



Annual Review of Vision Science

Visual Search: How Do We Find What We Are Looking For?

Jeremy M. Wolfe^{1,2,3}

¹Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts 02115, USA; email: jwolfe@bwh.harvard.edu

²Department of Radiology, Harvard Medical School, Boston, Massachusetts 02115, USA

³Visual Attention Lab, Brigham & Women's Hospital, Cambridge, Massachusetts 02139, USA

Annu. Rev. Vis. Sci. 2020. 6:2.1–2.24

The *Annual Review of Vision Science* is online at vision.annualreviews.org

<https://doi.org/10.1146/annurev-vision-091718-015048>

Copyright © 2020 by Annual Reviews.
All rights reserved

Keywords

visual attention, visual search, foraging, working memory, serial processing, parallel processing

Abstract

In visual search tasks, observers look for targets among distractors. In the lab, this often takes the form of multiple searches for a simple shape that may or may not be present among other items scattered at random on a computer screen (e.g., Find a red T among other letters that are either black or red.). In the real world, observers may search for multiple classes of target in complex scenes that occur only once (e.g., As I emerge from the subway, can I find lunch, my friend, and a street sign in the scene before me?). This article reviews work on how search is guided intelligently. I ask how serial and parallel processes collaborate in visual search, describe the distinction between search templates in working memory and target templates in long-term memory, and consider how searches are terminated.



1. INTRODUCTION

The crowd at Reykjavik airport was impressive, and somewhere in the view in front of me might be my wife. Determining whether or not she was in my current field of view was a visual search task. If we had an all-seeing eye, then search would be unnecessary. The image contained the information; either my wife was present or she was not. However, while we can see something at all locations in the visual field, it should be obvious that we cannot immediately identify everything in a scene. Our capacity is limited in a manner that makes search necessary. Some of those limits have to do with the structure of the visual system. For instance, the retina and visual nervous system are structured so that we have high-resolution vision only in the small foveal region at the point of fixation. Beyond those basic visual limitations, there are capacity limits that are computational. Even a limited version of an all-seeing eye that could process everything, everywhere in the visual field, would require an implausibly large brain (Tsotsos 2011, improving on Neisser 1967).

A substantial body of current research strives to understand how human searchers cope with these limits on the ability to perceive the world and to find what they are looking for. Classic laboratory tasks have asked observers to search for a target in an array of distractors scattered at random on an otherwise blank computer screen. Those tasks have taught us a great deal, although one theme in this review is the need to begin using tasks that are closer to those in real-world search.

This review concentrates on four topics: how visual search is guided, how serial and parallel processes collaborate in visual search, search templates and working memory (WM), and search termination and tasks beyond laboratory search.

2. HOW IS VISUAL SEARCH GUIDED?

2.1. Classic Visual Search

Tasks relying on classic search for a target among some number of distractors have been popular in vision research since the 1970s (Egeth 1977, Wolfe 1998a). These tasks eliminate the complications produced by real scenes. Much of classic visual search research has been done with stimuli that are deliberately made large enough and spaced widely enough that acuity and crowding limits do not constrain the task. However, these tasks can still show clear evidence for capacity limits and the need to search. An example, taken from Wolfe & Myers (2010), is shown in **Figure 1a,b**. Observers are looking for plastic among stone surfaces. Display set sizes of only one, two, three, or four large patches were presented to observers, each in a different quadrant of the visual field. These surface materials are easily recognized. Material properties turn out to be identifiable in a short fraction of a second (Sharan et al. 2014). Nevertheless, the response time (RT; equivalently referred to as the reaction time) increases roughly linearly with the number of patches shown. Moreover, the slope of the RT \times set size function is approximately twice as great when the target is absent than when it is present. If capacity is not limited, then the time required to make a response should be nearly independent of the number of items in the search array. This is what happens for simple feature searches, as is shown in **Figure 1c,d** (Wolfe et al. 2010b). Note that the set sizes are much larger in this case. Nevertheless, the ability to find a red target among green distractors does not depend on the number of green items (but see Lleras et al. 2020).

The slope of the RT \times set size function is a measure of the rate of processing of items. Consider the search shown in **Figure 1a,b**, in which RT increases by 35 ms for each additional item on target-present trials and by 81 ms/item on target-absent trials. One way to think about such searches is to conceive of them as serial, self-terminating searches (Sternberg 1966). Each item is examined in turn until the target is found or the search is abandoned. On average, such a search



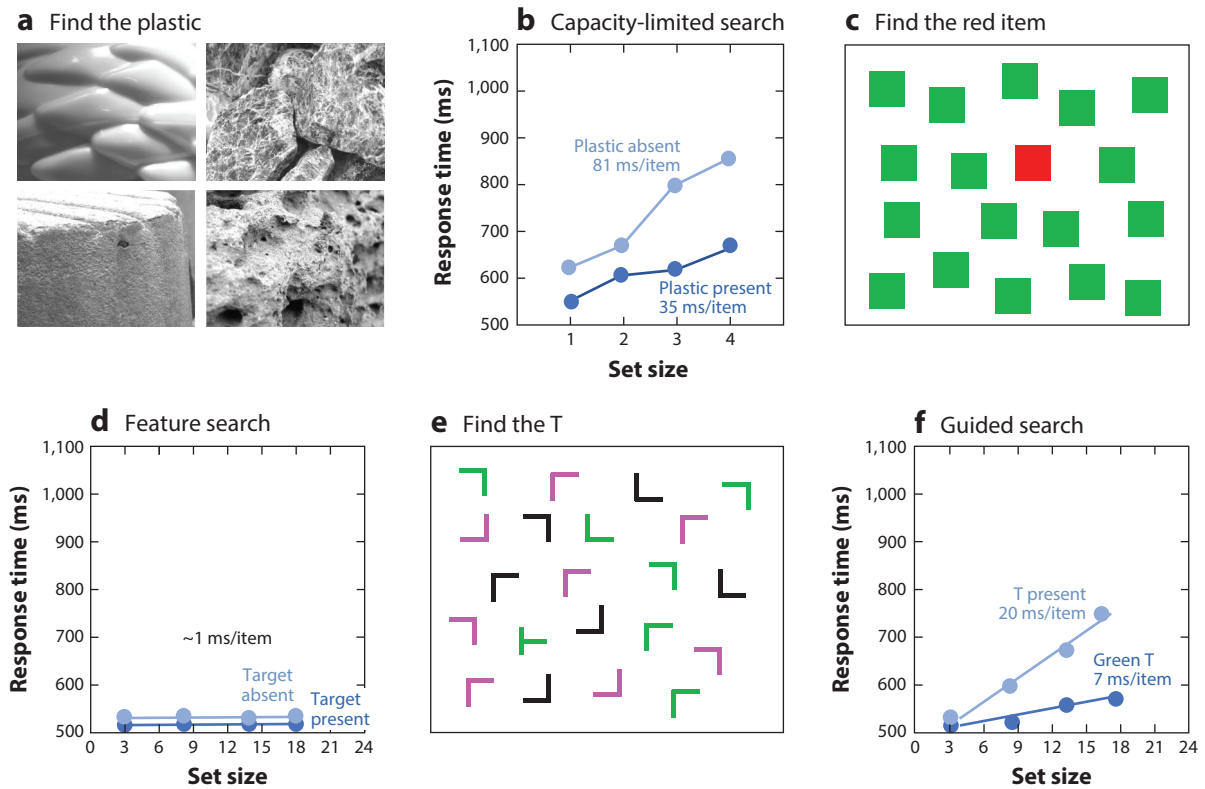


Figure 1

Classic visual search experiments. (a) Identifying plastic is easy even if defining plastic is hard. (b) Search for plastic requires an apparently serial search (data taken from Wolfe & Meyers 2010). (c) The red item pops out among green distractors. (d) In a search task, response time would be roughly independent of the number of green items in the display (data taken from Wolfe et al. 2010b). (e) Search for the T is more efficient if you know that the T is green. (f) This guidance by color is seen in the shallower slopes of the response time \times set size function for search for green Ts, as opposed to search for any T (hypothetical data; see Egeth et al. 1984).

will end after $(N + 1)/2$ items have been examined on target-present trials. If the target-absent trials were truly exhaustive, then all N items would need to be examined exactly once (search termination is more complex than this and is discussed below) (Chun & Wolfe 1996, Cousineau & Shiffrin 2004, Moran et al. 2013, Wolfe 2012). In this view, the target-absent slopes will be approximately twice the target-present slopes. The true cost per item would be twice the target-present slope. For the data shown, a simple serial model would argue that observers were processing items at a rate of $1/(0.035 \times 2) = \sim 14$ items per second. The same analysis would propose that all of the colored squares were being processed simultaneously (given slopes near 0 ms/item). Thus, the processing could be said to occur in parallel. Note that slopes would be much steeper if it were necessary to fixate on each item in turn. Voluntary saccadic eye movements occur at a rate of about 3–4 per second, translating to $RT \times$ set size slopes of 250–333 ms/item for target-absent trials and approximately half that for target-present trials.

In most real-world search tasks, initial parallel processing of the search display can be used to guide subsequent deployment of attention. For example, in **Figure 1e**, if you search for the letter T, then you will need to attend to each item until you stumble on the target. However, if you know that the T is green, then it should be intuitively clear that you will only (or, at least,



preferentially) attend to the green items (Egeth et al. 1984). Your search for a letter will be guided by the orthogonal information about its color. As a consequence, although there are 21 items in the display, the effective set size (Neider & Zelinsky 2008) will be 7, the number of green items. The slope of the $RT \times$ set size function for a search for the green T will be approximately one-third of the slope of the function when color is unknown (**Figure 1f**). The idea of guidance was present in the attention literature from at least the 1950s (Egeth et al. 1984, Green & Anderson 1956, Hoffman 1979). It was at the heart of the guided search (GS) model of Wolfe et al. (1989) that argued that differences in the efficiency of search tasks could be explained by differences in the amount of guidance; more guidance would lead to more efficient search. Multiple versions of GS have been developed (Wolfe 1994, 2007; Wolfe & Gancarz 1996; Wolfe et al. 1989, 2015), with the GS 2.0 (GS2) version being the best known. The idea of guidance is now generally accepted and is a part of virtually all models of search. Thus, taking the example in **Figure 1e**, it is hard to imagine a model that would not devote more processing resources to green items if the target was known to be green.

