



Major issues in the study of visual search: Part 2 of “40 Years of Feature Integration: Special Issue in Memory of Anne Treisman”

Jeremy M Wolfe^{1,2}

Published online: 10 April 2020
© The Psychonomic Society, Inc. 2020

When we announced a special issue of Attention Perception and Psychophysics to mark the 40th anniversary of Feature Integration Theory (A. Treisman & Gelade, 1980), we might have been getting a bit ahead of ourselves. It was 2018 and Treisman’s seminal paper had come out in 1980. In the end, our timing was perfect. Here it is, 2020, and here we offer up the second part of “40 Years of Feature Integration: Special Issue in Memory of Anne Treisman”. The choice to put a paper in the first or second issue was driven by acceptance date, rather than, for example, topic. Thus, many topics appear in both issues. Within this second issue, papers are organized (perhaps, imperfectly) by topic. This gives me the chance to highlight four topics that were of importance to Anne Treisman during her career and that remain important, as evidenced by the interesting papers they have generated here. Those topics are:

1. Feature search
2. Binding of Features
3. The Mechanics of Search
4. Ensemble statistics and non-selective processing

This introduction provides the opportunity to say a word about each paper but also to draw some connections between papers and to say a few words about the enduring interest in each of these topics. The first citation of each of the articles in the special issue is marked with an asterisk.

✉ Jeremy M Wolfe
jwolfe@bwh.harvard.edu

¹ Ophthalmology & Radiology, Harvard Medical School, Boston, MA, USA

² Visual Attention Lab, Department of Surgery, Brigham & Women’s Hospital, 65 Landsdowne St, 4th Floor, Cambridge, MA 02139, USA

Feature search

One of the pillars of Treisman’s Feature Integration Theory (FIT) was the idea that a limited set of basic features were processed in parallel, across the visual field or, at least, across large swathes of the field in a single step. The classic evidence for this was that the slopes of RT x set size functions were near zero for simple feature searches. Earlier models (including versions of my Guided Search model (Wolfe, 1994; Wolfe, Cave, & Franzel, 1989) tended to treat this feature processing step as essentially instantaneous and undifferentiated across tasks. That can’t be right. An intuitive appreciation of this fact can be obtained from Fig. 1.

In each panel, the search for a yellow ring is quite trivial. However, it should be clear that A will be a bit faster than B and B, in turn, will be a bit faster than C. In A, the target differs from the distractor items by color, shape/topology (Chen, 2005), and size. In B, only the shape is different. In C, the addition of blue distractors will slow search, even though the added stimuli are less similar to the target (J. Duncan & Humphreys, 1989).

How should we understand this? Even a simple perceptual decision (e.g. red vs green) is likely to be the product of some information accumulation process that can be modeled as something like a diffusion process (Purcell et al., 2010). That process might be very fast, but it will take some amount of time. Moreover, bigger set sizes will require more diffusers. Each of those might make an error if made its decision too fast, so decision thresholds should be a bit more conservative for bigger set sizes. This means that bigger set sizes should produce slower results (Leite & Ratcliff, 2010). In practice, RT x set size functions for feature searches do have slopes near zero msec/item, but, as a general rule, those slopes are not actually zero. Over the past few years, Alejandro Lleras and his lab have provided some of the most interesting data and theorizing about this issue. They found that RTs in simple feature searches increase in a reliably logarithmic manner with set size (Buetti, Cronin, Madison, Wang, & Lleras, 2016). In a series of papers, they have sought to explain feature searches

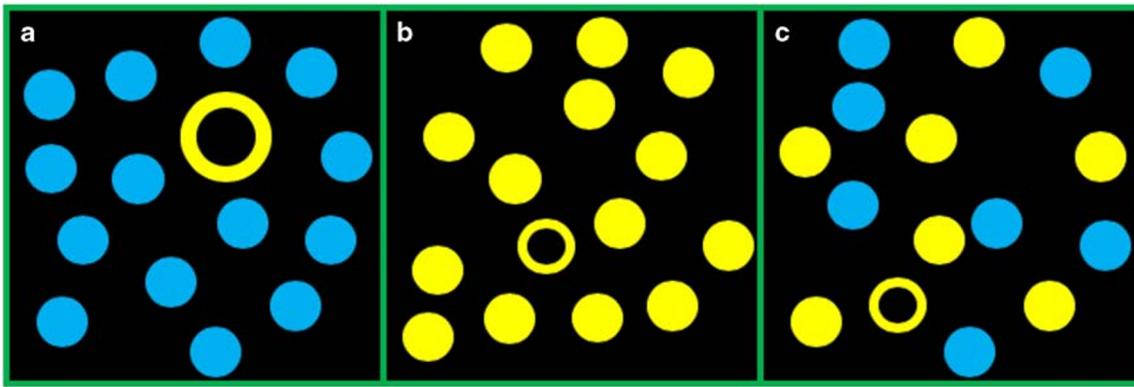


Fig. 1 Three “pop-out” searches. These are all easy, but they don’t seem to be equally easy.

in more complex collections of items (like Fig 1c). In the current issue, Lleras et al. (2020)* offer up their “Target Contrast Signal Theory” that stresses the role of comparison between a target template and each item in a search display. This can be seen as related to work from other labs, in other contexts, that also stresses the importance of the contrast between items over the absolute value of items (S. I. Becker, Harris, York, & Choi, 2017; Yu & Geng, 2019).

Glavan, Haggitt, and Houpt (2020)* bring a different methodology to understanding the processing of the most basic of basic features (in this case color and shape). Like Lleras et al., Glavan et al. note that the simple metric of near zero slopes in visual search experiments is not definitive proof that a feature is processed in parallel over multiple items at the same time. Their work employs Townsend and Nozawa’s “System Factorial Technology” (James T. Townsend & Nozawa, 1995). System Factorial Technology is Townsend’s response to the difficulty in making claims about serial and parallel processing (J.T. Townsend, 1971; J. T. Townsend, 2016). It is intended to provide a more rigorous way to discriminate between serial and parallel processes. Here, Glavan et al. have applied that SFT technology to the question of whether color and shape are processed in parallel at the same time. They conclude that SFT supports the parallel processing of more than one feature at a time. This is a relief since many models rely on that assumption.

A core aspect of Treisman’s FIT is the idea that preattentive features are, well, *preattentive*. Those featural properties can be extracted and used prior to the time that the item with that feature is selected for further processing. The paper by Bergmann, Tünnermann, and Schubö (2019)* can be seen as supporting that idea. Their observers were looking for a target of a specific shape. Nevertheless, the distribution of color in the display influenced the course of the search. In their task, Os looked for any diamond among circles. Two diamonds were present, in two different colors. The Os only needed to find one. The ratio of the colors changed over time and Os responded by favoring the smaller color subset,

showing that their behavior was shaped by a feature that they had not been instructed to attend to.

