which prevents triggering the formation of a new integrated structure (Kanwisher, 1987).

The type involved in RB is somewhat abstract, being invariant to size, orientation, and viewpoint (Kanwisher, Yin, & Wojciulik, 1999). This suggests it can be determined within about 100 ms, which in turn suggests that the process involved draws upon structures that already have some degree of binding. Only those items that are filtered in (e.g., those with a particular color or orientation) appear to be relevant for determining what constitutes a repetition (Kanwisher, Driver, & Machado, 1995).

iv) Integrative agnosia (Apperceptive agnosia; Visual space agnosia)

Damage to the inferior temporal lobe sometimes results in *integrative agnosia*, an inability to perceive overall shape or configuration, with perception being only of simple features, such as color or texture (Farah, 2004; Riddoch & Humphreys, 1987). This is likely caused—at least in part—by a disabling of attentional binding. The ability to determine which items belong together appears to be unaffected, consistent with the idea that several aspects of (nonattentional) binding are carried out prior to attentional binding (Behrmann & Kimchi, 2003).

2.3. Mechanism

The effects above involve bound structures that enable processes such as recognition of complex shapes and characters. Most indicate a mechanism with a time constant of roughly 150 ms, shorter than the 300 ms typical of a fixation, but longer than the 50 ms characteristic of filtering. The 50 ms value encountered in conjunction search may reflect a mechanism based on filtering (Treisman & Gelade, 1980), but it could also simply reflect the time needed to gate an item into a pipeline that operates on several items concurrently, with each processed for several hundred milliseconds (Wolfe, Klempen, & Dahlen, 2000).

The mechanism involved appears to operate upon segments and proto-objects that already have a degree of binding, much—if not exactly—the same as the structures operated upon by filtering. Indeed, *feature integration theory* (Treisman & Gelade, 1980) posits that filtering is the basis of a good part of attentional binding. For example, when searching for a conjunction of different kinds of properties, or *dimensions* (e.g., a red T among red As and blue Ts), it proposes that the spotlight of attention restricts gating to a small location, resulting in the simultaneous detection of all features there.

But for many tasks, attentional filtering and attentional binding appear to be carried out by different mechanisms (Bravo & Nakayama, 1992; Briand & Klein, 1986; Maddox, Ashby, & Waldron, 2002; VanRullen et al., 2004); these may involve respectively the *transient* and *sustained* forms of attention that are sometimes distinguished (Nakayama & Mackeben, 1989). In addition, the resolving of fine detail appears to involve still another mechanism, one that combines information from filters operating within a single dimension (Laberge, 1998).

The neural systems that implement (attentional) binding are largely unknown (Robertson, 2003). One possibility is the synchrony in firings of neurons that belong to the same group

(Singer, 1999). Another—not necessarily exclusive—account is based on the dynamic, selective routing of information from various areas of cortex (Olshausen, Anderson, & Van Essen, 1993), likely involving feedback connections between several levels of processing (Di Lollo, Enns, & Rensink, 2000; Lamme, 2003; Tsotsos, 2011, ch. 7). [Also see THEORIES OF ATTENTION.]

2.4. Relation to Consciousness

Just as conscious experience is associated with some aspects of attentional filtering, it may also be associated with some aspects of attentional binding. And as in the case of filtering, this could be in terms of both output and operation.

i) Conscious perception

Attentional binding appears necessary to experience structure containing integrated properties, spatial extension, and a well-defined position. This kind (or perhaps level) of *assembled experience* comprises ambient experience (providing basic sensory properties) along with a degree of additional (bound) structure; it may be related to the *access*- or *A-consciousness* involved with verbal reports (Block, 1995; also see Lamme, 2003). If attentional binding is diverted to some task, irrelevant visual stimuli can still be experienced, but only in ambient form, either the fragmentary features themselves (Braun & Sagi, 1990) or properties based directly upon them, such as scene gist (Li et al., 2002; VanRullen et al., 2004). Consistent with this, ambient experience alone appears to be encountered in integrative agnosia, likely due to the loss of attentional binding.

There is some possibility that attentional binding is involved in creating the "figure" in figure-ground organization (Vecera, Flavaris, & Filapek, 2004), while other kinds of attention—such as those involved change detection—are not (Kimchi & Peterson, 2008). Given that "figure" is generally experienced in an assembled way (and "ground" in an ambient one), it may be that attentional binding is associated at least to some extent with assembled experience.

ii) Nonconscious perception

In integrative agnosia, the overall configuration of an object is not experienced in an assembled way (i.e., as a bound structure), but is still available for motor tasks such as grasping (Goodale & Milner, 1992). This suggests that integrated structure is still captured in other processing streams, possibly as dark structure that is never experienced consciously. Indeed, even when background organization is not seen in an assembled way, it can still affect the perception of stimuli that are (Driver et al., 2001; Moore & Egeth, 1997). Such phenomena are consistent with the proposal that attentional binding is involved in the control of processes that have already captured a considerable degree of structure.

3. Attentional Holding

3.1. Function

Although binding is important for various tasks, it is not always enough. For example, when a physical object changes, it is helpful to see it as a single *visual object* that endures over time, rather than a sequence of unrelated structures. Such *continuity* likely plays a central role in the perception of physical objects and events. However, establishing this additional level of structure is a complex process, in that both permanence (persistence of identity) and change (transformation of properties) must be handled simultaneously (Kahneman, Treisman, & Gibbs, 1992). [Also see OBJECT PERCEPTION.]

Such continuity can be achieved via the idea of *coherent* representation. Here, properties are linked not only across space but also across time, so as to refer to a single persisting object

(Rensink, 2000a). As in the case of binding (and to some degree, filtering), this process may take time. Indeed, owing to the greater complexity of coherent structure, such *attentional holding* may require considerably more in the way of various computational resources.

3.2. Perceptual Effects

Similar to the situation for other kinds of attention, the limitations of attentional holding show up in various ways, with several common characteristics:

i) Object-specific preview benefit

When several items in a display are presented briefly (250 ms), a letter or digit located inside one of them can be identified more quickly the next time it appears, provided it remains inside the original item. This *object-specific preview benefit* lasts for over 600 ms, and can survive considerable displacements of item location (Kahneman et al., 1992) and even the addition of a border that does not overlap existing components (Gao & Scholl, 2010).

This effect weakens as the number of items increases from 2 to 4, suggesting that some aspects of the process involved are extremely limited, possibly to one item at a time. The formation of the representations underlying this effect is believed to be largely complete by about 250 ms (Kahneman et al., 1992).

ii) Change detection / change blindness

If a change is made while the accompanying motion signals are swamped (e.g., by an image flicker or eye movement), it can be quite difficult to notice, even when large and continually repeated (Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1997). This *change blindness* is believed to reflect a failure to attentionally hold the item at the moment of change (Rensink et al., 1997; Rensink, 2000a). Although reminiscent of inattentional blindness (Section

1.2), the two are not the same: for example, change blindness can still occur when the observer is aware that a change could happen, suggesting that the underlying processes are less flexible, or at least, slower (see Rensink, 2009).