2.2. Guidance and the Nature of Guiding Features

What guides attention in visual search? Our understanding of how attention is guided has evolved. Classic GS and other models concerned themselves with guidance by basic features like color and motion. Attention could be guided to features of the target and, to a lesser extent, away from features of the distractors (Gaspelin & Vecera 2019). Other forms of guidance are discussed below. In this section, I briefly review feature guidance.

A limited set of features are available to guide attention. These include uncontroversial properties like color, size, and motion, as well as less obvious properties like lighting direction (Adams 2008, Enns & Rensink 1990) or axis of rotation (Schill et al. 2020). The list of the one to two dozen guiding features is reviewed elsewhere (Wolfe 2014; Wolfe & Horowitz 2004, 2017). In this review, I focus on what we have learned about feature guidance in recent years. Feature guidance is derived from the basic sensory input that gives rise to the perception of attributes like color or orientation. However, guidance appears to be a specific abstraction from that input [below, I argue that this abstraction is the search template (Olivers et al. 2011)]. Thus, for example, small differences in orientation (e.g., 0° from 5°) may be easily discriminated if items are scrutinized, but a 0° target will not pop out in a display of 5° distractors. The orientation difference would need to be larger before it would summon attention (Foster & Ward 1991). Moreover, guiding signals appear to be categorical. For example, it is easier to search for oriented lines if they are uniquely steep or shallow (Wolfe et al. 1992). More recently, other aspects of feature guidance have been described. In a series of experiments, Stefanie Becker (2010; Becker et al. 2020) demonstrated that guidance is often usefully described as a relationship between targets and distractors, rather than as an absolute target value. Thus, observers might be more likely to define a target as the reddest item in a display, rather than as a specific red color. The orange target could be defined as the reddest item in one context but might be searched for as the yellowest item in another context. Another way to think about this is to imagine a boundary in feature space that puts targets on one side and distractors on the other. When targets and distractors are linearly separable in this way, search is efficient. When targets are flanked in the feature space by distractors (e.g., 0° among $+/-20^\circ$ distractors), search is inefficient (Bauer et al. 1996, Wolfe et al. 1992). When a target is close to distractors in feature space (e.g., a bluish green among greenish-blue distractors), attention will be guided by an internal representation or target template that is not centered on the actual target feature. Instead, the template will be centered a bit further from the distractors to avoid mistakenly categorizing a distractor as a target (Yu & Geng 2019). In signal detection terms,

2.4 Wolfe



Review in Advance first posted on April 22, 2020. (Changes may still occur before final publication.)

this is roughly like moving the decision criterion to a more conservative position to avoid false alarm (false positive) errors. The precise nature of the feature guidance is probably hidden from the searchers. They look for ripe blueberries without necessarily realizing that this involves setting a criterion in color space for items that will attract attention and then using another criterion to decide about the precise status of an attended berry.

Guidance by shape is, perhaps, the greatest puzzle in feature guidance. There are several basic features that are related to the shape of an object. These include closure, line termination, curvature, topology (holes), and aspect ratio (for a review, see Wolfe 2018). Unlike features like color and motion, it is not clear if it makes sense to talk about these aspects of shape as independent features. A new way forward might arise out of the remarkable progress in training deep learning algorithms to identify objects (e.g., Kriegeskorte 2015). At a high level in a deep learning network, an object or category (e.g., cat) might be defined by a high-dimensional vector. While that vector can be used to identify a cat, it probably does not guide attention toward candidate cats (but see Miconi et al. 2016). It could be that some other, coarser representation, derived from the same network, might be the basis for guidance by shape (analogous to coarse guidance by orientation, described above), but this hypothesis awaits experimental test.

2.3. Scene Guidance

Classic models of search like feature integration theory and S focus on the features of the target. Work over the past 25 years has made it clear that feature guidance is only part of the story. Indeed, for many tasks, feature guidance may not be the most important force that constrains the deployment of attention. That role would go to scene guidance. Classic feature guidance is guidance to locations or objects that appear to have target features. Scene guidance is guidance to locations that are more likely to contain a target, regardless of whether that target and/or its features are present. That is, as in **Figure 2**, search for a pillow will be directed to the top end of the bed, whether or not there is any preattentive feature evidence for the presence of a pillow.

Scene guidance is not a new idea (Biederman et al. 1973, Kingsley 1932), but it has taken on greater prominence in research on visual search, in part because it has become practical to use scenes as stimuli, rather than only random arrays of items.

Working out how scene content and structure guide attention is a project in its fairly early stages. Analogous to Itti et al.'s (1998) saliency map, Henderson & Hayes (2017, 2018) have developed meaning maps that show how the meaningfulness of a scene varies over the scene. Henderson & Ferreira (2004) and Vo & Wolfe (2013) argue that scenes, like language, have a grammar. That grammar is concerned with scene semantics (Is an object meaningful in this location in this scene?) and scene syntax (Where is it structurally plausible for an object to be? For instance, toasters do not float) (Vo & Henderson 2009). In a toaster search, attention would be guided toward places where toasters are physically plausible (e.g., horizontal surfaces) and away from places where toasters, while possible, are semantically improbable (e.g., the floor, under the table). Given the relative inefficiency of search for arbitrary, realistic objects in arrays (Vickery et al. 2005) and scenes (Wolfe et al. 2011a), it seems likely that scene guidance is doing the largest portion of attentional guidance in many real-world searches.

Scene guidance and feature guidance work together. In **Figure 3**, if you look for people, you will quickly find the man in the middle of the image. You may be slower to see the exact copy of his image, pasted on the path at the bottom. You guided your attention to objects or features that could plausibly be human, and the man at the bottom is the wrong size. He is only the wrong size, however, because the structure of the scene makes it so (Eckstein et al. 2017, Wolfe 2017).





Figure 2

Scene guidance: where observers look when told to look for a pillow. Figure reproduced with permission from Sage Boettcher and Melissa Vo, copyright 2019.



Figure 3

The interaction of scene and feature. The observer is asked to look for people.

Guidance mechanisms also know something about the structure of objects. For example, it is easier to find a red object with a yellow part than a red and yellow object, even if both items have the same number of red and yellow pixels (Wolfe et al. 1994, Vul et al. 2020).

In any scene search lasting for more than a fraction of a second, it seems evident that the nature of scene guidance will evolve during search. When a scene is first encountered, scene guidance (like feature guidance) appears coarse. Work on search for humans shows rapid, broad guidance toward the ground plane, for example (Ehinger et al. 2009). This is probably evident in **Figure 3**. Castelhana and colleagues (Castelhana & Heaven 2011, Castelhana & Henderson 2007, Pereira & Castelhana 2014) find that a very brief preview can bias subsequent search to relatively coarse regions of a scene. Moreover, this scene guidance only pertains to objects in scenes, not to objects pasted on top of scenes (Pereira & Castelhana 2019). This coarse initial stage of guidance is consistent with evidence that extremely brief exposure to a scene is enough to extract its gist, giving access to its rough structure and content (Greene & Oliva 2009, Oliva 2005, Oliva & Torralba 2001). As time progresses, more detailed information about a scene and its contents will develop, and search can then be guided by that information. Thus, in search for the pillow in **Figure 2**, attention is guided by the location of the bed (an anchor object; Boettcher et al. 2018). That guidance presumably requires the identification of the bed, which, in turn, probably requires deployment of attention to that bed. Even feature guidance takes some time to fully develop (Lleras et al. 2020, Palmer et al. 2019), although it remains reasonable to think of feature guidance as being based on the output of a fast, preattentive process. Scene guidance, by contrast, seems to be a process that evolves from initial information about the spatial envelope of the scene (Oliva & Torralba 2001) to guidance based on richer scene understanding and informed by attention to objects in the scene.

Another important difference between scene guidance and feature guidance is that much if not all scene guidance must be learned. The genes may code for a set of basic guiding features like color and size, but they do not code for the pillow–bed relationship. Contextual cueing (Chun & Jiang 1998) can be thought of as a very basic version of scene guidance in which repeated exposure to the same random arrays of stimuli leads to learning of the association of a random array with a target location. This type of guidance appears to be largely implicit, even in situations where real world stimuli are used (Jiang et al. 2014). Experts in specialized search tasks like those in radiology speak of a gestalt stage in processing when an initial assessment of the image guides attention to likely locations of a clinical finding (Nodine & Mello-Thoms 2019). This form of guidance is clearly learned. Moreover, it is reasonable to consider this a form of scene search, since radiologists certainly make use of the learned structure of the image as a whole and do not rely exclusively on their identification of specific features or objects.

2.4. Guidance by Search History

In a standard search experiment, and, one may suppose, in a string of searches in the world, performance on one search is influenced by performance on previous searches. The “priming of pop-out” phenomenon, described by Maljkovic & Nakayama (1994), is an important early example. Observers saw displays of red among green or green among red items. They needed to report the precise shape of the color singleton in the display. Maljkovic & Nakayama found that, even though the search for the color target is trivially easy, observers were faster if the preceding target was the same color as the current target. The effect was quite automatic. For instance, it did not matter if the observer knew the next color. This priming was based on the color of the previous target. The result has been replicated and extended to multiple dimensions (Hillstrom 2000, Kristjansson 2006). Priming is not a simple effect. For instance, priming effects for conjunction



stimuli can be longer lasting than single-feature priming of pop-out (Kruijine & Meeter 2015). Priming may be tied to episodic memory representations (Huang et al. 2004; see also Brascamp et al. 2011). Distractor properties can also be primed (Lamy et al. 2013), although the effects are smaller than those for targets (Wolfe et al. 2003).