Can a basic, preattentive feature be built from other basic features? If you look at Fig. 2, you will see four rectangles in a sea of smaller semi-circles. Two of these rectangles are defined by luminance edges. The other two are “subjective contours”, induced by the alignment of some of the semi-circles. Are those rectangles as *real* as the luminance-defined rectangles? Are they treated equivalently by later visual processes? In terms of visual search, Li, Cave, and Wolfe (2008) produced a somewhat mixed result, with some types of subjective contours producing easier search than others.

In this issue, Trick (Trick & Hardy, 2019)* extends this question to enumeration tasks as well as replicating the Li et al. (2008) search results. In her search tasks, performance depends on the task and on the specific nature of the subjective contours. The simple presence of a subjective contour may be preattentively available while the orientation of that contour may be less effective in guiding search.

Zupan and Watson (2019)* also follow up on Li et al. (2008). In their work, they are asking how subjects contours, of the sort illustrated in Fig. 2 function in a Watson and

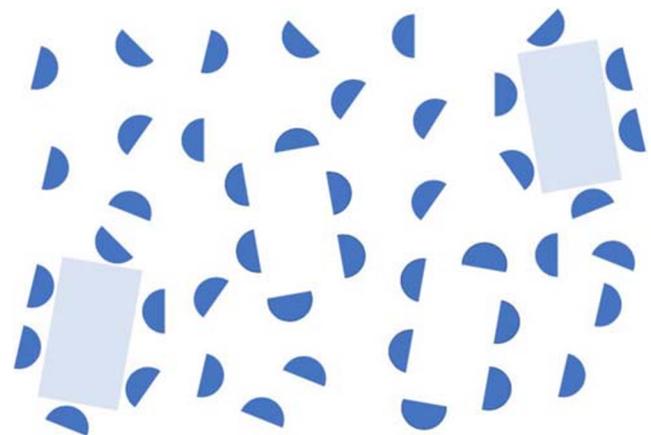


Fig. 2 Rectangles defined by luminance and by subjective contours. Can observers search for a subjective rectangle as effectively as they search for a luminance-defined rectangle?

Humphreys-style preview experiment (Watson & Humphreys, 1997). In those experiments, an initial glimpse of half of the stimuli from a search display tells the observer about objects/locations that they can ignore in the full display. Zupan and Watson find that subjective contours can be used for this purpose but they seem subjected to more stringent capacity limitations than other display elements defined by color or luminance borders. That is, it is easier to ignore eight previewed red rectangles than it is to ignore eight previewed subjective rectangles.

When we talk about parallel processing of preattentive features like color, shape, or orientation, we tend to speak of analysis of those features across the entire visual field in a single processing step. Of course, we know that acuity and crowding effects will limit processing in the periphery (Rosenholtz, 2020; Westheimer, 1982) but if we control for those factors by using appropriately scaled stimuli (Levi, Klein, & Aitsebaomo, 1985), the usual assumption is that a “pop-out” target would pop-out anywhere. Casteau and Smith (2019)* show that this is not quite true. Your visual field extends more than 90 deg to the left and right of straight ahead. However, you cannot make a saccadic eye movement to the edge of the field. You have an “effective oculomotor range” (EOMR), which is the range locations that can be reached with a “saccadic eye movement that does not also require a head movement” (Casteau & Smith, 2019). It turns out that you pay a reaction time (RT) cost if you are responding to a feature singleton outside the EOMR. Interestingly, you do not pay a cost if you are responding to a conjunction of these features. Conjunction tasks are slower overall but they do not show a specific penalty for crossing the EOMR. Overall, the message from the feature search papers in this special issue is that even the simplest feature searches are not as simple as we may have thought. If we just assume that a feature is processed more or less instantly across the visual field, regardless of what else is in that field, we will be ignoring important aspects of preattentive visual processing.

Binding of Features

The “integration” aspect of Treisman’s Feature *Integration* Theory has to do with the combination of basic features. If one piece of the visual system was processing color and another was processing shape, how did those two bits of information get together to give rise to the accurate perception of a colored shape? Figure 3 is an example of one of Treisman’s favorite ways of illustrating this question.

After a quick look at Fig. 3, you can answer many questions successfully. Is there a ceramic pig? (Yes) Is there an apple? (No). You can accurately list most if not all of the colors and shapes. Without looking back at the image, you are likely to be less accurate if asked, “Is there a small, red

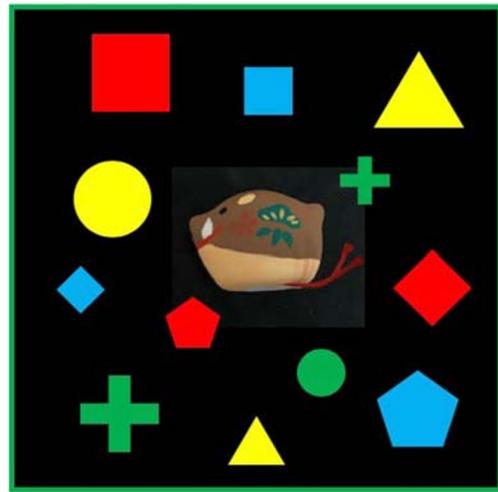


Fig. 3 Have a quick look at this image and then read the relevant text.

circle?” (No) or “Is there a big, blue pentagon?” (Yes). You saw all the colors and shapes, but Treisman would say that you did not “bind” color to shape and size in a way that allows you effortlessly attest to the presence or absence of combinations of features (A Treisman, 1996). If you thought you did see a small, red circle, you experienced what Treisman would call an “illusory conjunction” (A. M. Treisman & Schmidt, 1982).

Illusory conjunctions occur when a feature or features of an object in one location become associated with an object in another location. Reuther and colleagues wondered if this failure of veridical perception would get worse when the eyes were moving or planning to move (Reuther, Chakravarthi, & Hunt, 2020)*. It seems reasonable to imagine that the act of moving the eyes might make it more difficult to see features correctly tied to the correct location (see for example Dowd & Golomb, 2018) but Reuther et al. found that planning and carrying out eye movements didn’t interfere with observer’s ability to bind features.

