

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/295548823>

# Towards a Better Understanding of Parallel Visual Processing in Human Vision: Evidence for Exhaustive Analysis of Visual Information

Article in *Journal of Experimental Psychology General* · February 2016

DOI: 10.1037/xge0000163

CITATIONS

35

READS

1,167

5 authors, including:



**Simona Buetti**

University of Illinois, Urbana-Champaign

46 PUBLICATIONS 360 CITATIONS

[SEE PROFILE](#)



**Deborah A. Cronin**

University of Illinois, Urbana-Champaign

8 PUBLICATIONS 40 CITATIONS

[SEE PROFILE](#)



**Anna Madison**

University of Illinois, Urbana-Champaign

12 PUBLICATIONS 104 CITATIONS

[SEE PROFILE](#)



**Alejandro Lleras**

University of Illinois, Urbana-Champaign

137 PUBLICATIONS 2,373 CITATIONS

[SEE PROFILE](#)

# Towards a Better Understanding of Parallel Visual Processing in Human Vision: Evidence for Exhaustive Analysis of Visual Information

Simona Buetti, Deborah A. Cronin, Anna M. Madison, Zhiyuan Wang, and Alejandro Lleras  
University of Illinois at Urbana-Champaign

Most current models of visual processing propose that there are 2 main stages of visual processing, the first consisting of a parallel visual analysis of the scene and the second being a precise scrutiny of a few elements in the scene. Here, we present novel evidence that the first stage of processing adds systematic variance to visual processing times. When searching for a specific target, it has a behaviorally unique signature: RTs increase logarithmically with the number of items in the display and this increase is modulated by target-distractor similarity. This signature is characteristic of unlimited capacity parallel and exhaustive processing of all the elements in the scene. The function of this processing is to identify the locations in the scene containing items that are sufficiently similar to the target as to merit focused scrutiny, while discarding those that do not. We also demonstrate that stage-1 variability is sensitive to the observers' top-down goals: with identical displays, whereas RTs increase logarithmically with set size when observers are asked to find a specific target, they decrease exponentially when asked to find a unique item in the scene.

**Keywords:** attention, efficient search, exhaustive processing, parallel processing, similarity

The literature on visual processing and visual attention has, for the most part, converged on the idea that visual processing unfolds over two-stages, mostly independent and sequential in nature (e.g., Hoffman, 1979; Treisman & Gelade, 1980; Wolfe, 1994, but see Fific, Townsend, & Eidels, 2008, and Signal Detection Models of search such as Eckstein, Thomas, Palmer, & Shimozaki, 2000; Vergheze, 2001). First, there is a visual analysis of the scene, during which visual features are encoded across the scene (e.g., Itti

& Koch, 1999; Rosenholtz, Huang, Raj, Balas & Ilie, 2012; Wolfe, 1994). Processing proceeds in a spatially parallel manner: features are all encoded and analyzed simultaneously, at each location, by something akin to low-level feature detectors or feature contrast analyzers. The processing is also often assumed (either explicitly or implicitly) to be of unlimited capacity: the speed of processing at any one location is not affected by how many other features there are in the scene. Similarly, many assume that early on processing at each location is independent. That is, processing of visual information at any one location is not affected by the type of information present at other locations, except when performing computations like local feature contrasts and other immediate local interactions. The second stage of processing follows, a stage often described as being attentive and capacity limited, where one or a few locations are scrutinized in detail (e.g., Baldauf & Deubel, 2009; Bundesen, 1990; Hoffman, 1979; Itti & Koch, 1999; Jeffries, Enns, & Di Lollo, in press; Peters, Iyer, Itti & Koch, 2005; Rosenholtz et al., 2012; Treisman & Gelade, 1980; Wolfe, 1994; Zelinsky, 2008).

There is a parsimonious reason why all these theories have converged on this form of two-stage architecture for vision: vision begins with parallel coding of light information by the retina. Yet, when observers are asked to find an object in a visual search display, it is often the case the more objects there are in the scene, the longer it takes observers to find the target. Thus, it is widely assumed at some point there must be a switch between massively parallel processing of the input (stage 1) and a more capacity-limited, serial processing of individual items in a scene (stage 2).