For the purposes of this review, the important point is that finding something on one trial appears to guide attention toward subsequent instances of that target. Originally, Wolfe et al. (2003) argued that priming was a form of top-down guidance. The logic was that there were two forms of guidance. Bottom-up guidance came from the stimulus, and top-down guidance came from the observer. Since priming was a form of memory, it was internal to the observer and, thus, top down. More recently, Awh and colleagues (2012) have argued that this confuses two forms of guidance that should be considered to be distinct. There is top-down, volitional guidance to what you are looking for, which is distinct from the automatic effects of history that occur whether you are looking or not. Theeuwes (2018) argues that much of what passes for top-down guidance is driven by priming by targets. Even if one might not go that far (Ásgeirsson & Kristjánsson 2019), it seems correct that the recent history of search guides subsequent search.

2.5. Guidance by Value

A fifth form of guidance is related to, but probably distinct from, guidance by history. Attention is guided to features that have been associated with reward (Anderson et al. 2011). Like priming, effects of value can be quite automatic and can act against the interests of the observer (Hickey et al. 2010). Effects of value can be seen in monkeys (Ghazizadeh et al. 2016). The effects can be distinguished from the effects of just seeing and responding to a feature repeatedly, as in priming studies (Kim & Anderson 2018), so it probably makes sense to see value as a separate modulator of attentional priority. The roles of reward and value have been extensively reviewed by Failing & Theeuwes (2017).

2.6. Building a Priority Map

To briefly summarize, attention appears to be guided by at least five factors: bottom-up salience, top-down feature guidance, scene guidance, history, and value (Wolfe & Horowitz 2017). All of these forces may guide attention at any moment; however, selective attention generally has a single focus (for the debate on splitting attention, see Cave et al. 2010, Jans et al. 2010). Accordingly, the multiple sources of guidance need to be combined with some final common path. This is often imagined as a priority map consisting of some weighted average of all forms of guidance (Fecteau & Munoz 2006, Serences & Yantis 2006). The term salience or saliency map should be reserved for bottom-up, stimulus-driven guidance (Koch & Ullman 1985, Li 2002), although this terminological distinction is not always clear.

The basic ideas behind the operation of a priority map were established in the early discussions of saliency maps (Itti & Koch 2001). Information is pooled, and when it is time to deploy attention, a winner-take-all process directs attention to the point of highest activation in the map. This is not the only possibility. For instance, one could propose that attention is directed by one feature map without pooling information across maps. Chan & Hayward (2009) argue that it is possible to respond based on activity in an orientation map or a color map without the need to combine those signals into a priority map. Their dual-route account (Chan & Hayward 2014) acknowledges the combination of signals but does not require that combination. Buetti et al. (2019) provides evidence that the combination of signals does occur. They performed simple pop-out searches for color and shape targets. When the target was defined by both color and shape, RTs were faster than



responses to either feature alone. This indicates that combining features makes a bigger signal, driving faster responding.

The idea of a weighted average is particularly clearly articulated in the dimensional weighting theory of Müller and his group (Liesefeld & Müller 2020). As the name suggests, it puts particular emphasis on modulating the contribution of different dimensions like color and orientation, although, within a dimension like color, one can guide toward a specific feature like red. There appears to be a weaker but real ability to unweight or inhibit distractor features (Cunningham & Egeth 2016, Lamy et al. 2008, Neumann & DeSchepper 1991). It is likely that most of the weight-setting is implicit, based on priming or selection history effects. For a particularly fervid argument on this point, see Theeuwes (2013). This may be more relevant in a long series of trials in the lab than it is during a one-off search in the world. There needs to be some way to guide your search for bananas at the store without having to be primed by immediately preceding banana searches.

Attentional capture (Yantis 1993), for example, by abrupt-onset stimuli (Jonides & Yantis 1988), refers to situations where attention is summoned to a stimulus or location in spite of and, sometimes, in opposition to the observer's intentions. The phenomenon suggests that activity in the priority map is unlabeled, in the sense that attention can be captured by the wrong signal if that signal is incorporated into the priority map. For instance, in singleton mode, attention gets directed to any salient item, regardless of what makes it salient (Bacon & Egeth 1994), although this capture of attention by any oddball depends on the nature of the search task (Jung et al. 2019). Similarly, while the weighting system is flexible, it is probably impossible to set the bottom-up, stimulus-driven salience weight to zero. The classic statement of this principle comes from Sully (1892, p. 146), who said that "One would like to know the fortunate (or unfortunate) man who could receive a box on the ear and not attend to it," meaning that any irrelevant stimulus, if strong enough, will attract attention. Many of the more modern debates over the conditions of attentional capture can be seen as arguments over the strength of the blow to the metaphorical ear.

3. HOW DO SERIAL AND PARALLEL PROCESSES COLLABORATE IN VISUAL SEARCH?

3.1. Overt and Covert Attention: Eye Movements, Eccentricity, and the Functional Visual Field

Attention and eye movements are closely related. However, they can be readily dissociated. It is possible to look at one location while paying attention to another (Helmholtz 1924). It is often useful to refer to overt deployments of the eyes and covert deployments of attention. (Posner 1980). Under most circumstances, covert attention is deployed to the target of the next saccade before the saccade arrives at that location (Kowler et al. 1995). Perhaps the strongest effort to link attention to movements of the eyes is Rizzolatti et al.'s (1987) premotor theory, which holds that deployment of attention is just a weaker manifestation of the signals that direct action toward a location. Premotor theory is probably too strong (Belopolsky & Theeuwes 2012), but there is no doubt that eye movements and attention are closely linked in visual search (Hoffman 1996) even in cases in which search can be accomplished similarly with and without eye movements, i.e., when the items are big enough to avoid eccentricity and crowding effects (Klein & Farrell 1989, Zelinsky & Sheinberg 1997).

Under most real-world conditions, the effects of eccentricity and crowding cannot be avoided (Rosenholtz 2011, Rosenholtz et al. 2012), and as a result, overt eye movements and covert deployments of attention should be seen as participating in a complex, interactive dance during search.



The precise details are unclear because we cannot track covert attention as effectively as we can track the eyes. Under normal conditions, observers will fixate on 3–4 items every second. However, in a search like the search for a T among Ls in **Figure 2e–f**, the $RT \times \text{set size}$ functions indicate that those items are being processed at a rate of 20–40 items/s (details of mapping from slopes to rate depend on assumptions about memory in search; see below).

These numbers show that observers are processing approximately 5–10 items per fixation. How one understands the difference between the rate of eye movements and the rate of item processing in search is at the heart of most theoretical arguments in visual search today. At one extreme is the view that the eyes move to a new fixation every few hundred ms, and that covert attention is then serially deployed to 5–10 items before the eyes move elsewhere. This was the original GS position. Alternatively, attention might move to a location, permitting parallel processing of all items within a functional visual field (FVF) or useful field of view (UFOV) surrounding that point of fixation [there are many versions of this view, but they are well captured by Hulleman & Olivers (2017)]. In the strongest version of this FVF view, the difference in efficiency of, say, the search for a T and the search for a green T in **Figure 2e, f** is not a difference in guidance but a difference in the size of the FVF.

These models sound dramatically different, but they actually may be two ways of looking at the same dance of overt eye movements and covert attention. Once we leave the realm of large, widely spaced items on a computer screen, any GS-style model needs to incorporate the idea of an FVF. This is illustrated in **Figure 4** and described in the caption.

Upon fixation at some point, covert deployments of attention are going to be directed to recognizable items, given the current fixation. These items will fall within some FVF. Outside of the FVF, acuity and/or crowding will render the target unrecognizable. Each covert deployment will start a decision process (Is this a T or an L?) that can be modeled as diffusion of information toward a decision bound (Ratcliff 1978, Ratcliff & McKoon 2008). The covert deployments will

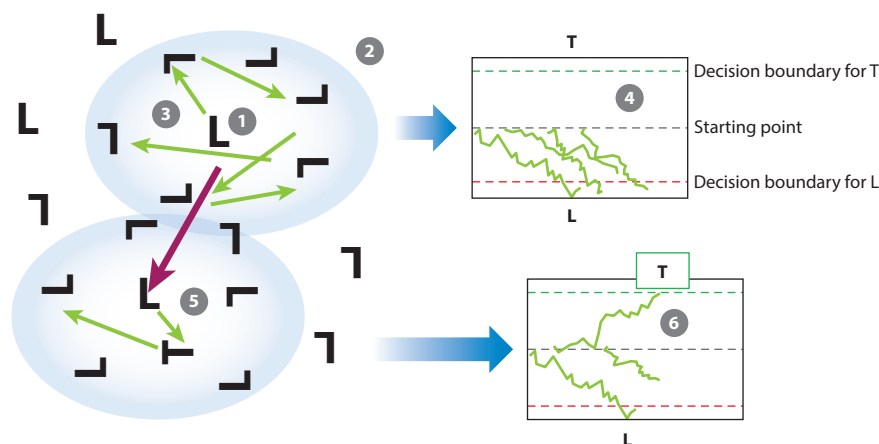


Figure 4

The dance of overt eye movements and covert attention. ① The observer fixates at some point. ② Within a functional visual field, shown by the dotted line, ③ they make covert deployments of attention (*green arrows*). ④ Each attentional selection starts a diffusion process, designed to identify the selected item as a T or an L. In this case, all processes reach the L decision bound. ⑤ Having failed to find a target, the eyes are redeployed to a new spot (*purple arrow*), surrounded by a new functional visual field. Again, covert deployments of attention are made. ⑥ This time, one of those items accumulates the information required to reach the decision boundary identifying it as a T.

occur every 50 ms or so. However, even the most dramatically fast object recognition processes take longer than that (Thorpe et al. 1996, VanRullen & Thorpe 2001). It follows that individual diffusion processes would overlap in time, making the recognition aspect of search into an asynchronous diffusion process. After 200–300 ms, the eyes are redeployed, and covert selection continues within a new FVF. In this case, the target is identified, and search ends.

Wolfe (2003) uses the analogy of a carwash to describe this asynchronous diffusion process. Cars go in one at a time, in series, but they are being processed in parallel in the sense that several cars are being washed at the same time. An experiment that looked at the contents of the diffuser at one instant in time would see the parallel processing of items within an FVF. The only real difference between this serial, covert attention account and a parallel FVF account is whether items within the FVF begin their diffusions synchronously or asynchronously. Given the lamentable lack of a covert attention tracker, mentioned above, and a welter of free parameters (e.g., diffusion rates for items as a function of eccentricity), these models are probably empirically indistinguishable.