When detecting the presence of change, up to 3-5 items can be monitored at a time (Luck & Vogel, 1997; Pashler, 1988; Rensink, 2000b). Relatively little information about each is stored (Alvarez & Cavanagh, 2004), and not all structural levels are represented (Austen & Enns, 2000). Filtering is able to control the properties considered relevant (Rensink, 2000b). When detecting an *absence* or *conjunction* of changing properties, only one item can be compared at a time, at a rate of about 300 ms per item. Thus suggests that information from the monitored items is not maintained in a set of independent representations, but is collected into a single *nexus* (see Rensink, 2002).

iii) Third-order motion

Three different kinds of motion are often distinguished. *First-order* motion is supported by differences in luminance, and results from the operation of low-level detectors (Cavanagh & Mather, 1989; but see Petersik & Rice, 2006). S*econd-order* is encountered for patterns with no luminance differences (e.g., textures or isoluminant colors); it appears to be based at least in part on the displacement of low-level structures (Seiffert & Cavanagh, 1998). Finally, *third-order* motion involves the tracking of movement patterns via higher-level operations. It acts on extended spatial structures (e.g., a group of lines), and can survive changes in their underlying properties (Lu & Sperling, 2001).

The perception of third-order motion is believed to require at least one kind of attention (Cavanagh, 1992). This differs from the kind that enables character recognition (presumably binding), although it may draw upon it. The mechanism involved has a time constant of about 300 ms; only one movement pattern can be seen at a time (Ho, 1998).

iv) Attentional blink

If two pre-specified targets in a stream of briefly-presented stimuli appear at slightly different times, an observer will often see the first target but not the second. This phenomenon is called the *attentional blink (AB)*; it is thought to result from a failure to attentionally hold the second target, possibly because the representation of the first is not yet complete (Raymond, Arnell, & Shapiro, 1992). The mechanisms underlying this effect are selective. They differ from those that underlie repetition blindness, but may involve coherent representation of some kind (Chun, 1997). The meaning of a word not consciously experienced during the blink can be implicitly perceived (Luck, Vogel, and Shapiro, 1996).

The separation needed for AB is 250-300 ms, a value known as the *attentional dwell time* (Duncan, Ward, & Shapiro, 1994; Theeuwes, Godijn, & Pratt, 2004). Only one such representation appears to be created at a time.

v) Simultanagnosia

Simultanagnosia is an inability to recognize more than one object at a time, resulting from damage to the brain (Coslett & Saffran, 1991). Two variants are known. The first, *dorsal* simultanagnosia, results from bilateral damage to the junction between the parietal and occipital lobes. It involves a failure to see more than one object or part of an object, with the rest of the input simply not experienced (Farah, 2004). The meaning of unseen items, however, can still have an effect (Jackson, Shepherd, Mueller, Husain, & Jackson, 2006). The second variant, *ventral* simultanagnosia, is caused by damage to the left inferior temporo-occipital region. It also involves a failure to recognize more than one object, but allows perception—including counting—of several simple shapes simultaneously (Farah, 2004). Both variants are believed to stem from an impairment of attentional processing (Coslett & Saffran, 1991; Duncan et al., 2003), likely including the engagement or disengagement of attentional hold, and—for dorsal simultanagnosia—attentional gating as well.

3.3. Mechanism

The effects above involve a mechanism—or set of mechanisms—with many of the following characteristics: relatively slow (c. 300 ms), access to only a few structures and a small amount of information from each, ability to filter inputs via simple properties, and only one overall structure in play at a time. Such a mechanism could represent at most only a few aspects of a physical object at any time. But if the relevant quantities could be created when requested, the result would be a *virtual representation* of the object, with any relevant aspect represented whenever needed (Ballard, Hayhoe, Pook, & Rao, 1997; Rensink, 2000a).

One model of this is the *object file,* a temporary representation of (bound) properties intended to capture the continuity of an object as it transforms or changes position (Kahneman et al., 1992). Containing only a handful of features and a single level of hierarchical structure, it forms a highly compact description of the object (including the spatial relations among its parts). Only a few object files are assumed to exist simultaneously, with only one updated at a time.

A proposal similar in many ways is the *coherence field,* in which information is "held" in reverberating circuits implemented via feedforward and feedback connections between selected proto-objects and a higher-level nexus, with only a few properties in play at any moment (Rensink, 2000a). The selected proto-objects become part of a coherent visual object; when the circuit is broken, the object-level structure simply dissipates, without any accumulation of information in coherent form (also see Wolfe et al., 2000). Some information can be retained—although no longer in coherent form—for a small number of items (perhaps 10 or so), possibly as a memory to help guide subsequent allocation (Hollingworth & Henderson, 2002; Rensink, 2000a).

A more detailed quantitative model that handles many of the issues involved with the representation of coherent (and bound) structure is the *Visual Theory of Attention* (Bundesen, 1990). This incorporates several aspects of attentional processing in a unified way, and provides a good fit to a considerable body of experimental data (e.g., Duncan et al., 2003).

In all these proposals a critical role is played by visual short-term memory (vSTM), which provides continuity over durations greater than a few hundred milliseconds. Although the contents of vSTM play an important role in coherent representation, the two are not necessarily the same, at least conceptually: the contents of vSTM, for example, can be maintained for several seconds after visual input disappears, whereas coherence collapses within a few hundred milliseconds afterwards (Rensink, 2000a). Bindings between different dimensions appear to be maintained in vSTM (Gajewski & Brockmole, 2006; but see Wheeler & Treisman, 2002), whereas the status of these in coherent representation (where they are held only over short durations) is unknown. And while more sophisticated structures, such as movement patterns—or *sprites*—may be accessed by vSTM to enable the recognition and anticipation of complex movements (Cavanagh, Labianca, & Thornton, 2001), it is likewise unclear whether these are also part of a coherent representation. More generally, the heavy involvement—if not outright embedding—of vSTM in long-term memory (Cowan, 2001) might enable access to considerable knowledge about possible movements and other transformations, which could *guide* the creation of coherent representation. But such knowledge need not be contained in the representation itself. [Also see THEORIES OF ATTENTION.]