Back in the 1990s, we had found that it was hard to search for an object defined by two colors (Wolfe et al., 1990). For example, a search was quite inefficient if the target was square divided into red and blue halves among distractors that were red-yellow and yellow-blue. The same task became markedly easier if observers search for a red object with a blue part among red objects with yellow parts and yellow objects with blue parts (Wolfe, Friedman-Hill, & Bilsky, 1994). We argued that something about the part-whole structure must be available preattentively. In their paper in the present issue, Vul, Rieth, Lew, & Rich, 2020 use illusory conjunctions to address the question of how much information about the structure of an object is available before attention reaches that object. They present a ring of two-part objects; for example, circles of one color with a smaller circle of another color inside it. They cue one item in the array and then ask the observer for the colors of the whole and the part. There are lots of ways to get the answer wrong. You could randomly name a color. You

could swap the color of the whole and the color of the part, suggesting that you had loosely bound the right colors to the object. The most interesting errors occur when the color from a neighboring object intrudes into the answer. If you get the color of the central part of the target object wrong, are you more likely to substitute the color of a neighboring central part or are you just as likely to get an intrusion from the color of the whole neighboring object? (This will all be clearer when you look at the pictures in the Vul et al paper). It turns out that part-to-part intrusions happen more often than chance would predict. For Vul et al., this shows that, prior to the arrival of attention, the object is not just a random collection of features. Instead, the preattentive objects must have some structure in which there is at least some information about whether a feature is the feature of a part or of the whole object. The Vul et al. dataset is rich and the paper also serves as a good example of how the right visualizations can help us to extract information from large volumes of data.

Henderson and McClelland (2020)* think that we need to distinguish between two kinds of illusory conjunctions. One type is due to crowding from near neighbors. The other, that they call ‘distal’, involves feature swaps between items that are quite far from each other and must involve some mechanism other than crowding. To look at these distal illusory conjunctions, they did a series of variations on a classic illusory conjunction experiment design. In the classic design, observers report on the properties of a couple of items (e.g. digits) that flank some other items (e.g. colored letters). Illusory conjunctions are found when observers are asked about those items in the middle. Treisman thought that spreading attention between the two flanking items led to misbindings in between. In the Henderson and McClelland studies, the position of two colored letters is varied. Sometimes they are inside the flanking digits. Sometimes they are outside. The observers are queried about one, target letter and the authors look for illusory conjunctions with the properties coming from the other, distracting letter. The critical finding is that illusory conjunctions occurred more often when the target letter was inside the flankers – as would be expected - but it really didn’t matter where the distractor letter was located. That is a bit hard to explain in the standard account. Their new account, involving attentional ‘shadows’ assumes that more than one item is processed at one time. We might be detecting a bit of a theme here. Like Vul et al., Henderson and McClelland are arguing for a preattentive world that is more complicated and/or structured than the unbound collection of features that Treisman sometimes proposed.

Part of the interest in putting together a special issue is that papers can ‘talk’ to each other in ways that might not have occurred without their accidental proximity in the issue. So Henderson and McClelland (2020) are proposing that those attended digits in an illusory conjunction experiment are casting attentional shadows in the direction from digit to fixation

(see their Figure 7). In the next article, Ceja, Jardine, & Franconeri, (2019)* present a new version of the illusory conjunction phenomenon. As noted, illusory conjunctions seem to occur more frequently when attention is spread out across the scene. Ceja et al. find that when attention is spread out, unattended objects or their features seem to migrate toward fixation; a phenomenon that they call “foveal gravity”. You can have a look for yourself at <https://osf.io/2bndg/>. Perhaps there is a connection between the foveal gravity of Ceja et al. and the attentional shadows of Henderson and McClelland or, perhaps, this is nothing but an illusory conjunction of ideas, produced by the proximity of the two articles. The reader can decide.

The final paper in this group concerned with binding explores the relationship between visual binding and mnemonic binding. Fiacconi, Cali, Lupiáñez, and Milliken (2019)* pit memory for the binding of a letter to a location against the perception of the current binding of the same or a different letter to the same or a different location. They find that factors that make the perceptual task faster make the memory task more accurate. Appropriately for a Treisman special issue, they tie the results together via the concept of an object file.

The Mechanics of Search

The largest group of papers in this Special Issue deal with the mechanics of search. How do search tasks unfold over time? What influences the slope of RT x set size functions, the standard measure of search efficiency? The basic form of Feature Integration Theory had a pretty simple story about search slopes. If an adequately salient, basic feature defined the target, search slopes would be near zero. All other searches would require serial search with slopes reflecting a rate of attentional deployment of a minimum of about 20-30 msec/item for target present trials and twice that for target absent trials. Slopes would be steeper if each item in the serial search took longer to process.

Some theoretical complications are illustrated in Fig. 4. In each panel, Os would be looking for a black “T”. Let’s suppose that the search stimuli in A produced the (hypothetical) data labelled “A” in the fourth panel. Search would be more efficient, if half the items were red and could be ignored (B) (Egeth, Virzi, & Garbart, 1984). Search would also be more efficient if the distractors were more easily rejected (C). Guidance by color (B) and speed of distractor rejection (C) are two different mechanisms but they could produce similar looking results. Horstmann, Becker, and Grubert (2020) investigate this issue and suggest that there are clues in eye movement data that can help to distinguish different routes to search efficiency.

In their review of the work on search for facial emotion, D. Vaughn Becker (as distinct from Stefanie Becker in

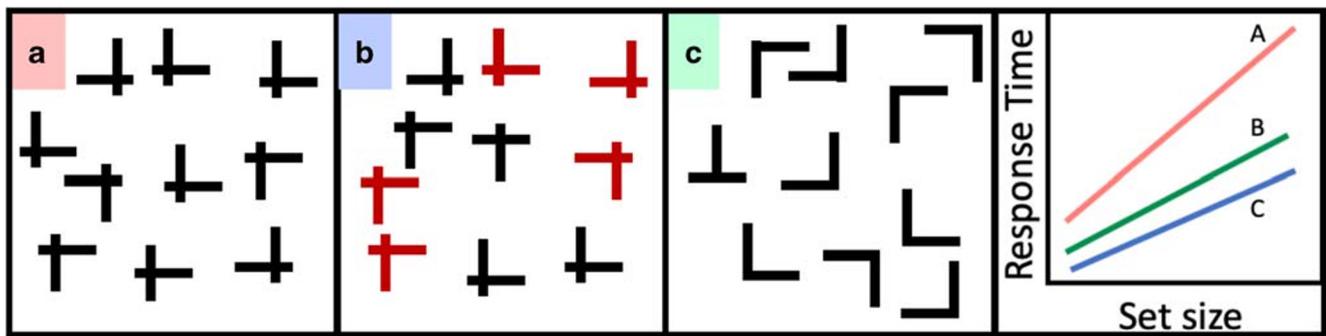


Fig. 4 In each of these searches, the target is a rotated T. The final panel cartoons the RT x set size functions for target present trials that might arise we used these stimuli in an actual experiment.