Significant progress has been made toward understanding each of these two stages, but much of the work on stage one and on stage two has come from different groups of investigators. Those working from a more computationally oriented approach to vision have produced biologically plausible models of stage-one processing where images are analyzed by detecting features and comput-

---

*Editor's Note.* Todd Horowitz served as the action editor for this manuscript.—IG

---

This article was published Online First April 14, 2016.

Simona Buetti, Deborah A. Cronin, Anna M. Madison, Zhiyuan Wang, and Alejandro Lleras, Department of Psychology, University of Illinois at Urbana-Champaign.

The data and Matlab code are available on the Open Science Framework website at <http://osf.io/xcw4f>. Author contributions: Simona Buetti and Alejandro Lleras wrote the paper and equally contributed to the paper. Simona Buetti and Alejandro Lleras developed the theoretical framework; Simona Buetti, Alejandro Lleras and Deborah A. Cronin designed experiments 2, 3A-D; Simona Buetti, Alejandro Lleras and Anna M. Madison designed experiments 1 and 4. Anna M. Madison programed and analyzed experiments 1 and 4. Deborah A. Cronin programed and analyzed experiments 2 and 3. Zhiyuan Wang performed the simulations and eccentricity analyses and wrote the appendices.

We thank Gordon Logan, Cathleen Moore, and Toby Mordkoff for valuable feedback during the development of this project, as well as James Townsend, Ruth Rosenholtz, Todd Horowitz, and two anonymous reviewers for their helpful comments on a previous version of the manuscript.

Correspondence concerning this article should be addressed to Simona Buetti, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820. E-mail: [sbuetti@gmail.com](mailto:sbuetti@gmail.com)

ing feature contrast signals (Itti & Koch, 1999; Itti & Koch, 2000; Rosenholtz et al., 2012). The goal here is to produce computational models that mimic how the human visual system ranks locations in terms of their visual conspicuity (e.g., Itti & Koch, 1999), and how it segments and recognizes objects in complex backgrounds (e.g., Walther & Koch, 2006). Further, understanding the low-level processing constraints in the human visual system also allow for a better understanding of the quality of visual representations (e.g., Pelli, 2008) and, in turn, what type of computations can and cannot be performed on these representations during subsequent processing (e.g., Rosenholtz, Huang & Ehinger, 2012). In contrast, researchers working from a visual cognition-oriented perspective have a different goal: their focus is often to better understand the nature of the limitations of visual attention. Thus, their theories tend to be centered on better understanding the second stage of processing, where visual attention is assumed to play a large role in determining behavior. These researchers focus on questions such as how attention moves in a scene toward a target (e.g., Treisman & Gelade, 1980; Wolfe, 1994) or what determines the speed at which attention processes items in a scene (e.g., Duncan & Humphreys, 1989, 1992). Critically, these models attempt to explain the factors that impact the variability observed in search times by manipulating factors that are presumed to impact visual attention. For instance, Wolfe (1994) proposed a mechanism by which top-down attentional settings could alter the efficiency with which focused attention lands on a target. Duncan and Humphreys (1989 and 1992) demonstrated that increasing target-distractor similarity slows down the rate of attentive processing in a display, as well as modulate the attentional weight of items in the display. Bundesen (TVA, 1990) also proposed an attentive weighting system that determines the likelihood of objects accessing visual short-term memory (STM). Certainly there are several investigators who straddle this divide (e.g., Eckstein et al., 2000; Rosenholtz et al., 2012; Verghese, 2001; Zelinsky, 2008). For instance, recent work by Rosenholtz and colleagues (Rosenholtz et al., 2012 and Zhang, Huang, Yigit-Elliott & Rosenholtz, 2015) has proposed that a critical factor in determining search efficiency is the lossy nature of the representations of peripheral objects: though peripheral objects are not precisely represented, they demonstrated that, with stimuli that produce inefficient searches, there is often enough information at those locations such that peripheral discriminability correlates with search efficiency. The lossy nature of peripheral representations is also related to the phenomenon of crowding and a critical factor in determining eye movements in a scene.