3.4. Relation to Consciousness

As in the case of other kinds of attention, conscious experience is likely involved with at least some aspects of attentional holding, in regards to both output and control of operation. Indeed, the likelihood of such involvement is highest for this kind of attention. Given that coherence is expensive in terms of computational resources, it is extremely important to guide its usage effectively, both when transmitting it to the appropriate processes and deciding when to create it in the first place. The more global the basis of this guidance, the better.

i) Conscious perception

Attentional holding appears necessary for a kind (or possibly level) of *coherent experience*, which includes not only ambient (and perhaps assembled) experience, but also the impression of a persisting substrate. This may be similar to the *object consciousness* proposed on the basis of verbal reports of structure, which is believed to involve vSTM (Iwasaki, 1993). In the absence of attentional hold, coherence does not exist; experience is of a succession of static structures, somewhat akin to perception under stroboscopic conditions.

Whether attentional hold is also sufficient to perceive a coherent object depends on what is meant by "object". If this refers to a physical object, hold is clearly insufficient: relatively little information is maintained in coherent form, and so most properties are not experienced in a coherent way. But if the term refers to a visual object—i.e., the contents of a coherent representation—it may be both necessary *and* sufficient (Rensink, 2002).

Two types of change can be distinguished on the basis of subjective impression (Rensink, 2002). *Dynamic* change is experienced as the transformation of a structure that is *continually present* throughout the process. In contrast, *completed* change is experienced as an event that has simply *happened* at some time in the past. These appear to involve different time constants: dynamic change cannot survive an interstimulus interval much beyond 300 ms, while completed change is far more robust; it also appears to be the only type experienced for changes made during eye movements (Rensink, 2002; also see Hollingworth, 2008). It may be that dynamic change involves coherent representation, and completed change, the contents of vSTM alone.

ii) Nonconscious perception

Although attentional hold may be necessary for the visual experience of change, implicit detection of change may still occur without it (Fernandez-Duque & Thornton, 2000; Laloyaux, Destrebecqz, & Cleeremans, 2006). The existence of this is controversial (Mitroff, Simons, & Franconeri, 2002). But if it does exist, it may reflect the operation of low-level feedback circuits that provide a degree of continuity over brief temporal gaps (cf. Di Lollo et al., 2000).

Another interesting phenomenon is the report of some observers of "sensing" or "feeling" a change without an accompanying visual experience. This may be a form of nonvisual—but still conscious—awareness involving mechanisms that differ to some extent from those underlying visual experience (Rensink, 2004). Such an account is controversial (Simons, Navarez, & Boot, 2005), but additional support for it has been found (Busch, Fründ, & Hermann, 2010; Galpin, Underwood, & Chapman, 2008).

4. Attentional Indexing

4.1. Function

It is sometimes important to individuate a physical object—to see it not just as *an* object (coherent or otherwise), but as a *particular* object. This is especially so when more than one object must be dealt with, for example, determining spatial relations such as "between", assessing the configuration of a group, or ensuring that items in an image are processed in an optimal—or at least, effective—sequence (Pylyshyn, 2003; Ullman, 1984).

Although this might be done via multiple coherent representations, it is difficult to create more than one such representation at any time (Section 3.3). But fortunately, simultaneous coherence is rarely needed: many operations involve only a few items, with each needing only to be distinguished in some way, such as by location. Consequently, an item only needs to contain enough information to be considered distinct at some point, and then kept track of as it moves about. If such *indexing* is successful, the item can be immediately accessed whenever needed.

4.2. Perceptual Effects

Although indexing can be useful, it also has limitations. As in the case of other kinds of attention, these need not all involve the same neural systems. But they nevertheless appear to exhibit a degree of similarity:

i) Multiple-object tracking

In *multiple-object tracking (MOT),* a set of identical items is displayed, a subset marked (e.g., briefly flashed), and the marked items tracked for several seconds as all items move around (Pylyshyn & Storm, 1988; Scholl, 2009). Accuracy is often good up to about 4 tracked items, this limit varying between 1 and 8, depending on task demands (Alvarez & Franconeri, 2007).

Accuracy also depends on eccentricity. It is limited by a resolution coarser than that due to purely sensory limitations, possibly reflecting the involvement of attentional gating (Intriligator & Cavanagh, 2001). Indeed, tracking may help with the control of gating (Bettencourt & Somers, 2009). But it does not improve change detection (Bahrami, 2003), suggesting that it cannot help control holding.

Although tracked items are often considered independent, they can be perceived as the corners of a virtual polygon that deforms over time (Yantis, 1992). Tracking is not extrapolated when an item becomes occluded (Franconeri, Pylyshyn, & Scholl, 2006), and does not appear to facilitate binding of the properties of the tracked item, or their encoding into vSTM (Bahrami, 2003; Scholl, 2009). Meanwhile, it is possible to track the center of a simple extended structure (Scholl, Pylyshyn, & Feldman, 2001), but not the intersection perceived to result from two moving structures (Anstis, 1990). This suggests that tracking is of structures rather than simple features, but that these structures are not highly integrated, perhaps being preattentive segments or proto-objects. Tracking is based on perceived rather than retinotopic space (Liu et al., 2005).

ii) Prioritization of search

A target item in visual search (Section 1.2) is more rapidly detected if in a set of items whose locations are shown ahead of time. This *prioritization of search* is believed to reflect the existence of a set of preferred locations (Yantis & Jones, 1991; but see Watson & Humphreys, 1997), each corresponding to an indexed item.

Up to 4 locations can be prioritized this way; the effect does not seem to be affected by the distance between them (Burkell & Pylyshyn, 1997; Yantis & Jones, 1991). Considerably more can be prioritized under some conditions, but this may simply reflect a group that is perceived as a single indexed item (Belopolsky, Theeuwes, & Kramer, 2005).

iii) Inhibition of return

When an item is attentionally gated, more time is needed to respond to it for a short while afterwards (Posner & Cohen, 1984). This *inhibition of return (IOR)* takes at least 200-300 ms to develop (depending somewhat on task demands), and can last several seconds. It appears only after the application of attentional operations involving space (Klein, 2000). It is believed to mark particular items or locations, so as to prevent attentional processes from returning too soon to a previously-visited input (see Klein, 2000).

At least four locations (items) can be inhibited simultaneously; distance of separation does not appear to matter (Snyder & Kingstone, 2001; Wright & Richard, 1996). IOR acts in an object-centered rather than a retinotopic frame (Tipper, Driver, and Weaver, 1991).

iv) Subitizing

Subitizing is the rapid, effortless, and accurate counting of a small number of items. Up to four can be enumerated this way, at a rate of about 50 ms per item; beyond this, a slower process is needed (Trick & Pylyshyn, 1994; but see Piazza, Mechelli, Butterworth, & Price, 2002). [Also see NUMERICAL COGNITION.] Individuation is clearly involved, so as to ensure that the subitizing is of distinct elements.