Horstmann, Becker and Grubert) and Rheem make use of the ideas about distractor rejection to explain why some have thought that angry faces “pop-out” in search displays (Becker & Rheem 2020*). Basically, the argument is that angry faces don’t grab attention but, once attended, an angry face holds attention a bit longer. Thus, it is harder to find a happy face among angry faces than an angry face among happy distractors because the angry distractors are holding your attention. Becker and Rheem also review the back and forth history of claims about face search followed by claims that the Os were using low-level features that had nothing to do with face followed by new stimuli, intended to correct the problem, and so forth. Honest disagreement is possible on this topic. I am inclined to think that faces might ‘pop-out’ from among non-face stimuli (Hershler & Hochstein, 2005) though one can argue about why (Hershler & Hochstein, 2006; VanRullen, 2006). Becker and Rheem’s review provides support for the view that, even if faces are a special class of stimulus, one kind of face does not pop-out from among other faces, as long as the low-level features are controlled.

Stefani, Sauter, and Mack (2020)* are investigating what can be seen as some of the internal workings of distractor rejection or, more specifically, of the ability to get attention and/or the eyes off of a distractor and on to a target. They have Os looking at a ring of discs. At some point, the discs are going to acquire colors and the Os are intended to report on the one disk with a target color. If an irrelevant disk at fixation has the target color, too, it takes a bit longer to move off that disk. Building on earlier work by Wright, Boot, and Brockmole (2015), they show that the original “Functional fixedness” effect generalizes to a range of conditions.

One of the sources of confusion in the search literature has to do with what we know about items that are not currently fixated in a visual search. Because attention typically precedes fixation on an object (Kowler, Anderson, Doshier, & Blaser, 1995), it is tempting to use fixation as a marker for the deployment of attention. At some level, everyone knows that this is incorrect. You can easily fixate on THIS word, while attending to some object or location away from fixation. Still, it could be the case that, in visual search, unless fixation is restricted,

attention and the eyes travel together. If overt movements of the eyes track covert deployments of attention in search, there are problems that arise. The rate of processing is one such problem. However, one interprets the slopes of RT x set size function, they can be read as measures of the rate at which items can be processed in search. So, a slope of 50 msec/item indicates that 20 items can be processed in a second. It doesn’t tell you how those 20 items were processed. It is just the rate.

In that one second, only 3–4 items will be fixated. It follows that more items are being processed than are being fixated, so it must be possible to process items away from fixation. Guided Search would argue that covert deployments of attention happen at a faster rate than overt deployments of the eyes and can visit items away from the point of fixation. Other accounts like the Functional Visual Field account of Hulleman and Olivers (2017) would hold that all of the items within some region (the Functional Visual Field or FVF) surrounding fixation would be processed in parallel. For present purposes, the important point is that both views would require that the meaning of items away from fixation must be processed during search. Cimminella, Della Sala, and Coco (2020)* provide some new evidence that this is the case. Using search for photographs of real objects, they found that RTs and eye movements could be driven by the meaning/ semantics of the objects, independent of their low-level features. These data would falsify models that permit only basic feature processing to occur away from fixation. They do not tell us exactly how meaning is extracted from unfixated items during search but they make it clear that searchers make use of those meanings.

Again, juxtaposition in the special issue highlights aspects of papers that might not have been as clear if the papers were seen in isolation. Nurit Gronau’s (2020)* search tasks involve only two items and a search for a target defined by its category (e.g. animals, clothes, transportation). The critical manipulation is whether attention is directed to both items or only to one. The key finding is that the categorical status of the items interacts when both items are attended but not if one of them is left out. This can be seen as refining what it means to process an item away from the point of fixation. It is not simply that

every item is processed if it is not too small and not too crowded. Clearly, the deployment of attention is critical when Gronau manipulates that deployment. During a less constrained visual search, it is still true that the deployment of attention is important but the nature of the deployment remains a topic of debate.

The Glavan et al. (2020) paper, discussed earlier, pointed out that there are limits to theorizing about search mechanisms, if you just rely on simple average RT data. The full distributions of RTs can be used to constrain models in ways that the average RTs alone cannot. With that in mind, Palmer, Horowitz, Torralba, and Wolfe (2011) et al. posted a big set of search RTs online. Panis, Moran, Wolkersdorfer, and Schmidt (2020)* have used those data for a new analysis based on Harzard functions. Their analysis highlights a point that must be true but that doesn't get enough attention. RTs cannot be simply divided into discrete pieces of time: this much time for search, this much time for decision, etc. A variety of operations must be occurring alongside attentional selection processes during search. Panis et al. see evidence for "object recognition, learning, and cognitive control processes" occurring at the same time. If our minds couldn't manage to do this sort of multitasking, we would never get anything done.

Reimer and Schubert (2020)* remind us of a different way in which search tasks are composed of multiple, simultaneous moving parts. Their emphasis is on the fact that attention is not a single thing, reflecting a single capacity limitation. There are multiple attentional mechanisms. There is a visual selection mechanism that is important in visual search. It is not the same as the multisensory attentional process that allows you to switch between visual stimuli as you read this, auditory stimuli as the phone rings, or tactile stimuli if you think about the pressure of your posterior on the chair. Reimer and Schubert use a dual task paradigm to show that demanding conjunction tasks involve visual attention but also a central capacity that is needed for auditory tasks as well (They didn't check tactile stimuli.).

Marcia Grabowecky was a student of Treisman's and her contribution (Smith & Grabowecky, 2020)* builds on work that she began many years ago with Anne. The question has to do the mechanics of search. When a stimulus appears on the screen, where does attention go first? In the case of simple pop-out stimuli, attention simply goes to the target. In more difficult tasks, Smith and Grabowecky find that attention prefers the center of mass of the stimulus array. They presented arcs of stimuli where all items are equidistant from fixation and their data show an RT benefit at the center of the arc. This might be an example of attention heading to where the chance of collecting useful information is the greatest (see, for example, Zelinsky, Rao, Hayhoe, & Ballard, 1997).

The first articles in this issue were investigating the basic, preattentive features that are available to guide search.