This somewhat simplistic dichotomy is meant to highlight how different goals have shaped the research on this topic. A consequence of this dichotomy is that investigators interested in visual attention and behavior, that is, those investigators that have focused largely on characterizing stage-two processing, have largely ignored the possibility that behavior may also be profoundly affected by variability during stage-one processing. Researchers focused on characterizing stage one have similarly avoided this possibility because their focus is less on performance than it is on computation and the outputs it produces. Thus, an important question has gone unaddressed by the literature so far: is there systematic variability in human vision during stage-one processing as there is in stage two? If so, how can it be characterized? Is it related to top-down goals? And how does it affect subsequent

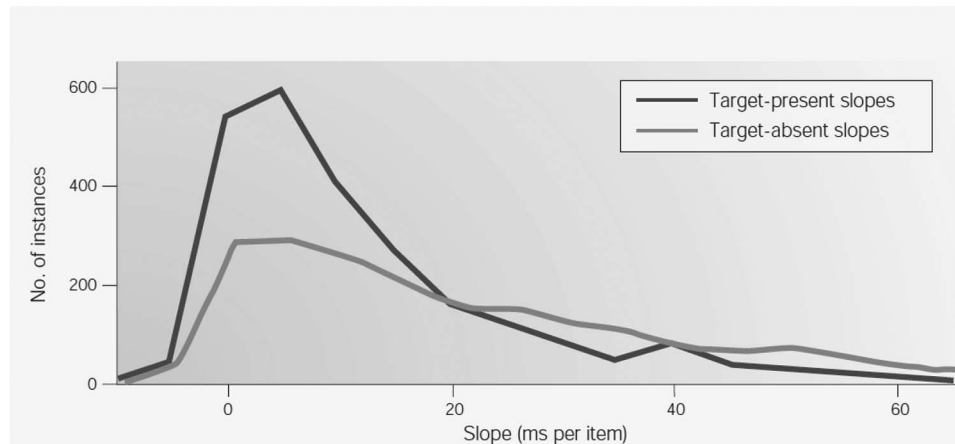
stage-two processing? The goal of the current manuscript is a first attempt at filling this gap.

### Current Assumptions About Stage-One Processing

The current assumption in many models is that the form of parallel, unlimited capacity processing characteristic of stage one should produce constant stage-one completion times. This assumption is explicit in theories like Guided Search (Wolfe, 1994), where 400 ms is the time allowed to process all displays irrespective of their visual complexity and to select and execute a response. In Bundesen's TVA (1990), there is a mathematically explicit goal to make search times independent of set size during searches thought to rely almost entirely on stage-one processing, so-called pop-out or feature searches. This assumption that stage-one processing produces no systematic variability is also implied in other computational models that attempt to account for human variability in search times (Itti & Koch, 2000). We propose that this state of affairs is likely the result of a preconception regarding parallel processing. As we will show below, and as was first demonstrated by Townsend and Ashby (1983) more than 30 years ago, the idea that parallel unlimited capacity processing produces RTs unaffected by set size does not always hold. In fact, we will argue that parallel, unlimited capacity processing in stage one systematically contributes to variance in RTs, even in the simplest of visual search tasks.

It must be said that sometimes variability has been observed and ascribed to stage-one processing (e.g., Wolfe, Oliva, Horowitz, Butcher & Bompas, 2002). However, it has also often been assumed that the underlying factor, like most factors impacting stage-one completion times, is orthogonal (or uninteresting) to the study of attention. Thus, for experimental design reasons, factors like eccentricity, display density, and crowding are often controlled for and not systematically investigated (but, see Balas, Nakano, & Rosenholtz, 2009; Rosenholtz et al., 2012).

In addition, when some presumed stage-one variability is found, as when search slopes are close to but significantly different from zero, the variability has been assumed to be theoretically unimportant (e.g., Wolfe & Horowitz, 2004). This follows because of a theoretical emphasis on differentiating two different forms of attentional behavior: when attention moves directly to the target in a scene (so called, "efficient" search or preattentive search) and when attention has to move to or inspect several locations before landing on and identifying the target (so called "inefficient" search or attentive search, see Treisman & Gelade, 1980; Wolfe, 1994). Thus, the emphasis is placed on understanding what determines whether the variability in a given search will be near zero or much larger. On closer inspection, this division of efficient and inefficient slope measurements is not so clear: when search slopes were measured across a wide ranging set of stimulus parameters, there was little evidence for two different types of slopes (Wolfe, 1998). That is, one would have expected a distribution of search slopes around zero (indexing preattentive searches) and a second distribution centered around a nonzero value (indexing the involvement of attention), but this was not observed (see Figure 1, but see Haslam, Porter & Rothschild, 2001, for a reanalysis of the same data that finds evidence for two distinct distributions of search slopes). Haslam et al.'s (2001) analysis notwithstanding, it was proposed that a continuum of search efficiency was a better way to