3.2. Varieties of Models of Visual Search

Even if it is hard to distinguish between different models of search, this does not mean that there is just one model of search operating under a range of names. Numerous approaches put different emphasis on factors like the role of eye movements or peripheral visual processing. Each model raises issues that can inform the development of the next generation of models.

Treisman & Gelade's (1980) seminal feature integration model (FIT) was built on the two-stage, preattentive–attentive architecture of Neisser (1967). Responding to the neurophysiology of the day, Treisman & Gelade envisioned the preattentive processing of simple features in parallel across the visual field. Binding those features into a recognizable object required serial deployment of attention (Treisman 1996, von der Malsburg 1981). Kristjánsson & Egeth (2020) recently reviewed the work that provided the grounding for the FIT. GS, in its original formulation, simply added the idea that basic, parallel feature information could be used to help deploy serial attention in an intelligent manner (**Figure 2e,f**). Other models can also be seen as variants of the FIT. The dimensional weighting model of Hermann Müller and colleagues (Liesefeld & Müller 2020, Müller et al. 1995, Rangelov et al. 2011) is quite like GS in giving an important role to feature guidance. Different features might have different potential to guide attention in a given task. In a search for a red, vertical target, it would not make much sense to guide attention based on irrelevant motion or size information. The dimensional weighting model emphasizes the need to weight the output of feature processors to get an intelligent guiding signal. Moreover, the model emphasizes the differences between switching from one feature to another within a dimension (red to green) compared to between dimensions (red/color to big/size). The same group has also developed the competitive GS model (Moran et al. 2013, 2015), which improves upon the mechanics of earlier GS models, especially by improving on the treatment of quitting rules, i.e., when an observer stops searching (discussed below). Several other more quantitative versions of GS-style models have been developed (Hubner 2001), with the integrative model of visual search of Schwarz & Miller (2016) being a recent, notable example.

Rensink's (2000) triadic architecture ties FIT-style search architecture more closely to conscious visual experience, adding the question of what we see to the problem of how we find what we are looking for. His parallel front-end stage carves the scene into proto-objects that can be selected by his version of a second, limited-capacity stage. Importantly, he adds a separate non-selective pathway, not subject to attentional limits, that produces some visual experience at all locations in the visual field. These ideas about nonselective gist processing are important in subsequent models like more recent versions of GS (Wolfe et al. 2011b).



Computational saliency models, with their roots in the path-breaking work of Koch & Ullman (1985) and Itti & Koch (2000), can be quite similar in spirit to GS, albeit with stronger ties to the underlying neurophysiology. A set of early-vision features, often modeled on neurophysiological data, are used to derive a saliency map that, in turn, directs attention and/or the eyes. Traditionally, such models have typically been more concerned with the bottom-up, stimulus-driven aspects of attention than with top-down, user-driven aspects. That is, they were modeling what aspects of the world captured attention with less emphasis on how observers got their attention to targets of interest to them. Computationally, modelers were not denying the existence of top-down factors. Initially, their models were just solving other problems. John Tsotsos (2011; Tsotsos et al. 1995) was relatively early in incorporating top-down processing into his computational models, and such factors are now part of standard models in the computational saliency map tradition (Hamker 2004, Navalpakkam & Itti 2006). Moreover, researchers like Tsotsos moved away from the FIT-style model to models much more concerned with how attention could modulate a network of neural connections to cause the right answer to emerge at the top of a pyramid of processing (Tsotsos 2011). For more on this topic, see Frintrop et al.'s (2010) excellent review and the 2015 special issue of *Vision Research* (Tsotsos et al. 2015).

Some models, like Zelinsky's (2008) Target Acquisition Model, place much more emphasis on modeling the overt deployments of the eyes than the more elusive covert deployments of attention. The data that lead to the conclusion that multiple items can be processed with each fixation also lead to models that propose that search progresses by the sequential analysis of clumps of items within an FVF, centered around the current point of fixation. Often, the assumption is that this analysis is parallel over the FVF. A notable example of such a model is that of Hulleman & Olivers (2017). They proclaim "the impending demise of the item in visual search" because they replace search by item with search by FVF clump. More recently, Liesefeld & Müller (2019) have argued that both processes—item by item and FVF clump processing—can be used to find targets in search.

The nature of the FVF that can be processed with each fixation takes a central role in models of search like those of Ruth Rosenholtz and her colleagues (2012). Rosenholtz argues that the inefficiency of searches like classic searches for a T among Ls is driven less by attentional capacity limits and more by limits imposed by crowding and image degradation in peripheral vision. These visual factors mean that a T that is present in the image may not be present in the representation in the early visual system or may be distorted to resemble a distracting L. Other aspects of the statistics of the scene can remain intact and can support good scene perception even if search for specific items in the scene is inefficient (Zhang et al. 2015).

From models that propose parallel processing of clumps of items, it is a conceptually short step to fully parallel models of search (Palmer 1995, Palmer & McLean 1995, Palmer et al. 2000, Vergheze 2001). In these models, all items are processed by either a limited- or an unlimited-capacity parallel processor, and much of the empirical work is devoted to determining the rule that allows the observer to determine if there is a target present and where it might be (e.g., Baldassi & Vergheze 2002). In practice, the fully parallel models have usually been applied to simple displays of a few items. If such a model were to be scaled up to deal with complex scenes, then it seems likely that it would become a model of the FVF, clump-processing variety; no one would propose that a fully parallel model could work over the whole visual field without eye movements. For example, a model like that of Najemnik & Geisler (2009) might combine parallel processing with the constraints imposed by eccentricity.

This section merely outlines the classes of attention models and notes the theoretical considerations that they raise. For a more extended discussion, the reader might consult the relevant chapters in the *Oxford Handbook of Attention* (Nobre & Kastner 2014).



4. SEARCH TEMPLATES AND WORKING MEMORY

A visual search presupposes that something is being searched for. The internal representation of the goal of search is often known as the search template (Gunseli et al. 2014), although everyone understands that this is a metaphor and that a search template does not look like the target in any literal sense. Several terms are used interchangeably, often in the same paper; these terms include search template (Rajsic et al. 2017), target template (Bravo & Farid 2016), memory template (Kristjánsson et al. 2018), and attentional template (Yu & Geng 2019). The jargon confuses two types of template that should be kept distinct. First, “visual search is thought to be guided by an active visual working memory (VWM) representation of the task-relevant features, referred to as the search template” (van Loon et al. 2017). Second, “during visual search, observers hold in mind a search template, which they match against the stimulus” (Bravo & Farid 2014, p. 32). These are not the same role and, in all probability, not the same internal representation. The first template is involved in guiding attention to likely targets and is only fairly coarsely mapped to the target (Anderson 2014, Kerzel 2019). Once the attention is directed to an item, it can be precisely matched to a template (this can be called the target template). A comparison to the target template, not guidance by the features of the guiding template, makes it possible to know that this gray tabby cat is similar to but not the same as your gray tabby cat.

The guiding and target templates are not the same template. The guiding template is thought to reside in WM (Olivers et al. 2011). Many experiments show that changing the contents of WM can bias the guidance of attention (Beck et al. 2011, Dowd & Mitroff 2013, Soto et al. 2005). The idea that the templates are maintained in WM goes back at least to neurophysiological work in the 1990s (Desimone & Duncan 1995). However, while it seems clear that WM serves an important role in search, it is reasonably clear that the target template resides elsewhere. The clearest evidence on this point comes from work on hybrid search (Wolfe 2012). In hybrid search, observers look for any of several targets. In experiments using photographs of objects as targets, observers have no difficulty searching for as many as 100 distinct targets at the same time (or even more; Cunningham & Wolfe 2014). Since the capacity of WM is approximately four items (Cowan 2001), no model of WM would propose that 100 target templates could be in residence in WM at the same time. It could be that target templates shuttle in and out of WM. If that were the case, then we might expect to see qualitatively different performance for hybrid searches with fewer than four items in the memory set and those with the larger memory sets that would require this shuttling. Moreover, we might expect response time to be a linear function of the memory set size. Instead, response time in hybrid search seems to be a smooth, logarithmic function of the memory set size (Wolfe 2012). Moreover, Drew et al. (2015) found that loading WM did not impact hybrid search performance, although a hybrid search did seem to reduce WM capacity by approximately one item.

It seems more plausible to propose that target templates reside in long-term memory (LTM), specifically, in activated long-term memory (Cowan 1995), a partition of LTM that is relevant to the current task. After all, the typical block of search trials and real-world search operate over longer times than those proposed for WM representations, so it is reasonable to assume that some representation of the goal of the search resides in some aspect of LTM. The guiding templates in WM are “representations of the features that distinguish targets from nontarget objects” (Grubert & Eimer 2018, p. 9527). It might be better to call this a “guiding representation” (Wolfe 2007, p. 100; underlining in original), but we are probably stuck with the term template, so perhaps the field can be persuaded to use guiding template or search template for a coarse WM representation that guides search and target template for the LTM representation of the precise goal of the search. The interaction of these types of templates can be seen in categorical hybrid search. Cunningham & Wolfe (2014) had observers memorize several animal images. They then had observers search



for those targets among distractors that could be other animals or letters. By manipulating the numbers of animals and letters in the visual display, they could show that the observers guided attention to the animals—using some set of animal features in a WM guiding template—and then matched those animals against the specific target templates in LTM.

5. SEARCH TERMINATION AND TASKS BEYOND LABORATORY SEARCH

One of the more vexing problems in search is the question of search termination. When is it time to quit the current search task? The original feature integration account assumed that parallel pop-out tasks could be thought of as a single two-alternative forced choice decision: Either the target was present or it was not. For search tasks that the FIT considered to be serial and self-terminating, search ended when the target was found or when all distractors were examined and rejected. Thus, absent trials were imagined to be serial exhaustive searches (or almost exhaustive, given that some miss errors would be made when observers quit before finding the target). GS complicated this picture by having the parallel guidance mechanism eliminate some distractors without serial examination, so GS2 proposed that search ended when the target was found or when all distractors having activation or priority above some threshold had been examined. That activation threshold was set adaptively. It became higher when an observer missed a target and lower after a correct response (Chun & Wolfe 1996). GS2 also proposed that “some very long searches are terminated when the . . . observer . . . concludes that it rarely takes this long to find a target” (Wolfe 1994, p. 210).