However, just because a feature can guide attention doesn't mean that it should on every occasion. For instance, if the target is blue and all the distractors are blue, there is not much point to creating an attention-guiding priority map that highlights all the blue items. You want to guide to diagnostic features – the ones that provide information about the current task. This idea is part of the Guided Search model (Wolfe, 1994) and it is central to the dimension-weighting approach of Herman Muller's group (Found & Muller, 1996; Liesefeld & Müller, 2019). In the current issue, (Lee & Geng, 2019)* find that you can do this weighting on the fly. That is, if you know that the target is blue, you will prioritize blue. If you don't know, then your initial look at the stimulus will tell you if color is likely to be diagnostic and you can and plan your weighting on that basis. The ability to gain information about what to weight from the initial appearance of a stimulus is related to ensemble perception, the topic of papers that appear later in this issue.

Before we get there, however, let us consider how much of a brain it takes to perform these basic visual search tasks. The answer is that the small brain of the archer fish will do the task in a surprisingly human fashion (Reichenthal, Segev, & Ben-Shahar, 2020)*. Among their other virtues, the experiments, described here, are methodologically wonderful. The archer fish is so-named because it shoots bugs off of leaves by spitting a stream of water at them. In the lab, the fish can be persuaded to spit at targets on a computer monitor, placed above a fish tank. The RTs are longer than human RTs and you are not going to collect 1000 trials an hour from a fish, as you might in a human study. Nevertheless, the data from spitting fish can successfully generate RT x set size functions for the fish. The results show that the pattern of results for feature and conjunction searches in humans is approximated well by the fish. The other virtue of this paper is that it reviews visual search results from a wide variety of non-human species. It is a testimony to the importance of visual search as a behavior and to the ingenuity of the research community that we have data from dung beetles, bees, cuttlefish, pigeons, monkeys and more.

Grieben, Tekülve, Zibner, Lins, Schneegans, & Schöner (2020)* offer a new model to implement search which they test using conjunction tasks not unlike those used with the archer fish. As the authors say, the Grieben et al. model is in the tradition of FIT and Guided Search. It is distinguished from those models by being based on dynamic field theory. Basically, search behavior emerges from the right set of neurally-inspired equations. This is distinct from Guided Search. Guided Search provides a wiring diagram of processing steps. However, it tends to assert that a process-like template matching occurs. It does not give much in the way of detail about how to implement such a process. The Grieben et al. paper does propose those details. The paper is particularly concerned with the interaction of memory and search. In

their model, it should be the case that, if you have been looking at a scene, you will be more efficient at finding a target in the scene than if you had not previewed the scene. Your memory should guide your attention. It has been hard to find evidence of that improved efficiency. They provide some evidence. They propose that earlier failures had at least two causes. One is methodological. If you show a preview, take it away, and then show the search array, the system treats the search array as a new stimulus and there is no benefit. The other reason is that the benefit is a working memory benefit. Working memory is limited in capacity so the effects are small. Thus, you might miss a benefit because your experiment was underpowered. Actually, I don't think they exactly make this point about statistical power, but that was the lesson I took home for why some of my earlier efforts in this realm might have failed to find benefits (Wolfe, Klempe, & Dahlen, 2000). Regardless, the architecture described by Grieben et al. looks like it could yield explanations for many effects and it might yield predictions for new, interesting search phenomena.

Figure 5 illustrates a set of questions that are relevant to the next few articles in this issue. Starting from the left, search for vertical lines in each patch and, then, move to the next panel when you are done.

One question, illustrated here is the question of when it is time to terminate a search. This quitting question was not a major issue in the early days of FIT. In FIT terms, parallel searches for feature targets ended when that single step, parallel process returned “no” for an answer. All other searches were assumed to be serial and more-or-less exhaustive. When search was less exhaustive, observers would make more errors; a form of speed-accuracy tradeoff. With the rise of models having a less dichotomous serial/parallel structure, quitting rules became a topic of research (Chun & Wolfe, 1996; Moran, Zehetleitner, Mueller, & Usher, 2013; Wolfe, 2012), represented in the current issue by Shi, Allenmark, Zhu, Elliott, and Müller (2019)*. In a standard search task, where observers determine if a target is present or absent, the quitting problem is

the problem of when to terminate target-absent trials (like, for example, panel 3 of Figure 5). In their paper, Shi et al. are revisiting a problem we introduced in Horowitz and Wolfe (1998). If the target absent searches for a task were really, serial and exhaustive, then the observer would need to have some way of keeping track of which distractor items had been examined. Otherwise, they would not know when they had exhausted the set. To thwart this marking of rejected distractors, Horowitz and I replotted every item on the screen every 110 msec. This was a bit like searching for the target “T” in a snowstorm of “L”s. Surprisingly, this made almost no difference to the slopes of the target present trials. We argued that “visual search has no memory” by which we meant that observers were not, in fact, keeping track of the rejected distractors.

Substantial debate ensued. For instance, there is Peterson, Kramer, Wang, Irwin, and McCarley's (2001) response, entitled “Visual search has memory”. However, that is not the issue here. Shi et al. (2019) focus on the absent trials in dynamic search. Normally, the slope of target absent trials is roughly twice the slope of present trials. Indeed, that 2:1 ratio was part of Treisman's original evidence for serial, self-terminating searches. In Horowitz and Wolfe (1998), our slope ratios were more like 1:1 as they are in Shi et al's replication. We didn't exactly sweep that fact under the rug but nor did we really deal with it. Shi et al. point out that, for target absent trials, our dynamic task is not strictly comparable to the standard static search. The two tasks are asking different questions. In a static search, the question to ask before quitting is something like “Have I looked at more or less everything?” With everything randomly appearing and disappearing in dynamic search, that question doesn't make sense. In dynamic search, the right question is more like “Have I searched long enough?”. Am I reasonably sure that the target is not going to turn up if I keep looking? Shi et al. go from that observation to some clever modeling that bears on the general question of quitting; not just the specific issue of quitting in dynamic search tasks.