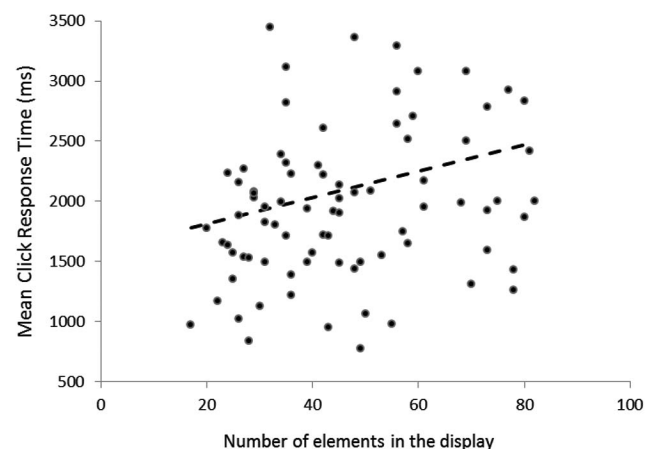
*Figure 1.* Distribution of search slopes (ms per item) as observed in a wide range of search tasks. Reprinted with permission from “What Attributes Guide the Deployment of Visual Attention and How Do They Do It?” by J. M. Wolfe and T. S. Horowitz, 2004, *Nature Reviews Neuroscience*, 5, p.4. Copyright 2004 by Nature Publishing Group. From this distribution, one can see that there are as many (if not more) so-called “efficient” searches as inefficient ones in this database. The distribution is also not clearly bimodal, rejecting a strong preattentive versus attentive search dichotomy. Reanalysis for these data by Haslam, Porter, and Rothschild (2001) demonstrated that these histograms were more likely the result of two distinct underlying distributions (with different frequencies of occurrence and some overlap) than the result of a uni-modal (skewed) distribution.

understand attentive processing than in terms of a dichotomy, and it is in practice what researchers have used since. People referred to relatively “efficient” search when slopes were generally in the vicinity of 0 ms/item (with generally an upper bound around 10 ms/item) and to “inefficient” search, when slopes were larger than 10 ms/item. In practical terms, efficient slopes indicated that attention was not critical to performance in the task. Consequently, there was no need to directly compare two search slopes against each other when they fall in the efficient search range, because they were both basically reflecting the same nonattentive processing. To study attention, the emphasis today remains on studying search functions that fall in the inefficient search range.

A significant problem with the current approach to visual processing is that existing theories have a very difficult time accounting for performance variability when search is studied in real scenes (e.g., Itti & Koch, 2000; Wolfe, Alvarez, Rosenholtz, Kuzmova & Sherman, 2011). In the case of Wolfe et al.’s study (see Figure 2), there is a great deal of variability in the data that cannot be accounted for by a linear slope indexing attentional scrutiny (that is, stage-two processing) of a real scene. In fact, a linear slope only accounts for about 4% of the variance in those experiments. This suggests that there may be important variability in visual search behavior yet to be accounted for. Some have proposed this source of variance is knowledge: the fact we know where things tend to be in real world scenes may allow us to search more efficiently in real scenes than in abstract lab displays (Wolfe, Alvarez, et al., 2011; Wolfe, Vö, Evans, & Greene, 2011). Another possibility is that other purely visual sources of variance may be important contributors to performance. Clearly, work by Rosenholtz and colleagues (e.g., Zhang et al., 2015) and Duncan and Humphreys (1989) already demonstrated that this is the case for stage-two processing. But, the issue of whether this is also the case for stage-one processing remains open. As we will demonstrate

here, the similarity of nontarget elements to the target can influence response times, even when those items are not scrutinized during stage-two attentive processing. In the General Discussion, we provide a proof of concept that perhaps stage-one variability combined with stage-two variability can help explain this disconnect between visual search tasks with simple abstract stimuli and visual search in real-world scenes.