Arbitrary combinations of features cannot be subitized, although localized groups can (Trick & Enns, 1997), suggesting that the process acts on structures (e.g., proto-objects) formed prior to attentional binding (Pylyshyn, 2003; Scholl, 2009). There appears to be no effect of spatial separation, but performance is slower for items on the same segment (or bound structure), indicating that these items are not considered separate, at least initially (Davis & Holmes, 2005).

4.3. Mechanism

Each of the effects described above appears to involve a mechanism that can index about 4 items. Position is referenced not to an arbitrary point in retinotopic space, but to a structure—a segment, proto-object, or frame—created at early levels of processing. Selectivity can be allocated on the basis of global considerations, showing that this process is attentional (also see Scholl, 2009). But it operates rapidly and appears to be concerned with facilitating the control of other processes, perhaps explaining why it has not always been considered a form of attention (cf. Pylyshyn, 2003).

Coherent representations do not appear to be automatically indexed (van Dam & Hommel, 2010), suggesting that indexing may be carried out by mechanisms largely separate from those used for holding (also see Bahrami, 2003; Scholl, 2009). One proposed mechanism is the *fingers of instantiation* (*FINST)* model*,* which posits that 4-6 items can be simultaneously tracked (Pylyshyn, 2003; Pylyshyn & Storm, 1988). A variant is the *flexibly-allocated index* (*FLEX*) model, in which 1-8 items can be tracked, with a minimal separation between them (Alvarez & Franconeri, 2007). The memory involved in indexing may be related to a spatial short-term memory concerned exclusively with location (e.g., Darling, Della Sala, & Logie, 2007; Xu & Chun, 2006). [Also see THEORIES OF ATTENTION.]

4.4. Relation to Consciousness

At the subjective level, indexing requires effort, consistent with it being a form of attention (see Scholl, 2009). It also appears to involve—at least during tracking—an experience of continuity, one similar to that experienced during holding, where objects maintain coherence over both time and space.

Whether indexing is also associated with a distinct kind of perceptual experience (corresponding to individuation, say) is an open issue. Given that it largely acts by facilitating the control of other attentional processes, the kinds of experience associated with indexing may not extend beyond those associated with other kinds of attention. More precisely: although indexing may enable more effective control of attentional processing, its effects may be exhibited mostly—if not entirely—via the facilitation of mechanisms found elsewhere; no new kinds of output or operational control may be involved. As such, no new kinds of conscious experience would be expected. The definitive resolution of this issue will likely require additional work in which the effects of indexing can be separated out from those of the processes it facilitates.

Conclusion

Although the relation between perception and attention is complex, it is possible to clarify—and potentially, simplify—several of the issues involved. For visual perception, a *function-centered taxonomy* can distinguish at least five kinds of attention, each having a distinct function and characteristic mode of operation. Many of the behavioral effects associated with attention (including perceptual deficits) can be identified with one of these. In addition, at least three kinds of visual experience can be distinguished, involving increasingly sophisticated levels of represented structure, and each associated with a particular kind of attention. The results of many studies also indicate that processes of considerable sophistication occur both prior to and independent of any kind of attention, in the complete absence of conscious visual experience.

Future Directions

1. *What exactly is an attentional kind?* The taxonomy used here centers around function, with each function appearing to be associated with a distinct set of effects and mechanisms[2]. But how real is this? Does each kind of attention correspond to an aspect or stage of a unitary process, or to something more? To answer this would include determining whether all effects believed to reflect the same kind of attention actually involve the same mechanism; this in turn would require careful measurement of things such as timing parameters and variables involved. It would also include determining the extent to which each mechanism is also used for other functions (e.g., a mechanism for filtering that is also used for binding.) Some of these issues could be investigated via approaches such as the *attention operating characteristic (AOC),* which shows the degree to which the same mechanism (or "resource") is used in two concurrent tasks (Sperling & Dosher, 1985).

2. *What kinds of attention exist?* Over the years, various processes have been lumped together under the name of "attention" (Allport, 1993; Chun, Golumb, & Turk-Browne, 2011; Tsotsos, 2011). It is important to determine what these are in a rigorous way. The approach used here might be extended to other kinds of contingently selective process, e.g., assignment of semantic associations, or control of motor actions. Additional kinds of attention might also be found via the splitting of existing categories—e.g., binding separated into the linking of features and the formation of extended configurations across space. Progress here would include developing more rigorous criteria for determining what counts as a distinct kind of attentional process, as well as a more careful examination of effects currently considered to involve the same kind of attention.

3. *How do the various kinds of attention relate to each other?* This is essentially the generalization of the question of how eye movements relate to covert kinds of attention. A key step is to determine which attentional processes depend on which. Although there could in principle be interactions among several processes simultaneously, much could likely be done simply by investigating pairs of interactions. If there is some dependence between particular processes, it becomes important to determine how closely their operation is aligned, both in space (operands selected) and in time (kinds of delays).

4. *What are the characteristics of the processes operating prior to attention*? The kinds of attention discussed here all appear to operate on the same preattentive structures: extended segments, localized proto-objects, and possibly frames of some kind. It is important to get a better understanding of these. Do they form a common substrate for all attentional processes? Are they distinct, or different aspects of the same thing? To what extent can they be overridden—or at least, set—by attentional control allocated ahead of time? These questions can be investigated via modifications of the techniques outlined here.

5. *How do the different kinds of visual experience relate to the different kinds of attention?* Interest in how attention relates to conscious awareness spans centuries. Most proposals (e.g., Koch & Tsuchiya, 2007; Lamme, 2003) assume only one kind of attention and one kind of visual experience. However, given the distinctions that appear to be emerging, this issue needs to be recast into a more articulated form: relationships may only exist between some kinds of attention and some kinds of experience. Given the number of possible combinations and the number of issues to consider in each, establishing such relationships will take considerable work. But once determined, they will likely give us important new insights into the nature of both attention and consciousness.

Endnotes

[1] These have often been referred to as "objects". But this word is overworked, having been used to refer to virtually any level of structure. Following Driver et al. (2001), simple large-scale background structures at early levels are referred to here as "segments"; following Rensink & Enns (1995), more complex localized structures are referred to as "proto-objects". "Object" (as in "visual object") is reserved for the highest level of visual structure, viz., a coherent representation in which all components refer—over both space and time—to a single underlying substrate, ideally capturing the structure of a physical object in the world.

[2] The function-centered categorization developed here echoes to a large extent the approach advocated by Marr (1982) and Dennett (1994). This begins by analyzing the process in terms of its function, followed by analysis of the mechanisms/representations that carry it out, followed by analysis of the underlying neural implementation. Given the focus of this article, only the first two are discussed in the framework here.