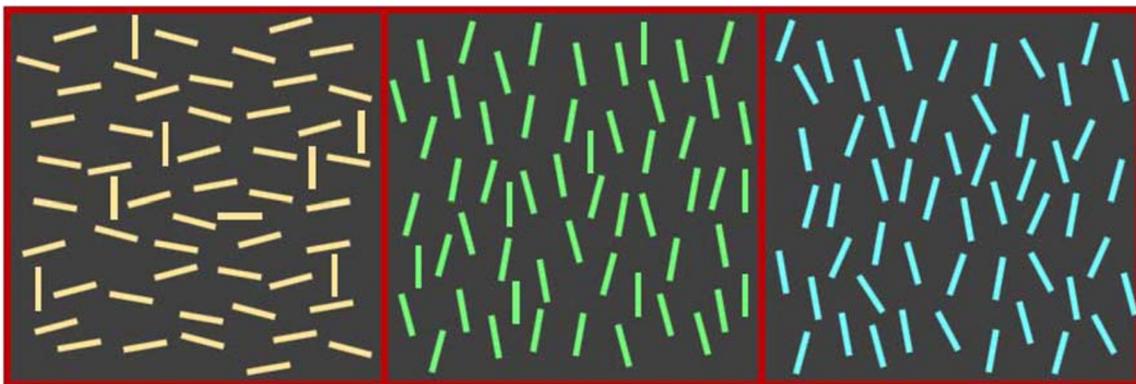


Fig. 5 Working from left to right, look for vertical lines

Returning to Fig. 5, the task illustrated here is not a standard laboratory search because there is an unknown number of targets present as opposed to the usual zero or one. These multitarget tasks are versions of foraging tasks, borrowing the term from the animal foraging literature (Stephens & Krebs, 1986). Foraging has attracted some recent interest in the human search literature (Cain, Vul, Clark, & Mitroff, 2012; Wolfe & Danielson, 2012) including from Arni Kristjánsson's group (Jóhannesson, Thornton, Smith, Chetverikov, & Kristjánsson, 2016; Á. Kristjánsson, Jóhannesson, & Thornton, 2014; Kristjánsson & Kristjánsson, 2018) and it is that group that contributes a foraging study to this issue (Kristjánsson, Björnsson, & Kristjánsson, 2019)*. The questions in foraging studies are a bit different than those in standard search. In foraging studies, we want to know the rate with which targets are acquired, the time when observers choose to leave one patch/trial to move on to the next, and the number of items that are left behind when the observer leaves the patch. In Kristjánsson's versions of these tasks, observers look for instances of two different target types; for example, red and green dots among blue and yellow dots. He has been interested in the pattern of picking. Did observers pick one type of target first and then go back for the other or did observers alternate at random between the two types. In earlier work, Kristjánsson required observers to pick every target so the questions of patch leaving time and error rate were not addressed. In the new paper, published here, observers are free to leave the patch when they will and the paper describes how responses change when observers have this freedom.

Ensemble statistics

Figure 5 illustrates one more topic that was of interest to Treisman and has become a topic of increasing interest today. Looking at the second panel in the figure, even before we locate specific vertical targets, we know that the overall orientation of that region is vertical. That vertical 'gist' will influence how long we search for specifically vertical items as can be seen by comparing that second panel to the first panel where the overall horizontal structure is immediately apparent. Clearly, we can glean information about the average value of a feature over a region. We are also sensitive to the variance of a feature, as a comparison of panel two with panel three will reveal. This ability to assess the global properties of a scene or a region is typically referred to as gist (Oliva, 2005) or ensemble (Whitney & Yamanashi Leib, 2018) processing.

The paper by Corbett and Munneke (2019)* sits at the intersection of research on ensembles and visual search. The core observation is that, over trials, search is faster if the statistics of the scene remain stable. This is true even if what you are looking for (e.g. conjunction of spatial frequency and

orientation) is unrelated to the variation being manipulated (e.g. size of the items). In a pair of papers, Aysecan Boduroglu and co-authors look at interactions between properties in ensemble tasks. In the first paper, they show that observers can make both centroid and center-of-mass judgments about sets of disks of different sizes. Observers' assessments of the sizes of dots influence their estimates of the position of the centroid, even though, in principle, the centroid estimate should not make use of size (Boduroglu & Yildirim, 2020)*. The second paper presents a series of experiments showing that observers can make independent decisions about the average size and orientation of ensembles of items, suggesting that independent underlying mechanisms support those decisions (Yoruk & Boduroglu, 2019)*.

Raidvee, Toom, Averin, and Allik (2020)* are investigating the limits on ensemble processing. If you have several circles, how good are you at assessing their mean diameter, their total area, or the total of their diameters? Mean diameter is an average size judgement. Observers are quite good at that. They are not accurate when it comes to the perception of total length or cumulative area. The authors draw a line to the interpretation of graphs (e.g. bubble charts), suggesting that it is the diameter of the dot that should be linearly related to the underlying data and not the area of the dot.

Takano and Kimura (2019)* are studying judgements about mean luminance and finding that observers are biased toward the brighter disks in a display if they are asked to choose the brighter of two arrays and biased toward darker disks if asked to choose the darker of the same two arrays. Their work has interesting parallels with texture segmentation work by Chubb and colleagues (e.g. Chubb & Landy, 1994; Chubb, Landy, & Econopouly, 2004) which raises an interesting question; do the panels of Fig. 5 show "ensembles" or "textures". Ensemble research is typically done using displays with relatively few elements. Textures are often composed of many (think of the average orientation of a field of wheat). Do textures and ensembles produce the same patterns of results in judgements about properties like average luminance.

Ensemble perception can be taken as evidence for two pathways to visual awareness (Wolfe, Vo, Evans, & Greene, 2011). Returning to Panel Two of Fig. 5, an attentionally selective pathway allows you to determine that some specific items are vertical while others are tilted. At the same time, a non-selective pathway (or in other views, a state of global or diffuse attention) allows you to determine that the entire ensemble is oriented vertically. In a rather different way, the penultimate paper in this issue examines the interaction of these realms. Duncan and Van der Stigchel (2019)* are doing change detection experiments and finding that we are more likely to guess that something has changed if the background has changed, even if that is not relevant. They argue that "researchers have tended to ignore inattentional visual processes" and they are, no doubt, correct that that vast body of

vision research concerns perception of attended stimuli. Their paper is a useful reminder that a lot of important vision is happening outside of the current “spotlight” of attention.

Coda

I am quite sure that Anne Treisman would have enjoyed the papers assembled in these two volumes. She would be pleased that FIT was still inspiring research. She would have argued with those who were ‘disproving’ her theory in a manner that would be both intellectually rigorous and unflinchingly polite and constructive. And she would have absorbed and remembered pretty much all of the work published here. In my early days as an attention researcher, Anne was my world wide web. She had read everything, remembered most of it, and thought about what it all meant. When asked, she was remarkably generous with that knowledge. Google Scholar and the Web of Science are marvelous tools, but I am sorry not to have the Treisman search engine to consult.