A key assumption of this class of model is that rejected distractors or regions are marked in some fashion so that they are not attended again. Inhibition of return (IOR) has been a popular proposed mechanism for this marking (Klein 1988) even though IOR is merely a bias against returning to an attended location, not an absolute prohibition (Klein & MacInnes 1999). However, Horowitz & Wolfe (1998, 2001, 2003) challenged this idea in a series of experiments from which they concluded that “visual search has no memory” (Horowitz & Wolfe 1998)—meaning no memory for rejected distractors. In one set of experiments, they used a dynamic search display where they randomly replotted all of the items every 110 ms. They found that the slope of the $RT \times$ set size function was the same in this dynamic search as it was in an equivalent static search (Horowitz & Wolfe 1998). Perfect memory for rejected distractors would have predicted twice the slope in the dynamic case, where no memory can be used. Subsequent research has pointed out that dynamic search conditions are not identical to static search conditions. Observers could adopt different strategies in the two situations. If items are truly reappearing in random locations, then observers could just sit at one location and wait for the target to appear (von Muhlenen et al. 2003). Although Horowitz & Wolfe (1998) tried to thwart a sit-and-wait strategy, observers might still have been moving attention around the display less widely in dynamic search. Shi et al. (2020) replicated the dynamic search results but focused on the relationship between target-present and -absent slopes. In static search, the absent:present ratio is typically near (or a bit more than) 2:1 (see Wolfe 1998b, figure 2). Shi et al. found that this ratio was consistently near 1:1 for dynamic search, arguing for different strategies in the two conditions. It may well be that observers are answering different questions in the two cases. In a static search, the question might be, “Shall I quit, because I have looked at more or less everything that needs to be looked at?” In dynamic search, that question does not make sense, and the right question might be, “Have I searched long enough that it is unlikely I would have missed the target?”

Horowitz & Wolfe (2001) used methods other than dynamic search to argue for a failure to keep track of rejected distractors, but others have found evidence for some, albeit fairly limited,

2.14 Wolfe



Review in Advance first posted on April 22, 2020. (Changes may still occur before final publication.)

memory (Dickinson & Zelinsky 2005, McCarley et al. 2003). Of course, in extended search tasks over longer time scales (e.g., searching the house for your keys), there must be some memory for where you have looked, but it seems safe to argue that search is not terminated by checking when all relevant items have been marked as rejected because observers do not keep reliable track of what distractors have been rejected.

The search termination question becomes more complicated when the family of search tasks expands beyond the search for a single target among discrete distractors. These laboratory tasks do not capture the richness of real-world searches, including socially important artificial search tasks like those in radiology or security screening. For instance, what happens if the number of targets is unknown? In radiology, the phenomenon of satisfaction of search (SoS) describes a situation in which finding one target makes it less likely that an observer will find a second target (Nodine et al. 1992, Tuddenham 1962). The original thought was that SoS was caused by early search termination, but that turns out to be incorrect (Berbaum et al. 1991). A body of research shows that there are multiple causes (Berbaum et al. 2019, Cain et al. 2013), possibly including early termination.

More generally, search tasks in areas such as medical image perception, airport security (Mitroff et al. 2014, Wolfe et al. 2013), and driving (Robbins & Chapman 2018) are different from laboratory studies of search for one target that is present on 50% of trials (for recent reviews of medical image perception, see Williams & Drew 2019, Wolfe 2016, Wolfe et al. 2018, Wu & Wolfe 2019; most notably, see Samei & Krupinski 2018).

When the number of targets becomes large, search tasks become foraging tasks, as, for example, when you are picking berries. You are performing a search task for each berry, but the more interesting question in foraging becomes, “When is it time to move to the next berry bush?” The animal literature on foraging provides rules for patch leaving, like the marginal value theorem (MVT; Charnov 1976). These can be seen as variants of search termination rules. They predict that foragers should leave some targets behind as the rate of collection drops and the next patch, bush, or display seems more promising. Humans roughly follow these optimal foraging predictions (Wolfe 2013), although one can imagine times when it would be important to find every target (e.g., cancer metastases) and where an MVT-style strategy would not be good.

Tasks like cancer and airport screening raise a different issue: target prevalence. Laboratory search tasks tend to have targets present on 50% of trials (sometimes 100% if the task is to localize the target). Screening tasks, by contrast, have very low prevalence (e.g., 3–5 cancers per 1,000 breast cancer screening cases; Jackson et al. 2012). Low prevalence causes observers to miss targets that they would have found in a higher-prevalence context (Wolfe et al. 2005), even in true clinical settings (Evans et al. 2013). Observers do tend to quit more quickly in low-prevalence search, but the main driver of the prevalence effect is a criterion shift: Observers become more reluctant to identify an ambiguous item as a target (for a review, see Horowitz 2017).

Can a generalized quitting model apply across these different types of search tasks? Perhaps we can formulate quitting across tasks as a negative answer to the question, “Is continuing to search this stimulus a good use of my time?” Thus, a basic quitting model could be a quitting signal that rises toward a threshold as a function of time or of successive deployments of attention. This is roughly equivalent to a collapsing decision bound that terminates the search early, typically with a target-absent response. In the quitting signal version, search ends when the target is found, when enough targets are found, or when the quitting signal reaches a quitting threshold. This is the rule that was implemented by Wolfe & Van Wert (2010). But how is that quitting threshold set? There are some complications. Chun & Wolfe (1996) proposed an adaptive rule where observers search longer after making an error and for a shorter period after correctly quitting, and versions of this idea have been part of GS ever since (see also Shi et al. 2020). Simply searching for approximately



N ms or searching through approximately N items will not work, even in a standard search experiment, if the set size is varied. The quitting time on a trial must vary with the set size on that trial. Interestingly, $RT \times \text{set size}$ functions and error rates look substantially the same whether different set sizes are mixed or run as separate blocks (Wolfe et al. 2010a). Thus, it seems that the quitting threshold on a given trial is based on a rapid assessment of the stimuli present on that trial. You cannot search through some percentage of the items if you do not have some notion of how many items are present. Roughly equivalently, you cannot know how long to search if you do not have some ideas about the stimulus that is being searched.

This initial assessment of the nature of the search array must be done very quickly given that RTs for simple target-absent trials can be just a few hundred ms long, including the motor response time. Thus, the first step in setting a quitting threshold probably involves assessing the gist of the scene via the nonselective pathway, mentioned above. A glance will give an observer some idea of how long to search before giving up on the hunt for granola at the hotel breakfast buffet. Importantly, this assessment and the setting of that initial quitting threshold do not require experience with a block of practice trials. After all, although lab experiments may involve hundreds of essentially similar trials, many real-world searches are unique, one-time events. If there are multiple searches, then observers can adaptively refine quitting thresholds based on feedback about errors (Chun & Wolfe 1996), target prevalence (Wolfe & Van Wert 2010), and developing expertise (Brams et al. 2019). Moreover, observers probably adjust quitting thresholds within a single search based on an evolving assessment of the current stimulus. This is nicely captured in the competitive GS model of Moran et al. (2013). They model the probability of quitting at each step in a search as $P(\text{quit}) = W_{\text{quit}} / (W_{\text{quit}} + \Sigma(\text{stimuli}))$, where W_{quit} is a quitting signal that grows on each step, and $\Sigma(\text{stimuli})$ is some assessment of the likelihood that a target could be present (Moran's et al.'s definition is more precise.) Early in a search, $P(\text{quit})$ will be dominated by stimulus. As time goes on, W_{quit} will push $P(\text{quit})$ toward 1.0.

If we assume that the stimulus signal is bigger when a target is present (even if that target has not been found; see, for example, Evans et al. 2013), then this formulation can help explain a curious mystery of search termination. If we look at RT distributions, then we can compare the mean quitting time to the target present distribution. One would expect that the percentage of target-present RTs falling above that target-absent mean would be roughly the percentage of miss errors. In fact, the percentage of target-present RTs above the target-absent mean is much greater than the miss error rate (Wolfe et al. 2010b, figure 5). One possibility, captured by the Moran et al. (2013) model, is that a larger stimulus signal on target-present trials keeps observers searching longer when there is, in fact, a target to be found. This topic deserves further research.

6. CONCLUSIONS

Writing a review of visual search is daunting. For an apparently narrow topic, the literature is very large, and the demands of space have resulted in the omission of many topics. Whole reviews could be and have been written about the underlying neurophysiology (e.g., Treue 2014), and much more could be written about topics that were only glancingly discussed, like attentional capture. This review tries to emphasize areas with significant potential for future progress. A noncomprehensive list of research topics would include:

1. Guidance, notably, scene guidance. Scene guidance does a great deal of work in the real world. It is not entirely clear how that is accomplished.
2. Search templates and the internal workings of search. There is work to be done if we want to understand how an intention to find the cat is translated into target templates and guiding templates that allow you to perform the task.



3. Search termination. Do different tasks use versions of the same termination strategy, or do we need to figure out different rules for each task?
4. Use-inspired search tasks. We are only at the beginning of drawing research projects from real-world search problems. Most of what we have studied to date has involved real-world tasks that conveniently mimic the repetitive structure of trial-based lab paradigms. For practical reasons, we know very little about search tasks that evolve over minutes, hours, or days. We also know little about one-shot search tasks. For example, when you are searching for a place to eat in Campo de Fiore in Rome, you are not going to do that hundreds of times. How do you get it right (or right enough) the first time? (The answer, by the way, is to eat at Roscioli's.)

The questions will not run out before it is time for the next review.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Work on this chapter was supported by the National Institutes of Health (grants EY017001 and CA207490) and the National Science Foundation (grant 1848783). I thank Wanyi Lu, Sneha Suresh, Hsing-Fen Tu, and Farahnaz Wick for their help.