Related Chapters

- DISORDERS OF ATTENTION
- EVENT PERCEPTION
- EXECUTIVE FUNCTIONS
- EYE MOVEMENTS
- OBJECT RECOGNITION
- ORGANIZING THE SENSORY INPUT
- SPATIAL ATTENTION
- THEORIES OF ATTENTION
- THE NATURE AND STATUS OF MENTAL RESOURCES

References

Albrecht, A.R., List, A., & Robertson, L.C., (2008). Attentional selection and the represention of holes and objects. *Journal of Vision*, *8:8*, 1-10.

Allport, A. (1989). Visual attention. In M.I. Posner (Ed.), *Foundations of Cognitive Science* (pp. 631-682). Cambridge, MA: MIT Press.

Allport, A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In D.E. Meyer & S. Kornblum (Eds.), *Attention And Performance XIV: Synergies In Experimental Psychology, Artificial Intelligence, And Cognitive Neuroscience* (pp. 183-218). Cambridge, MA: MIT Press.

Alvarez, G.A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. *Psychological Science*, *15*, 106-111.

Alvarez, G.A., & Franconeri, S. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7:14*, 1-10.

Anstis, S. (1990). ``Imperceptible intersections: The chopstick illusion". In A. Blake, T. Troscianko (Eds.), *AI and the Eye* (pp. 105-117). London: John Wiley.

Anton-Erxleben, K., Abrams, J., & Carrasco, M. (2010). Evaluating comparative and equality judgments in contrast perception: Attention alters appearance. *Journal of Vision*, *10:6,* 1–22.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*, 157-162.

Austen, E., & Enns, J.T. (2000). Change detection: Paying attention to detail. *Psyche*, *6* (11).

Baars, B.J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Prog. Brain Res*., *150*, 45–53.

Bahrami, B. (2003) Object property encoding and change blindness in multiple object tracking. *Visual Cognition*, *10*, 949–963.

Ballard, D.H., Hayhoe, M.M., Pook, P.K., & Rao, R.P.N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723-767.

Barlow, H.B. (1981). Critical Limiting Factors in the Design of the Eye and Visual Cortex. *Proc. R. Soc. Lond. B*, *212*, 1-34.

Barlow, H.B., Mollon, J.D. (Eds.). (1982). *The Senses*. Cambridge: University Press.

Bartolomeo, P., & Chokron, S. (2002). Orienting of attention in left unilateral neglect. *Neuroscience and Biobehavioral Reviews, 26*, 217-234.

Behrmann, M., & Kimchi, R. (2003). What does visual agnosia tell us about perceptual organization and its relationship to object perception? *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 19-42.

Belopolsky, A.V., Theeuwes, J., & Kramer, A. (2005). Prioritization by transients in visual search. *Psychonomic Bulletin & Review*, *12*, 93-99.

Bettencourt, K.C., & Somers, D.C. (2009). Effects of target enhancement and distractor suppression on multiple object tracking capacity. *Journal of Vision*, *9:9*, 1-11.

Bichot, N.P., Cave, K.R., & Pashler, H. (1999). Visual selection mediated by location: Feature-based selection of noncontiguous locations. *Perception & Psychophysics*, *61*, 403-423.

Bisiach, E. (1993). Mental representation in unilateral neglect and related disorders: The twentieth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *46*, 435-561.

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and Mind*, *2*, 5-38.

Block, N. (1995). On a confusion about a function of consciousness. *Behav. Brain Sci.***,** *18*, 227-247.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition*. Nature, 226, 177–178.*

Bravo, M.J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics*, *51*, 465-472.

Briand, K.A., & Klein, R.M. (1986). Is Posner's “Beam” the Same as Treisman's “Glue”? On the Relation Between Visual Orienting and Feature Integration Theory. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 228-241.

Broadbent, D.E. (1982). Task combination and selective intake of information. *Acta Psychologica, 50*, 253-290.

Braun, J., & Sagi, D. (1990). Vision outside the focus of attention. *Perception & Psychophysics*, *48*, 45-58.

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

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523–547.

Burkell, J.A., & Pylyshyn, Z. W. (1997). Searching through subsets: A test of the visual indexing hypothesis. *Spatial Vision*, *11*, 225-258.

Busch, N.A., Fründ, I, & Hermann, C.S, (2010). Electrophysiological evidence for different types of change detection and change blindness. *Journal of Cognitive Neuroscience*. *22*, 1852-1869.

Carpenter, R.H.S. (1988). *Movements of the Eyes* (2nd rev. & enlarged ed.). London: Pion.

Carrasco, M., Evert, D.L., Change, I., & Katz, S.M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, *57*, 1241-1261.

Carrasco, M., & Frieder, K.S. (1997). Cortical magnification neutralized the eccentricity effect in visual search. *Vision Research*, *37*, 63-82.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*, 308–313.

Cartwright-Finch, U., & Lavie, N. (2007). The role of perceptual load in inattentional blindness. *Cognition*, *102*, 321-340.

Cavanagh, P. (1992). Attention-based motion perception. *Science*, *257*, 1563–1565.

Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: Sprites. *Cognition*, *80*, 47-60.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4,* 103-129.

Cave, K.R. (1999). The FeatureGate model of visual selection**.** *Psychological Research*, *62*, 182-194.

Cave, K.R., & Bichot, N.P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin &* Review, 6, 204-223.

Chong, S.C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*, 393-404.

Chun, M.M. (1997). Types and tokens in visual processing: A double dissociation between the attentional blink and repetition blindness. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 738-755.

Chun, M.M., Golomb, J.D., & Turk-Browne, N.B. (2011). A Taxonomy of External and Internal Attention. *Annual Review of Psychology*, *62*, 73-101.

Cohen, A., Ivry, R. (1989). Illusory Conjunctions Inside and Outside the Focus of Attention. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 650-663.

Cohen, M.A., Alvarez, G.A, & Nakayama, K. (2011). Natural-Scene Perception Requires Attention**.** *Psychological Science, 22*, 1165-1172.

Cohen, M.A., & Dennett, D.C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, *15*, 358-364.

Coslett, H.B., & Saffran, E. (1991). Simultanagnosia: To see but not two see. *Brain*, *114*, 1523-1545.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-18.

Crowne, D.P. (1983). The frontal eye field and attention. *Psychological Bulletin*, *93*, 232-260.

Darling, S., Della Sala, S., & Logie, R.H. (2007). Behavioural evidence for separating components within visuo-spatial working memory. *Cognitive Processing, 8*, 175-181.

Davis, E.T., Kramer, P., & Graham, N. (1983). Uncertainty about spatial frequency, spatial position, or contrast of visual patterns. *Perception & Psychophysics*, *33*, 20-28.