References

- Becker, D. V., & Rheem, H. (2020). Searching for a face in the crowd: pitfalls and unexplored possibilities. *Atten Percept Psychophys*.
- Becker, S. I., Harris, A. M., York, A., & Choi, J. (2017). Conjunction Search is Relational: Behavioral and Electrophysiological Evidence. *Journal of Experimental Psychology: Human Perception and Performance*, 43(10), 1828–1842.
- Bergmann, N., Tünnermann, J., & Schubö, A. (2019). Which search are you on? Adapting to color while searching for shape. *Attention, Perception, & Psychophysics*. doi: <https://doi.org/10.3758/s13414-019-01858-6>
- Boduroglu, A., & Yildirim, I. (2020). Statistical summary representations of bound features. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01944-9>
- Buetti, S., Cronin, D. A., Madison, A. M., Wang, Z., & Lleras, A. (2016). Towards a Better Understanding of Parallel Visual Processing in Human Vision: Evidence for Exhaustive Analysis of Visual Information. *Journal of Experimental Psychology: General*, 145(6), 672–707. doi: <https://doi.org/10.1037/xge0000163>
- Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (2012). A Bayesian optimal foraging model of human visual search. *Psychol Sci*(23), 1047–1054. doi:<https://doi.org/10.1177/0956797612440460>
- Casteau, S., & Smith, D. T. (2019). Is pre-attentive search restricted to the range of eye-movements? *Atten Percept Psychophys*, msAT18_044.
- Ceja, C. R., Jardine, N. L., & Franconeri, S. L. (2019). Foveal gravity: A robust illusion of color-location misbinding. *Attention, Perception, & Psychophysics*. doi: <https://doi.org/10.3758/s13414-019-01853-x>
- Chen, L. (2005). The topological approach to perceptual organization. *Visual Cognition*, 12(4), 553–637.
- Chubb, C., & Landy, M. S. (1994). Orthogonal distribution analysis: A new approach to the study of texture perception. In M. S. Landy & J. A. Movshon (Eds.), *Computational Models of Visual Processing* (pp. 291–301). Cambridge, MA: MIT Press.
- Chubb, C., Landy, M. S., & Econopouly, J. (2004). A visual mechanism tuned to black. *Vision Research*, 44(27), 3223–3232.
- Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology*, 30, 39–78.
- Cimminella, F., Della Sala, S., & Coco, M. I. (2020). Extra-foveal Processing of Object Semantics Guides Early Overt Attention During Visual Search. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01906-1>
- Corbett, J. E., & Munneke, J. (2019). Statistical stability and set size exert distinct influences on visual search. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01905-2>
- Dowd, E. W., & Golomb, J. D. (2018). The Binding Problem after an eye movement. *Atten Percept Psychophys*, in press
- Duncan, D., & Van der Stigchel, S. (2019). Constancy bias: When we “fill in the blanks” of unattended or forgotten stimuli. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01838-w>
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 32–39.
- Fiacconi, C. M., Cali, J. N., Lupiáñez, J., & Milliken, B. (2019). Coordinating the interaction between past and present: Visual working memory for feature bindings overwritten by subsequent action to matching features. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01880-8>
- Found, A., & Muller, H. J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a ‘dimension weighting’ account. *Perception & Psychophysics*, 58(1), 88–101.
- Glavan, J. J., Haggit, J. M., & Houpt, J. W. (2020). Temporal organization of color and shape processing during visual search. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01863-9>
- Grieben, R., Tekülve, J., Zibner, S. K. U., Lins, J., Schneegans, S., & Schöner, G. (2020). Scene memory and spatial inhibition in visual search. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01898-y>
- Gronau, N. (2020). Vision at a glance: The role of attention in processing object-to-object categorical relations. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01940-z>
- Henderson, C. M., & McClelland, J. L. (2020). Intrusions into the shadow of attention: A new take on illusory conjunctions. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01893-3>
- Hershler, O., & Hochstein, S. (2005). At first sight: a high-level pop out effect for faces. *Vision Research*, 45(13), 1707–1724.
- Hershler, O., & Hochstein, S. (2006). With a careful look: Still no low-level confound to face pop-out. *Vision Research*, 46(18), 3028–3035.
- Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, 394(Aug 6), 575–577.
- Horstmann, G., Becker, S., & Grubert, A. (2020). Dwelling on Simple Stimuli in Visual Search. *Atten Percept Psychophys*.
- Hulleman, J., & Olivers, C. N. L. (2017). The impending demise of the item in visual search. *The Behavioral and Brain Sciences*, 1–20. doi: <https://doi.org/10.1017/S0140525X15002794>, e132
- Jóhannesson, O., Thornton, I. M., Smith, I. J., Chetverikov, A., & Kristjánsson, A. (2016). Visual foraging with fingers and eye gaze. *i-Perception*, 7(2). <https://doi.org/10.1177/2041669516637279>.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35(13), 1897–1916.
- Kristjánsson, Á., Jóhannesson, O. I., & Thornton, I. M. (2014). Common Attentional Constraints in Visual Foraging. *PLoS One*, 9(6), e100752. doi: <https://doi.org/10.1371/journal.pone.0100752>