LITERATURE CITED

- Adams WJ. 2008. Frames of reference for the light-from-above prior in visual search and shape judgements. *Cognition* 107:137–50
- Anderson BA. 2014. On the precision of goal-directed attentional selection. *J. Exp. Psychol. Hum. Percept. Perform.* 40:1755–62
- Anderson BA, Laurent PA, Yantis S. 2011. Value-driven attentional capture. *PNAS* 108:10367–71
- Ásgeirsson ÁG, Kristjánsson Á. 2019. Attentional priming does not enable observers to ignore salient distractors. *Vis. Cogn.* 27:595–608
- Awh E, Belopolsky AV, Theeuwes J. 2012. Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends Cogn. Sci.* 16:437–43
- Bacon WF, Egeth HE. 1994. Overriding stimulus-driven attentional capture. *Percept. Psychophys.* 55:485–96
- Baldassi S, Verghese P. 2002. Comparing integration rules in visual search. *J. Vis.* 2:559–70
- Bauer B, Jolicœur P, Cowan WB. 1996. Visual search for colour targets that are or are not linearly-separable from distractors. *Vis. Res.* 36:1439–66
- Beck VM, Hollingworth A, Luck SJ. 2011. Simultaneous control of attention by multiple working memory representations. *Psychol. Sci.* 23:887–98
- Becker S, Atalla M, Folk CL. 2020. Conjunction search: Can we simultaneously bias attention to features and relations? *Atten. Percept. Psychophys.* In press
- Becker SI. 2010. The role of target-distractor relationships in guiding attention and the eyes in visual search. *J. Exp. Psychol. Gen.* 139:247–65
- Belopolsky AV, Theeuwes J. 2012. Updating the premotor theory: The allocation of attention is not always accompanied by saccade preparation. *J. Exp. Psychol. Hum. Percept. Perform.* 38:902–14
- Berbaum KS, Franken EA, Caldwell RT, Shartz K, Madsen M. 2019. Satisfaction of search in radiology. In *The Handbook of Medical Image Perception and Techniques*, ed. E Samei, EA Krupinski, pp. 121–66. Cambridge, UK: Cambridge Univ. Press. 2nd ed.



- Berbaum KS, Franken EA Jr., Dorfman DD, Rooholamini SA, Coffman CE, et al. 1991. Time course of satisfaction of search. *Invest. Radiol.* 26:640–48
- Biederman I, Glass AL, Stacy EW. 1973. Searching for objects in real-world scenes. *J. Exp. Psychol.* 97:22–27
- Boettcher SEP, Draschkow D, Dienhart E, Vö MLH. 2018. Anchoring visual search in scenes: assessing the role of anchor objects on eye movements during visual search. *J. Vis.* 18:11
- Brams S, Ziv G, Levin O, Spitz J, Wagemans J, et al. 2019. The relationship between gaze behavior, expertise, and performance: a systematic review. *Psychol. Bull.* 145(10):980–1027
- Brascamp JW, Pels E, Kristjansson A. 2011. Priming of pop-out on multiple time scales during visual search. *Vis. Res.* 51:1972–78
- Bravo MJ, Farid H. 2014. Informative cues can slow search: the cost of matching a specific template. *Attention Percept. Psychophys.* 76:32–39
- Bravo MJ, Farid H. 2016. Observers change their target template based on expected context. *Atten. Percept. Psychophys.* 78:829–37
- Buetti S, Xu J, Lleras A. 2019. Predicting how color and shape combine in the human visual system to direct attention. *Sci. Rep.* 9:20258
- Cain MS, Adamo SH, Mitroff SR. 2013. A taxonomy of errors in multiple-target visual search. *Vis. Cogn.* 21:899–921
- Castelhano M, Heaven C. 2011. Scene context influences without scene gist: eye movements guided by spatial associations in visual search. *Psychon. Bull. Rev.* 18:890–96
- Castelhano MS, Henderson JM. 2007. Initial scene representations facilitate eye movement guidance in visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 33:753–63
- Cave KR, Bush WS, Taylor TGG. 2010. Split attention as part of a flexible attentional system for complex scenes: comment on Jans, Peters, and De Weerd 2010. *Psychol. Rev.* 117:685–95
- Chan LK, Hayward WG. 2014. No attentional capture for simple visual search: evidence for a dual-route account. *J. Exp. Psychol. Hum. Percept. Perform.* 40:2154–66
- Chan LKH, Hayward WG. 2009. Feature integration theory revisited: dissociating feature detection and attentional guidance in visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 35:119–32
- Charnov EL. 1976. Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* 9:129–36
- Chun M, Jiang Y. 1998. Contextual cuing: Implicit learning and memory of visual context guides spatial attention. *Cogn. Psychol.* 36:28–71
- Chun MM, Wolfe JM. 1996. Just say no: How are visual searches terminated when there is no target present? *Cogn. Psychol.* 30:39–78
- Cousineau D, Shiffrin RM. 2004. Termination of a visual search with large display size effects. *Spat. Vis.* 17:327–52
- Cowan N. 1995. *Attention and Memory: An Integrated Framework*. Oxford, UK: Oxford Univ. Press
- Cowan N. 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24:87–114
- Cunningham CA, Egeth HE. 2016. Taming the white bear: initial costs and eventual benefits of distractor inhibition. *Psychol. Sci.* 27:476–85
- Cunningham CA, Wolfe JM. 2014. The role of object categories in hybrid visual and memory search. *J. Exp. Psychol. Gen.* 143:1585–99
- Desimone R, Duncan J. 1995. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18:193–222
- Dickinson CA, Zelinsky GJ. 2005. Marking rejected distractors: a gaze-contingent technique for measuring memory during search. *Psychon. Bull. Rev.* 12:1120–26
- Dowd EW, Mitroff SR. 2013. Attentional guidance by working memory overrides saliency cues in visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 39:1786–96
- Drew T, Boettcher SP, Wolfe JM. 2015. Searching while loaded: Visual working memory does not interfere with hybrid search efficiency but hybrid search uses working memory capacity. *Psychon. Bull. Rev.* 23:201–12
- Eckstein M, Koehler K, Welbourne LE, Akbas EP. 2017. Humans but not deep neural networks often miss giant targets in scenes. *Curr. Biol.* 27:2827–32.e3



- Egeth H. 1977. Attention and preattention. In *The Psychology of Learning and Motivation*, ed. GH Bower, pp. 277–320. New York: Academic
- Egeth HE, Virzi RA, Garbart H. 1984. Searching for conjunctively defined targets. *J. Exp. Psychol. Hum. Percept. Perform.* 10:32–39
- Ehinger KA, Hidalgo-Sotelo B, Torralba A, Oliva A. 2009. Modeling search for people in 900 scenes: a combined source model of eye guidance. *Vis. Cogn.* 17:945–78
- Enns JT, Rensink RA. 1990. Influence of scene-based properties on visual search. *Science* 247:721–23
- Evans KK, Birdwell RL, Wolfe JM. 2013. If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening. *PLOS ONE* 8(5):e64366
- Failing M, Theeuwes J. 2017. Selection history: how reward modulates selectivity of visual attention. *Psychon. Bull. Rev.* 25:514–38
- Fecteau JH, Munoz DP. 2006. Saliency, relevance, and firing: a priority map for target selection. *Trends Cogn. Sci.* 10:382–90
- Foster DH, Ward PA. 1991. Horizontal-vertical filters in early vision predict anomalous line-orientation frequencies. *Proc. R. Soc. Lond. B* 243:83–86
- Frintrop S, Rome E, Christensen HI. 2010. Computational visual attention systems and their cognitive foundations: a survey. *ACM Trans. Appl. Percept.* 7:6
- Gaspelin N, Vecera S. 2019. An introduction to the special issue on “Dealing with Distractors in Visual Search.” *Vis. Cogn.* 27:183–84
- Ghazizadeh A, Griggs W, Hikosaka O. 2016. Object-finding skill created by repeated reward experience. *J. Vis.* 16:17
- Green BF, Anderson LK. 1956. Color coding in a visual search task. *J. Exp. Psychol.* 51:19–24
- Greene MR, Oliva A. 2009. The briefest of glances: the time course of natural scene understanding. *Psychol. Sci.* 20:464–72
- Grubert A, Eimer M. 2018. The time course of target template activation processes during preparation for visual search. *J. Neurosci.* 38:9527–38
- Gunseli E, Meeter M, Olivers CNL. 2014. Is a search template an ordinary working memory? Comparing electrophysiological markers of working memory maintenance for visual search and recognition. *Neuropsychologia* 60:29–38
- Hamker FH. 2004. A dynamic model of how feature cues guide spatial attention. *Vis. Res.* 44:501–21
- Helmholtz HV. 1924. *Treatise on Physiological Optics*. Rochester, NY: Opt. Soc. Am.
- Henderson JM, Ferreira F. 2004. Scene perception for psycholinguists. In *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*, ed. JM Henderson, F Ferreira, pp. 1–58. New York: Psychol. Press
- Henderson JM, Hayes TR. 2017. Meaning guides attention in real-world scenes. *Nat. Hum. Behav.* 1:743–47
- Henderson JM, Hayes TR. 2018. Meaning guides attention in real-world scene images: evidence from eye movements and meaning maps. *J. Vis.* 18:10
- Hickey C, Chelazzi L, Theeuwes J. 2010. Reward changes saliency in human vision via the anterior cingulate. *J. Neurosci.* 30:11096–103
- Hillstrom AP. 2000. Repetition effects in visual search. *Percept. Psychophys.* 62:800–17
- Hoffman JE. 1979. A two-stage model of visual search. *Percept. Psychophys.* 25:319–27
- Hoffman JE. 1996. Visual attention and eye movements. In *Attention*, ed. H Pashler, pp. 119–54. London: Univ. Coll. Lond. Press
- Horowitz TS. 2017. Prevalence in visual search: from the clinic to the lab and back again. *Jpn. Psychol. Res.* 59:65–108
- Horowitz TS, Wolfe JM. 1998. Visual search has no memory. *Nature* 394:575–77
- Horowitz TS, Wolfe JM. 2001. Search for multiple targets: Remember the targets, forget the search. *Percept. Psychophys.* 63:272–85
- Horowitz TS, Wolfe JM. 2003. Memory for rejected distractors in visual search? *Vis. Cogn.* 10:257–98
- Huang L, Holcombe AO, Pashler H. 2004. Repetition priming in visual search: episodic retrieval, not feature priming. *Mem. Cogn.* 32:12–20
- Hubner R. 2001. A formal version of the Guided Search (GS2) model. *Percept. Psychophys.* 63:945–51