Davis, G., & Holmes, A. (2005). What is enumerated by subitization mechanisms? *Perception & Psychophysics, 67,* 1229-1241.

Dennett, D.C. (1994). Cognitive science as reverse engineering: Several meanings of "top-down" and "bottom-up". In D. Prawitz, B. Skyrms, D. Westerstähl (Eds.), *Logic, Methodology,a nd Philosophy of Science IX* (pp. 679-689). Amsterdam: Elsevier.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193-222.

Deubel, H., & Schneider, W.X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36,* 1827-1837.

Di Lollo, V., Enns, J.T., & Rensink, R.A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*, 481-507.

Driver, J., & Baylis, G.C. (1989). Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 448-456.

Driver, J., Davis, G., Russell, C., Turrato, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition, 80*, 61-95.

Driver, J., & Halligan, P.W. (1991). Can Visual Neglect Operate in Object-centred Co-ordinates? An Affirmative Single-case Study. *Cognitive Neuropsychology*, *8*, 475-496.

Duncan, J. (1984). Selective Attention and the Organization of Visual Information. *Journal of Experimental Psychology: General*, *113*, 501-517.

Duncan, J., & Humphreys, G.W. (1989). Visual Search and Stimulus Similarity. *Psychological Review*, *96*, 433-458.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature, 369*, 313-315.

Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Ward, R., Kyllingsbæk, S., van Raamsdonk, M. Rorden, C., & Chavda, S. (2003). Attentional functions in dorsal and ventral simultanagnosia*, Cognitive Neuropsychology*, *20*, 675-701.

Eckstein, M.P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science, 9,* 111-118.

Egeth, H.E., Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48,* 269-297.

Egly, R., Driver, J., & Rafal, R.D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, *123*, 161-177.

Eriksen, B.A., & Eriksen, C.W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Perception and Psychophysics***,** *16,* 143–149.

Eriksen, C.W., & St. James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40,* 225-240.

Eriksen, C.W., & Yeh, Y.-Y. (1985). Allocation of attention in the visual field*.* *Journal of Experimental Psychology: Human Perception and Performance*, *11,* 583-597.

Farah, M.J. (2004). *Visual Agnosia* (2nd ed.). Cambridge MA: MIT Press.

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, 323-344.

Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. *Journal of Neurophysiology, 106*, 1389-1398.

Franconeri, S.L., Bemis, D.K., & Alvarez, G.A. (2009). Number estimation relies on a set of segmented objects. *Cognition*, *113,* 1-13.

Franconeri, S.L., Pylyshyn, Z.W., & Scholl, B.J. (2006). Spatiotemporal cues for tracking multiple objects through occlusion. *Visual Cognition*, *14,* 100-103.

Gajewski, D.A., & Brockmole, J.R. (2006). Feature bindings endure without attention: Evidence from an explicit recall task. *Psychonomic Bulletin & Review*, *13,* 581-587.

Galpin, A., Underwood, G., & Chapman, P. (2008). Sensing without seeing in comparative visual search. *Consciousness and Cognition*, *17,* 672-687.

Gao, T., & Scholl, B.J. (2010). Are objects required for object-files? Roles of segmentation and spatiotemporal continuity in computing object persistence. *Visual Cognition*, *18,* 82-109.

Giordano, A.M., McElree, B., & Carrasco, M. (2009). On the automaticity and flexibility of covert attention: A speed-accuracy trade-off analysis. *Journal of Vision*, *9:30,* 1-10.

Goodale, M.A., & Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15, 20*-22.

Harris, C.R., & Pashler, H. (2004). Attention and the processing of emotional words and names: Not so special after all. *Psychological Science*, *15,* 171-178.

Hatfield, G. (1998). Attention in early scientific psychology. In R.D. Wright (Ed.), *Visual Attention* (pp. 3-25). Oxford: University Press.

Hayhoe, M. M., & Ballard, D. H. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9,* 188–194.

He, Z.J., & Nakayama, K. Visual attention to surfaces in 3-D space. (1995). *Proceedings of the National Academy of Sciences USA, 92*, 11155-11159.

Henderson, J.M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. Eye guidance in reading and scene perception. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (pp. 269-293). Oxford: Elsevier.

Hirst, W., & Kalmar, D. (1987). Characterizing attentional resources. *Journal of Experimental Psychology: General*, *116*, 68-81.

Ho, C.E. (1998). Letter recognition reveals pathways of second-order and third-order motion. *Proceedings of the National Academy of Sciences USA*, *95*, 400-404.

Hollingworth, A. (2008). Visual memory for natural scenes. In S. J. Luck & A. Hollingworth (Eds.), *Visual Memory* (pp. 123-162). New York: Oxford University Press.

Hollingworth, A., & Henderson, J.M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 113-136.

Humphreys, G.W. (1981). Flexibility of attention between stimulus dimensions. *Perception & Psychophysics*, *30*, 291-302.

Intriligator, J., & Cavanagh, P. (2001) The spatial resolution of visual attention. *Cognitive Psychology*, *43*, 171–216.

Iwasaki, S. (1993). Spatial attention and two modes of consciousness. *Cognition*, *49*, 211-233.

Jackson, G.M., Shepherd, T., Mueller, S.C., Husain, M., & Jackson, S.R. (2006) Dorsal simultanagnosia: an impairment of visual processing or visual awareness? *Cortex*, *42*, 740-749.

Jiang, Y., Costello, P., & He, S. (2007). Processing of invisible stimuli: advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological Science*, *18*, 349-55.

Juan, C.-H., Shorter-Jacobi, S.M., & Schall, J. (2004). Dissociation of spatial attention and saccade preparation. *Proceedings of the National Academy of Sciences USA*, *101, 15541-15544.*

Kahneman, D., Treisman, A.**, &** Gibbs, B.J. (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.

Kanwisher, N., Driver, J., & Machado, L. (1995). Spatial repetition blindness is modulated by selective attention to color or shape. *Cognitive Psychology*, *29*, 303-337.

Kanwisher, N., Yin, C., & Wojciulik, E. (1999). Repetition blindness for pictures: Evidence for the rapid computation of abstract visual descriptions. In V. Coltheart (Ed.), *Fleeting Memories*: *Cognition Of Brief Visual Stimuli* (119-150). Cambridge MA: MIT Press.

Kaplan, E. (1991). The receptive field structure of retinal ganglion cells in cat and monkey. In . A.G. Leventhal (Ed.), *Vision and Visual Dysfunction* (pp. 10–40). Boston, MA: CRC Press.

Kaunitz, L., Fracasso, A., & Melcher, D. (2011). Unseen complex motion is modulated by attention and generates a visible aftereffect. *Journal of Vision,11:10,* 1–9.

Kimchi, R. (2009). Perceptual organization and visual attention. *Prog Brain Res*, 176, 15-33.