- Kristjánsson, A. r., Björnsson, A. S., & Kristjánsson, T. (2019). Foraging with Anne Treisman: Patch leaving, features versus conjunctions and memory for foraged locations. *Atten Percept Psychophys*.
- Kristjánsson, T., & Kristjánsson, Á. (2018). Foraging through multiple target categories reveals the flexibility of visual working memory. *Acta Psychologica*, 183, 108–115. doi: <https://doi.org/10.1016/j.actpsy.2017.12.005>
- Lee, J., & Geng, J. J. (2019). Flexible weighting of target features based on distractor context. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01910-5>
- Leite, F. P., & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception, & Psychophysics*, 72(1), 246–273. doi: <https://doi.org/10.3758/APP.72.1.246>
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, 25, 963–977.
- Li, X., Cave, K., & Wolfe, J. M. (2008). Kanisza-style subjective contours do not guide attentional deployment in visual search but line termination contours do. *Perception & Psychophysics*, 70(3), 477–488.
- Liesefeld, H. R., & Müller, H. J. (2019). Distractor handling via dimension weighting. *Current Opinion in Psychology*, 29, 160–167. doi: <https://doi.org/10.1016/j.copsyc.2019.03.003>
- Lleras, A., Wang, Z., Ng, G. J. P., Ballew, K., Xu, J., & Buetti, S. (2020). A target contrast signal theory of parallel processing in goal-directed search. *Attention, Perception, & Psychophysics*. doi: <https://doi.org/10.3758/s13414-019-01928-9>
- Moran, R., Zehetleitner, M. H., Mueller, H. J., & Usher, M. (2013). Competitive Guided Search: Meeting the challenge of benchmark RT distributions. *Journal of Vision*, 13(8), 24. doi: <https://doi.org/10.1167/13.8.24>.
- Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees & J. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251–257). San Diego, CA: Academic Press / Elsevier.
- Palmer, E. M., Horowitz, T. S., Torralba, A., & Wolfe, J. M. (2011). What are the shapes of response time distributions in visual search? *Journal of Experimental Psychology. Human Perception and Performance*, 37(1), 58–71. doi: <https://doi.org/10.1037/a0020747>
- Panis, S., Moran, R., Wolkersdorfer, M. P., & Schmidt, T. (2020). Studying the dynamics of visual search behavior using RT hazard and micro-level speed–accuracy tradeoff functions: A role for recurrent object recognition and cognitive control processes. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01897-z>
- Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science*, 12(4), 287–292.
- Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychological Review* doi: <https://doi.org/10.1037/a0020311>
- Raidvee, A., Toom, M., Averin, K., & Allik, J. (2020). Perception of means, sums, and areas. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01938-7>
- Reichenthal, A., Segev, R., & Ben-Shahar, O. (2020). Feature integration theory in non-humans: Spotlight on the archerfish. *Attention, Perception, & Psychophysics*. doi: <https://doi.org/10.3758/s13414-019-01884-4>
- Reimer, C. B., & Schubert, T. (2020). Visual and central attention share a capacity limitation when the demands for serial item selection in visual search are high. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01903-4>
- Reuther, J., Chakravarthi, R., & Hunt, A. R. (2020). The eye that binds: Feature integration is not disrupted by saccadic eye movements. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01873-7>
- Rosenholtz, R. E. (2020). What modern vision science reveals about the awareness puzzle: Summary-statistic encoding plus limits on decision complexity underlie the richness of visual perception and its quirky failures. *Atten Percept Psychophys*.
- Shi, Z., Allenmark, F., Zhu, X., Elliott, M. A., & Müller, H. J. (2019). To quit or not to quit in dynamic search. *Atten Percept Psychophys*, in press.
- Smith, M. K., & Grabowecky, M. (2020). Exogenous Orientation of Attention to the Center of Mass in a Visual Search Task. *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01908-z>
- Stefani, M., Sauter, M., & Mack, W. (2020). Delayed disengagement from irrelevant fixation items: Is it generally functional? *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01926-x>
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory*. Princeton, NJ: Princeton U. Press.
- Takano, Y., & Kimura, E. (2019). Task-driven and flexible mean judgment for heterogeneous luminance ensembles. *Atten Percept Psychophys*, *msAt18_052*.
- Townsend, J. T. (1971). A note on the identification of parallel and serial processes. *Perception & Psychophysics*, 10, 161–163.
- Townsend, J. T. (2016). A Note on Drawing Conclusions in the Study of Visual Search and the Use of Slopes in Particular. A reply to Kristjánsson and Wolfe. *i-Perception*, *ms*.
- Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal Properties of Elementary Perception: An Investigation of Parallel, Serial, and Coactive Theories. *Journal of Mathematical Psychology*, 39(4), 321–359. doi: <https://doi.org/10.1006/jmps.1995.1033>
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6, 171–178.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Treisman, A. M., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107–141.
- Trick, L. M., & Hardy, N. L. A. (2019). Does the standard search task predict performance in related tasks for Kanisza-style illusory contours? *Attention, Perception, & Psychophysics* doi: <https://doi.org/10.3758/s13414-019-01890-6>
- VanRullen, R. (2006). On second glance: still no high-level pop-out effect for faces. *Vision Research*, 46(18), 3017–3027.
- Vul, E., Rieth, C., Lew, T. F., & Rich, A. N. (2020). The structure of illusory conjunctions reveals hierarchical binding of multi-part objects. *Atten Percept Psychophys*.
- 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(1), 90–122.
- Westheimer, G. (1982). The spatial grain of the perifoveal visual field. *Vision Research*, 22, 157–162.
- Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. *Annual Review of Psychology*, 69, 105–129. doi: <https://doi.org/10.1146/annurev-psych-010416-044232>
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1(2), 202–238.
- Wolfe, J. M. (2012). When do I quit? The search termination problem in visual search. *Nebraska Symposium on Motivation*, 59, 183–208.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided Search: An alternative to the Feature Integration model for visual search. *Journal of Experimental Psychology. Human Perception and Performance*, 15, 419–433.
- Wolfe, J. M., & Danielson, J. R. (2012). Visual Foraging Behavior: When are the berries riper on the other side of the screen? *J. Vis., Talk given at the annual meeting of the Vision Sciences Society in Naples, FL*.
- Wolfe, J. M., Friedman-Hill, S. R., & Bilsky, A. B. (1994). Parallel processing of part/whole information in visual search tasks. *Perception & Psychophysics*, 55(5), 537–550.

- Wolfe, J. M., Klempen, N., & Dahlen, K. (2000). Post-attentive vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26(2), 693-716.
- Wolfe, J. M., Vo, M. L., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, 15(2), 77-84. doi: <https://doi.org/10.1016/j.tics.2010.12.001>
- Wolfe, J. M., Yu, K. P., Stewart, M. I., Shorter, A. D., Friedman-Hill, S. R., & Cave, K. R. (1990). Limitations on the parallel guidance of visual search: Color X color and orientation X orientation conjunctions. *Journal of Experimental Psychology: Human Perception and Performance*, 16(4), 879-892.
- Wright, T. J., Boot, W. R., & Brockmole, J. R. (2015). Functional fixedness: The functional significance of delayed disengagement based on attention set. *Journal of Experimental Psychology: Human Perception and Performance*, 41(1), 17-21. doi: <https://doi.org/10.1037/xhp0000016>
- Yoruk, H., & Boduroglu, A. (2019). Feature-Specificity in Visual Statistical Summary Processing. *Atten Percept Psychophys*.
- Yu, X., & Geng, J. J. (2019). The attentional template is shifted and asymmetrically sharpened by distractor context. *Journal of Experimental Psychology: Human Perception and Performance*, 45(3), 336-353. doi: <https://doi.org/10.1037/xhp0000609>
- Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatio-temporal dynamics of visual search. *Psychological Science*, 8(6), 448-453.
- Zupan, Z., & Watson, D. G. (2019). Perceptual grouping constrains inhibition in time-based visual selection. *Attention, Perception, & Psychophysics*. doi: <https://doi.org/10.3758/s13414-019-01892-4>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.