- Hulleman J, Olivers CNL. 2017. The impending demise of the item in visual search. *Behav. Brain Sci.* 40:e132
- Itti L, Koch C. 2000. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis. Res.* 40:1489–506
- Itti L, Koch C. 2001. Computational modelling of visual attention. *Nat. Rev. Neurosci.* 2:194–203
- Itti L, Koch C, Niebur E. 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 20:1254–59
- Jackson SL, Cook AJ, Miglioretti DL, Carney PA, Geller BM, et al. 2012. Are radiologists' goals for mammography accuracy consistent with published recommendations? *Acad. Radiol.* 19:289–95
- Jans B, Peters JC, De Weerd P. 2010. Visual spatial attention to multiple locations at once: The jury is still out. *Psychol. Rev.* 117:637–82
- Jiang YV, Won B-Y, Swallow KM, Mussack DM. 2014. Spatial reference frame of attention in a large outdoor environment. *J. Exp. Psychol. Hum. Percept. Perform.* 40:1346–57
- Jonides J, Yantis S. 1988. Uniqueness of abrupt visual onset in capturing attention. *Percept. Psychophys.* 43:346–54
- Jung K, Han SW, Min Y. 2019. Search efficiency is not sufficient: The nature of search modulates stimulus-driven attention. *Atten. Percept. Psychophys.* 81:61–70
- Kerzel D. 2019. The precision of attentional selection is far worse than the precision of the underlying memory representation. *Cognition* 186:20–31
- Kim H, Anderson BA. 2018. Dissociable components of experience-driven attention. *Curr. Biol.* 29:841–45.e2
- Kingsley HL. 1932. An experimental study of 'search.' *Am. J. Psychol.* 44:314–18
- Klein R. 1988. Inhibitory tagging system facilitates visual search. *Nature* 334:430–31
- Klein R, Farrell M. 1989. Search performance without eye movements. *Percept. Psychophys.* 46:476–82
- Klein RM, MacInnes WJ. 1999. Inhibition of return is a foraging facilitator in visual search. *Psychol. Sci.* 10:346–52
- Koch C, Ullman S. 1985. Shifts in selective visual attention: towards the underlying neural circuitry. *Hum. Neurobiol.* 4:219–27
- Kowler E, Anderson E, Doshier B, Blaser E. 1995. The role of attention in the programming of saccades. *Vis. Res.* 35:1897–916
- Kriegeskorte N. 2015. Deep neural networks: a new framework for modeling biological vision and brain information processing. *Annu. Rev. Vis. Sci.* 1:417–46
- Kristjánsson A. 2006. Simultaneous priming along multiple feature dimensions in a visual search task. *Vis. Res.* 46:2554–70
- Kristjánsson A, Egeth HE. 2020. How feature integration theory integrated cognitive psychology, neurophysiology, and psychophysics. *Atten. Percept. Psychophys.* In press
- Kristjánsson T, Thornton IM, Kristjánsson Á. 2018. Time limits during visual foraging reveal flexible working memory templates. *J. Exp. Psychol. Hum. Percept. Perform.* 44:827–35
- Kruijine W, Meeter M. 2015. The long and the short of priming in visual search. *Atten. Percept. Psychophys.* 77:1558–73
- Lamy D, Antebi C, Aviani N, Carmel T. 2008. Priming of pop-out provides reliable measures of target activation and distractor inhibition in selective attention. *Vis. Res.* 48:30–41
- Lamy D, Yashar A, Ruderman L. 2013. Orientation search is mediated by distractor suppression: evidence from priming of pop-out. *Vis. Res.* 81:29–35
- Li Z. 2002. A salience map in primary visual cortex. *Trends Cogn. Sci.* 6:9–16
- Liesefeld H, Mueller HJ. 2020. A theoretical attempt to revive the serial/parallel-search dichotomy. *Atten. Percept. Psychophys.* In press
- Liesefeld HR, Müller HJ. 2019. Distractor handling via dimension weighting. *Curr. Opin. Psychol.* 29:160–67
- Lleras A, Wang Z, Ng GJP, Ballew K, Xu J, Buetti S. 2020. A target contrast signal theory of parallel processing in goal-directed search. *Atten. Percept. Psychophys.* In press
- Maljkovic V, Nakayama K. 1994. Priming of popout: I. Role of features. *Mem. Cogn.* 22:657–72
- McCarley JS, Wang RF, Kramer AF, Irwin DE, Peterson MS. 2003. How much memory does oculomotor search have? *Psychol. Sci.* 14:422–26



- Miconi T, Groomes L, Kreiman G. 2016. There's Waldo! A normalization model of visual search predicts single-trial human fixations in an object search task. *Cereb. Cortex* 26:3064–82
- Mitroff SR, Biggs AT, Adamo SH, Dowd EW, Winkle J, Clark K. 2014. What can 1 billion trials tell us about visual search? *J. Exp. Psychol. Hum. Percept. Perform.* 41:1–5
- Moran R, Zehetleitner M, Liesefeld H, Müller H, Usher M. 2015. Serial vs. parallel models of attention in visual search: accounting for benchmark RT-distributions. *Psychon. Bull. Rev.* 23:1300–15
- Moran R, Zehetleitner MH, Mueller HJ, Usher M. 2013. Competitive guided search: meeting the challenge of benchmark RT distributions. *J. Vis.* 13:24
- Müller HJ, Heller D, Ziegler J. 1995. Visual search for singleton feature targets within and across feature dimensions. *Percept. Psychophys.* 57:1–17
- Najemnik J, Geisler WS. 2009. Simple summation rule for optimal fixation selection in visual search. *Vis. Res.* 49:1286–94
- Navalpakkam V, Itti L. 2006. Top-down attention selection is fine grained. *J. Vis.* 6:1180–93
- Neider MB, Zelinsky GJ. 2008. Exploring set size effects in scenes: identifying the objects of search. *Vis. Cogn.* 16:1–10
- Neisser U. 1967. *Cognitive Psychology*. New York: Appleton, Century, Crofts
- Neumann E, DeSchepper BG. 1991. Costs and benefits of target activation and distractor inhibition in selective attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 17:1136–45
- Nobre AC, Kastner S. 2014. *Oxford Handbook of Attention*. Oxford, UK: Oxford Univ. Press
- Nodine CF, Krupinski EA, Kundel HL, Toto L, Herman GT. 1992. Satisfaction of search (SOS). *Invest. Radiol.* 27:571–73
- Nodine CF, Mello-Thoms C. 2019. Acquiring expertise in radiologic image interpretation. In *The Handbook of Medical Image Perception and Techniques*, ed. E Samei, EA Krupinski, pp. 139–56. Cambridge, UK: Cambridge Univ. Press. 2nd ed.
- Oliva A. 2005. Gist of the scene. In *Neurobiology of Attention*, ed. L Itti, G Rees, J Tsotsos, pp. 251–57. New York: Academic
- Oliva A, Torralba A. 2001. Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vis.* 42:145–75
- Olivers CN, Peters J, Houtkamp R, Roelfsema PR. 2011. Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn. Sci.* 15:327–34
- Palmer EM, Van Wert MJ, Horowitz TS, Wolfe JM. 2019. Measuring the time course of selection during visual search. *Atten. Percept. Psychophys.* 81:47–60
- Palmer J. 1995. Attention in visual search: distinguishing four causes of a set size effect. *Curr. Dir. Psychol. Sci.* 4:118–23
- Palmer J, McLean J. 1995. *Imperfect, unlimited-capacity, parallel search yields large set-size effects*. Paper presented at the Meeting of the Society for Mathematical Psychology, Irvine, CA
- Palmer J, Verghese P, Pavel M. 2000. The psychophysics of visual search. *Vis. Res.* 40:1227–68
- Pereira EJ, Castelano MS. 2014. Peripheral guidance in scenes: the interaction of scene context and object content. *J. Exp. Psychol. Hum. Percept. Perform.* 40:2056–72
- Pereira EJ, Castelano MS. 2019. Attentional capture is contingent on scene region: using surface guidance framework to explore attentional mechanisms during search. *Psychon. Bull. Rev.* 26:1273–81
- Posner MI. 1980. Orienting of attention. *Q. J. Exp. Psychol.* 32:3–25
- Rajsic J, Ouslis NE, Wilson DE, Pratt J. 2017. Looking sharp: Becoming a search template boosts precision and stability in visual working memory. *Atten. Percept. Psychophys.* 79:1643–51
- Rangelov D, Muller HJ, Zehetleitner M. 2011. Dimension-specific intertrial priming effects are task-specific: evidence for multiple weighting systems. *J. Exp. Psychol. Hum. Percept. Perform.* 37:100–14
- Ratcliff R. 1978. A theory of memory retrieval. *Psychol. Rev.* 85:59–108
- Ratcliff R, McKoon G. 2008. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20:873–922
- Rensink RA. 2000. Seeing, sensing, and scrutinizing. *Vis. Res.* 40:1469–87
- Rizzolatti G, Riggio L, Dascola I, Umiltà C. 1987. Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia* 25:31–40