Kimchi, R., & Peterson, M.A. (2008). Figure-ground segmentation can occur without attention. *Psychological Science, 19*, 660-668.

Klein, R.M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4***,** 138-147.

Kleiss, J.A., & Lane, D.M. (1986). Locus and persistence of capacity limitations in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 200-210.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*, 16-22.

Kramer, A. F., & Hahn, S. (1995). Splitting the beam: Distribution of attention over noncontiguous regions of the visual field. *Psychological Science*, *6,* 381-386.

Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception & Psychophysics, 50,* 267-284.

LaBerge, D. (1998). Attentional emphasis in visual orienting and resolving. In R.D. Wright (Ed.), *Visual Attention* (pp. 417-454). Oxford: University Press.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96,* 101-124.

Laloyaux, C., Destrebecqz, A., & Cleeremans, A. (2006). Implicit change identification: A replication of Fernandez-Duque and Thornton (2003). *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1366-1379.

Lamme, V.A.F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, *15*, 1605-1615.

Lamme, V.A.F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, ***7,*** 12-18.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 451-468.

Li, F.F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences USA*, *99*, 9596-9601.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*, 529–66.

Liu, G., Austen, E.L., Booth, K.S., Fisher, B.D., Argue, R., Rempel, M.I., & Enns, J.T. (2005). Multiple-object tracking is based on scene, not retinal, coordinates. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 235-247.

Lu, Z.-L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: review and update. *J. Optical Society of America A*, *18*, 2331-2370.

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

Luck S.J., Vogel E.K., & Shapiro K.L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, *383*, 616-618.

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

Maddox, W.T., Ashby, F.G, & Waldron, E.M. (2002). Multiple attention systems in perceptual categorization. *Memory & Cognition*, *30*, 325-329.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (pp. 8-38). San Francisco: W.H. Freeman.

Maunsell, J.H.R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*, 317-322.

Mitroff, S.R., Simons, D.J., & Franconeri, S.L. (2002). The siren song of implicit change detection. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 798-815.

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.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*, 782-784.

Most, S.B., Simons, D.J., Scholl, B.J., Jimenez, R., Clifford, E., & Chabris, C.F. (2001). How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science*, *12*, 9-17.

Naccache, L. Blandin, E., & Dehaene, S. (2002). Unconscious Masked Priming Depends on Temporal Attention. *Psychological Science, 13*, 416-424.

Nakayama, K. (1990). The iconic bottleneck and the tenuous link between early visual processing and perception. In C. Blakemore (Ed.), *Vision: Coding and Efficiency* (pp. 411-422). New York: Cambridge University Press.

Nakayama K, & MacKeben M. (1989) Sustained and transient components of focal visual attention. Vision Research, *29*, 1631-1647.

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

Niebur, E., & Koch, C. (1998). Computational architectures for attention. In Parasuraman, R. (Ed.), *The Attentive Brain* (pp. 163-186). Cambridge MA: MIT Press.

Noton, D., & Stark, L. (1971). Eye movements and visual perception. *Scientific American*, *224*(6), 34-43.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, *155*, 23-36.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*, 520-527.

Olshausen, B.A., Anderson, C.H., & Van Essen, D.C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, *13*, 4700-4719.

O'Regan, J.K., & Noë, A. (2001.) A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*, 939-73.

Palmer, J., Ames, C.T., & Lindsey, D.T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 108-130.

Parasuraman, R., (Ed.). (1998). *The Attentive Brain*. Cambridge MA: MIT Press.

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

Pashler, H.E. (1999). *The Psychology of Attention*. Cambridge MA: MIT Press.

Pelli, D. G., Cavanagh, P., Desimone, R., Tjan, B., & Treisman, A. (2007). Crowding: Including illusory conjunctions, surround suppression, and attention. *Journal of Vision*, *7:i,* 1.

Petersik, T.J, & Rice, C.M. (2006). The evolution of explanations of a perceptual phenomenon: A case history using the Ternus effect. *Perception*, *35*, 807-821.

Piazza, M., Mechelli A., Butterworth, B., & Price, C.J. (2002). Are Subitizing and Counting Implemented as Separate or Functionally Overlapping Processes? *NeuroImage, 15***,** 435-446.

Posner, M.I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma, & D. Bouwhuis (Eds.)., *Attention and Performance X* (pp. 531-556). Hillsdale, NJ: Erlbaum.

Posner, M.I., & Snyder, C.R.R. (1975). Attention and cognitive control. In R.L. Solso (Ed.) *Information Processing and Cognition* (pp. 55-85). Hillsdale, NJ: Erlbaum.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*, 160-174.

Posner, M.I., Walker, J.A., Friedrich, F.J., & Rafal, R.D. (1984). Effects of parietal injury on covert orienting of attention. *Journal of Neuroscience, 4,* 1863-1874.

Pylyshyn, Z.W. (2003). *Seeing and Visualizing: It's Not What You Think*. (201-279). Cambridge MA: MIT Press.

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

Prinzmetal, W. (1995). Visual feature integration in a world of objects. *Current Directions in Psychological Science, 4,* 90-94.

Prinzmetal, W., Henderson, D., & Ivry, R. (1995). Loosening the constraints on illusory conjunctions: The role of exposure duration and attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1362-1375.

Rafal, R. (1994). Neglect. *Current Opinion in Neurobiology*, *4,* 231-236.

Ramachandran, V.S. (1992). Filling in gaps in perception: I. *Current Directions in Psychological Science*, *1,* 199-205.

Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature*, *350*, 699-702.

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

Rensink RA. (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. (2002). Change detection. *Annual Review of Psychology*, *53,* 245-277.

Rensink, R.A. (2004). Visual sensing without seeing. Psychological Science, *15,* 27-32.

Rensink, R.A. (2005). Robust inattentional blindness. *Journal of Vision, 5:* 790a.

Rensink, R.A. (2009). Attention: Change blindness and inattentional blindness. In W. Banks (Ed). *Encyclopedia of Consciousness,* Vol 1 (pp. 47-59). New York: Elsevier.

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

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

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

Riddoch, M.J. Humphreys, G.W. (1987). A case of integrative visual agnosia. *Brain*, *110*, 1431-1462.

Robertson, L.C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience*, *4*, 93-102.

Robertson, L.C., & Kim, M.-S. (1999). Effects of perceived space on spatial attention. *Psychological Science*, *10,* 76-79.

Rossi, A.F., & Paradiso, M. (1995). Feature-specific effects of selective visual attention. *Vision Research*, *35,* 621–634.

Sagi, D., & Julesz, B. (1986). Enhanced detection in the aperture of focal attention during simple discrimination tasks. *Nature*, *321,* 693-695.