- Robbins CJ, Chapman P. 2018. Drivers' visual search behavior toward vulnerable road users at junctions as a function of cycling experience. *Hum. Factors* 60:889–901
- Rosenholtz R, Huang J, Ehinger KA. 2012. Rethinking the role of top-down attention in vision: effects attributable to a lossy representation in peripheral vision. *Front. Psychol.* 3:13
- Rosenholtz RE. 2011. What your visual system sees where you are not looking. In *Proc. SPIE: Human Vision and Electronic Imaging, XVI*, ed. BE Rogowitz, TN Pappas, art. 786510. San Francisco, CA: SPIE
- Samei E, Krupinski EA. 2018. *The Handbook of Medical Image Perception and Techniques*. Cambridge, UK: Cambridge Univ. Press. 2nd ed.
- Schill HM, Cain MS, Josephs EL, Wolfe JM. 2020. Axis of rotation as a basic feature in visual search. *Atten. Percept. Psychophys.* In press
- Schwarz W, Miller JO. 2016. GSDT: an integrative model of visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 42:1654–75
- Serences JT, Yantis S. 2006. Selective visual attention and perceptual coherence. *Trends Cogn. Sci.* 10:38–45
- Sharan L, Rosenholtz R, Adelson EH. 2014. Accuracy and speed of material categorization in real-world images. *J. Vis.* 14:12
- Shi Z, Allenmark F, Zhu X, Elliott MA, Müller HJ. 2020. To quit or not to quit in dynamic search. *Atten. Percept. Psychophys.* In press
- Soto D, Heinke D, Humphreys GW, Blanco MJ. 2005. Early, involuntary top-down guidance of attention from working memory. *J. Exp. Psychol. Hum. Percept. Perform.* 31:248–61
- Sternberg S. 1966. High-speed scanning in human memory. *Science* 153:652–54
- Sully J. 1892. *The Human Mind: A Text-Book of Psychology*. New York: D. Appleton & Co.
- Theeuwes J. 2013. Feature-based attention: It is all bottom-up priming. *Philos. Trans. R. Soc. B* 368:20130055
- Theeuwes J. 2018. Visual selection: usually fast and automatic; seldom slow and volitional. *J. Cogn.* 1:29
- Thorpe S, Fize D, Marlot C. 1996. Speed of processing in the human visual system. *Nature* 381:520–52
- Treisman A. 1996. The binding problem. *Curr. Opin. Neurobiol.* 6:171–78
- Treisman A, Gelade G. 1980. A feature-integration theory of attention. *Cogn. Psychol.* 12:97–136
- Treue S. 2014. Object- and feature-based attention: monkey physiology. In *Oxford Handbook of Attention*, ed. AC Nobre, S Kastner, pp. 573–600. Oxford, UK: Oxford Univ. Press
- Tsotsos J. 2011. *A Computational Perspective on Visual Attention*. Cambridge, MA: MIT Press
- Tsotsos JK, Culhane SN, Wai WYK, Lai Y, Davis N, Nuflo F. 1995. Modeling visual attention via selective tuning. *Artif. Intell.* 78:507–45
- Tsotsos JK, Eckstein MP, Landy MS. 2015. Computational models of visual attention. *Vis. Res.* 116(Pt. B):93–94
- Tuddenham WJ. 1962. Visual search, image organization, and reader error in roentgen diagnosis. Studies of the psycho-physiology of roentgen image perception. *Radiology* 78:694–704
- van Loon AM, Olmos-Solis K, Olivers CNL. 2017. Subtle eye movement metrics reveal task-relevant representations prior to visual search. *J. Vis.* 17:13
- VanRullen R, Thorpe SJ. 2001. Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception* 30:655–68
- Vergheze P. 2001. Visual search and attention: a signal detection approach. *Neuron* 31:523–35
- Vickery TJ, King L-W, Jiang Y. 2005. Setting up the target template in visual search. *J. Vis.* 5:81–92
- Vo ML, Wolfe JM. 2013. Differential ERP signatures elicited by semantic and syntactic processing in scenes. *Psychol. Sci.* 24:1816–23
- Vo MLH, Henderson JM. 2009. Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *J. Vis.* 9:24
- von der Malsburg C. 1981. The correlation theory of brain function. In *Models of Neural Networks, Vol. II: Temporal Aspects of Coding and Information Processing in Biological Systems*, ed. E Domany, JL van Hemmen, K Schulten, pp. 95–119. Berlin: Springer
- von Muhlenen A, Muller HJ, Muller D. 2003. Sit-and-wait strategies in dynamic visual search. *Psychol. Sci.* 14:309–14
- Vul E, Rieth C, Lew TF, Rich AN. 2020. The structure of illusory conjunctions reveals hierarchical binding of multi-part objects. *Atten. Percept. Psychophys.* In press



- Williams L, Drew T. 2019. What do we know about volumetric medical image search? A review of the basic science and medical image perception literatures. *Cogn. Res. Princ. Implic.* 4:21
- Wolfe JM. 1994. Guided Search 2.0: a revised model of visual search. *Psychon. Bull. Rev.* 1:202–38
- Wolfe JM. 1998a. Visual search. In *Attention*, ed. H Pashler, pp. 13–74. Hove, UK: Psychol. Press
- Wolfe JM. 1998b. What do 1,000,000 trials tell us about visual search? *Psychol. Sci.* 9:33–39
- Wolfe JM. 2003. Moving towards solutions to some enduring controversies in visual search. *Trends Cogn. Sci.* 7:70–76
- Wolfe JM. 2007. Guided Search 4.0: current progress with a model of visual search. In *Integrated Models of Cognitive Systems*, ed. W Gray, pp. 99–119. Oxford, UK: Oxford Univ. Press
- Wolfe JM. 2012. Saved by a log: How do humans perform hybrid visual and memory search? *Psychol. Sci.* 23:698–703
- Wolfe JM. 2013. When is it time to move to the next raspberry bush? Foraging rules in human visual search. *J. Vis.* 13:10
- Wolfe JM. 2014. Approaches to visual search: feature integration theory and guided search. In *Oxford Handbook of Attention*, ed. AC Nobre, S Kastner, pp. 11–55. Oxford, UK: Oxford Univ. Press
- Wolfe JM. 2016. Use-inspired basic research in medical image perception. *Cogn. Res. Princ. Implic.* 1:17
- Wolfe JM. 2017. Visual attention: size matters. *Curr. Biol.* 27:R1002–3
- Wolfe JM. 2018. Visual search. In *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, ed. J Wixted, pp. 569–623. Hoboken, NJ: Wiley
- Wolfe JM, Alvarez GA, Rosenholtz R, Kuzmova YI, Sherman AM. 2011a. Visual search for arbitrary objects in real scenes. *Atten. Percept. Psychophys.* 73:1650–71
- Wolfe JM, Brunelli DN, Rubinstein J, Horowitz TS. 2013. Prevalence effects in newly trained airport checkpoint screeners: Trained observers miss rare targets, too. *J. Vis.* 13:33
- Wolfe JM, Butcher SJ, Lee C, Hyle M. 2003. Changing your mind: on the contributions of top-down and bottom-up guidance in visual search for feature singletons. *J. Exp. Psychol. Hum. Percept. Perform.* 29:483–502
- Wolfe JM, Cain MS, Ehinger KA, Drew T. 2015. *Guided Search 5.0: meeting the challenge of hybrid search and multiple-target foraging*. Paper presented at the Annual Meeting of the Vision Science Society, St. Petersburg, FL, May 15–20
- Wolfe JM, Cave KR, Franzel SL. 1989. Guided Search: an alternative to the Feature Integration model for visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 15:419–33
- Wolfe JM, Evans KK, Drew T. 2018. The first moments of medical image perception. In *The Handbook of Medical Image Perception and Techniques*, ed. E Samei, EA Krupinski, pp. 188–96. Cambridge, UK: Cambridge Univ. Press
- Wolfe JM, Friedman-Hill SR, Bilsky AB. 1994. Parallel processing of part/whole information in visual search tasks. *Percept. Psychophys.* 55:537–50
- Wolfe JM, Friedman-Hill SR, Stewart MI, O'Connell KM. 1992. The role of categorization in visual search for orientation. *J. Exp. Psychol. Hum. Percept. Perform.* 18:34–49
- Wolfe JM, Gancarz G. 1996. Guided Search 3.0: a model of visual search catches up with Jay Enoch 40 years later. In *Basic and Clinical Applications of Vision Science*, ed. V Lakshminarayanan, pp. 189–92. Dordrecht, Neth.: Kluwer Acad.
- Wolfe JM, Horowitz TS. 2004. What attributes guide the deployment of visual attention and how do they do it? *Nat. Rev. Neurosci.* 5:495–501
- Wolfe JM, Horowitz TS. 2017. Five factors that guide attention in visual search. *Nat. Hum. Behav.* 1:0058
- Wolfe JM, Horowitz TS, Kenner NM. 2005. Rare targets are often missed in visual search. *Nature* 435:439–40
- Wolfe JM, Horowitz TS, Palmer EM, Michod KO, VanWert MJ. 2010a. Getting into Guided Search. In *Tutorials in Visual Cognition*, ed. V Coltheart, pp. 93–120. Hove, UK: Psychol. Press
- Wolfe JM, Myers L. 2010. Fur in the midst of the waters: visual search for material type is inefficient. *J. Vis.* 10:8
- Wolfe JM, Palmer EM, Horowitz TS. 2010b. Reaction time distributions constrain models of visual search. *Vis. Res.* 50:1304–11



- Wolfe JM, Van Wert MJ. 2010. Varying target prevalence reveals two dissociable decision criteria in visual search. *Curr. Biol.* 20:121–24
- Wolfe JM, Vo ML, Evans KK, Greene MR. 2011b. Visual search in scenes involves selective and nonselective pathways. *Trends Cogn. Sci.* 15:77–84
- Wu C-C, Wolfe JM. 2019. Eye movements in medical image perception: a selective review of past, present and future. *Vision* 3:32
- Yantis S. 1993. Stimulus-driven attentional capture. *Curr. Dir. Psychol. Sci.* 2:156–61
- Yu X, Geng JJ. 2019. The attentional template is shifted and asymmetrically sharpened by distractor context. *J. Exp. Psychol. Hum. Percept. Perform.* 45:336–53
- Zelinsky G. 2008. A theory of eye movements during target acquisition. *Psychol. Rev.* 115:787–835
- Zelinsky GJ, Sheinberg DL. 1997. Eye movements during parallel/serial visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 23:244–62
- Zhang X, Huang J, Yigit-Elliott S, Rosenholtz R. 2015. Cube search, revisited. *J. Vis.* 15:9

2.24 Wolfe



Review in Advance first posted on April 22, 2020. (Changes may still occur before final publication.)