Scholl, B.J. (2009). What have we learned about attention from multiple-object tracking (and vice versa)? In D. Dedrick & L. Trick (Eds.), *Computation, Cognition, and Pylyshyn* (pp. 49-77). Cambridge MA: MIT Press.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple-object tracking. *Cognition, 80*(1/2), 159–177.

Seiffert, A.E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli. *Vision Research*, *38,* 3569-3582.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception*, *28,* 1059-1074.

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

Simons, D.J., Nevarez, G., Boot, W.R. (2005). Visual sensing *is* seeing: Why "mindsight," in hindsight, is blind. *Psychological Science*, *16,* 520-524.

Sincich, J.C., & Horton, J.C. (2005). The circuitry of V1 and V2: Integration of color, form, and motion. *Annual Review of Neuroscience*, *28,* 303-326.

Singer, W. (1999). Neuronal Synchrony: A Versatile Code for the Definition of Relations? *Neuron*, *24,* 49–65.

Snyder, J.J., & Kingstone, A. (2000). Inhibition of return and visual search: How many separate loci are inhibited? *Perception & Psychophysics*, *62,* 452-458.

Sperling, G., & Dosher, B.A. (1986). Strategy and optimization in human information processing. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of Perception and Human Performance, Vol 1.* (pp. 1-65). New York: Wiley.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, *5,* 1024–1037.

Steinman, B.A., Steinman, S.B., & Lehmkuhle, S. (1995). Visual attention mechanisms show a center—surround organization. *Vision Research, 35,* 1859-1869.

Theeuwes, J., Godijn, R., & Pratt, J. (2004). A new estimation of the duration of attentional dwell time. *Psychonomic Bulletin & Review*, *11,* 60-64.

Thornton, T.L., & Gilden, D.L. (2007). Parallel and serial processes in visual search. *Psychological Review, 114,* 71-103.

Tipper, S.P., Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *43,* 289-298.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, *255*(Nov), 114B-125.

Treisman, A. (1999). Solutions to the binding problem: Progress through controversy and convergence. *Neuron*, **24**, 105-110.

Treisman, A.M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12,* 97-136.

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

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 459-478.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14,* 107-141.

Trick, L.M., & Enns, J.T. (1997). Clusters precede shapes in perceptual organization. *Psychological Science*, *8,* 124-129.

Trick, L.M., & Pylyshyn, Z.W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*, 80-102.

Tsal, Y. (1983). Movement of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 523-530.

Tsotsos, J.K. (1990). Analyzing vision at the complexity level. *Behavioral and Brain Sciences*, *13,* 423-469.

Tsotsos, J.K. (2011). *A Computational Perspective on Visual Attention*. Cambridge, MA: MIT Press.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8,* 1096 - 1101.

Tsuchiya, N., Koch, C., Gilroy, L.A., & Blake, R. (2006) Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry *Journal of Vision, 6,* 1068–1078.

Ullman, S. (1984). Visual routines. *Cognition, 18,* 97-159.

van Dam, W.O., & Hommel, B. (2010). How object-specific are object files? Evidence for integration by location. *Journal of Experimental Psychology: Human Perception and Performance*, *36,* 1184-1192.

VanRullen, R., Reddy, L., & Koch, C. (2004). Visual search and dual tasks reveal two distinct attentional resources. *Journal of Cognitive Neuroscience*, *16,* 4-14.

Vecera, S.P., Flevaris, A.V., & Filapek, J.C. (2004). Exogenous spatial attention influences figure-ground assignment. *Psychological Science*, *15,* 20-26.

Watson, D.G., & Humphreys, G.W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review*, *104,* 90-122.

Wheeler, M.E., & Treisman, A.M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131,* 48-64.

Wickens, C.D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and Performance VIII* (pp. 239-257). Hillsdale, NJ: Erlbaum.

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

Wolfe, J.M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*, 33-39.

Wolfe, J.M. (1999). Inattentional amnesia. In Coltheart, V. (Ed.), *Fleeting Memories: Cognition of Brief Visual Stimuli* (pp. 71-94). Cambridge, MA: MIT Press.

Wolfe, J.M., Klempen, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 693-716.

Wright, R.D. (Ed.). (1998). *Visual Attention*, Oxford: University Press.

Wright, R.D., & Richard, C.M. (1996). Inhibition-of-return at multiple locations in visual space. *Canadian Journal of Experimental Psychology*, *50*, 324-327.

Wright, R.D., & Ward, L.M. (2008). *Orienting of Attention*. Oxford: University Press.

Xu, Y., Chun, M.M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*, 91-95.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*, 295-340.

Yantis, S., & Jones, E. (1991). Mechanisms of attentional selection: Temporally modulated priority tags. *Perception & Psychophysics*, *50*, 166-178.

Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *Journal of Vision, 10:16,* 1-12.

Further Reading

Coltheart, V. (Ed.). (1999). *Fleeting Memories: Cognition of Brief Visual Stimuli*. Cambridge, MA: MIT Press

Itti, L., Rees, G., Tsotsos, J. (2005). *The Neurobiology of Attention*. Burlington MA: Academic.

Jensen, M.S., Yao, R., Street, W.N., Simons, D.J. (2011). Change blindness and inattentional blindness. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, pp. 529–546.

Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Englewood, CO: Roberts & Co.

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

Parasuraman, R. (Ed.). (1998). *The Attentive Brain*. Cambridge MA: MIT Press.

Pashler, H.E. (1999). *The Psychology of Attention*. Cambridge MA: MIT Press.

Rensink, R.A. (2002). Change detection. *Annual Review of Psychology*, 53, 245-277.

Scholl, B.J. (2009). What have we learned about attention from multiple-object tracking (and vice versa)? In D. Dedrick & L. Trick (Eds.), *Computation, Cognition, and Pylyshyn* (pp. 49-77). Cambridge MA: MIT Press.

Wright, R.D. (Ed.). (1998). *Visual Attention*, Oxford: University Press.

Abstract

The relation between perception and attention is discussed, via a taxonomy of attention centered around function. Five kinds are distinguished. The first is sampling, involving the pickup of information by the eye. Second is filtering (or gating), the control of information considered relevant. The third is binding, involving the formation of more integrated structure over space. Fourth is holding, which creates the coherent structure necessary to perceive continuity over time. Fifth is indexing, which enables individuation of selected items. The function of each is discussed, along with several associated behavioral effects and perceptual deficits, the mechanisms believed to carry it out, and possible relation to conscious visual experience. Discussion is also given of the implicit processing carried out both prior to and independent of any kind of attention, in the complete absence of conscious visual experience.

Key words

attention, awareness, binding, change blindness, consciousness, implicit perception, inattentional blindness, selection, taxonomy, tracking