Finally, it is also important to remember that the current conceptualization of attention as a limited capacity, stage-two mechanism has led researchers to propose that the rich visual sensation that we get when we open our eyes, our seemingly full awareness



*Figure 2.* Replotted data from Wolfe, Alvarez, et al. (2011) Experiment 2 showing time to click with a mouse on a target as a function of the number of elements in a real world scene. The slope is approximately 11 ms/item and the  $R^2 = 0.044$ .

of the world, is a Grand Illusion (Blackmore, Brelstaff, Nelson & Troscianko, 1995; Dennett, 1991, 1992, 1998; Mack & Rock, 1998; O'Regan, 1992; Rensink, 2000; Rensink, O'Regan & Clark, 1997; Treisman & Gelade, 1980). The claim is that we are only aware of the information we attend to, and that attention precedes awareness (though this claim has been challenged elsewhere: Lamme, 1995; Lamme, Supér, & Spekreijse, 1998; Landman, Spekreijse, & Lamme, 2003, 2004; Sligte, Scholte, & Lamme, 2008, 2009; Sligte, Vandenbroucke, Scholte & Lamme, 2010; Supér, Spekreijse, & Lamme, 2001, 2003; Vandenbroucke, Sligte, Fahrenfort, Ambroziak, & Lamme, 2012). The rapid movement of attention around the scene before our awareness produces the illusion of detailed visual awareness. Many propose that without attention, visual processing fails quite dramatically (e.g., Ward & Scholl, 2015). For instance, some propose that without attention, the collection of features that make up objects in scenes are not bound into coherent representations (Rensink, 2000; Treisman & Gelade, 1980; Treisman & Schmidt, 1982). In the General Discussion, we will argue that our results are consistent with a form of parallel, unlimited-capacity attentive processing during stage one (Experiments 1–3) that can theoretically produce rich percepts of our entire visual field. Additionally, as suggested by Experiment 4, this early attentive processing is capable of performing different computations, and thus could play a role in a number of tasks and domains other than visual search (such as scene processing, gist extraction and the extraction of statistical properties in a scene, see Chong & Treisman, 2005a, 2005b; Haberman & Whitney, 2009; Oliva, 2005; Potter, 1975, 1976, 2012; Torralba & Oliva, 2003; Oliva & Torralba, 2006). At any rate, theoretical proposals that rely critically on attention being only a stage-two, capacity-limited mechanism, including theories of the Grand Illusion, may require revision.

### Some Important Definitions

For clarity's sake it is worth being explicit in the definitions of the terms we will be using throughout the present paper as readers may associate slightly different concepts with those terms. Thus, several definitions follow. They were largely adapted from Townsend and Ashby (1983) to the specific task of visual search.

#### Parallel Processing

In the case of visual processing, by parallel processing we mean that processing will unfold *concurrently* at each location in the display where an item is present. The processing begins simultaneously at all locations when the display appears. Note that this is an architectural definition of processing (i.e., how things are wired), not a functional definition of processing (i.e., what is being computed). Importantly, that elements in a scene are processed in parallel does *not* imply that all elements will be processed at the same speed, nor that processing will complete for each element at the same time or that all elements will be fully processed. Such processing assumptions are entirely independent from the claim that processing is progressing in parallel.

#### Unlimited Capacity

By unlimited processing capacity we mean that the rate of processing of individual items (or "channels" of processing) is not

restricted by the number of other concurrent channels that are simultaneously at work. The rate of processing might be affected by other factors (like the signal-to-noise ratio of visual information in the display), but *not* by the number of channels doing processing. Townsend and Ashby (1983) for instance proposed that, given the stochastic nature of neural processing, the finishing time for a given channel can be understood as being sampled from an exponential distribution. The finishing time reflects something like the inverse of the resource or efficiency with which a particular channel is processed. Changing factors like eccentricity or signal-to-noise ratio for each item might affect the shape of the exponential distribution from which the finishing time is sampled. The concept of capacity, however, is only related to whether or not the distribution (for each individual item) is a function of number of items in the display. Let's consider for example the case of a display containing two items. When processing is unlimited capacity, the finishing time of each of the two items is initially determined by sampling from the same distribution twice, once for each item. The sampling itself is identical as what it would be when only one item appears on the display. As a result, the actual finishing times between the two items will differ, but that difference is purely attributable to random sampling of the same underlying distribution. In contrast, if processing were *capacity limited*, the exponential distribution for each of the two items would be different from when that item appears alone. More specifically, when capacity is limited, capacity becomes shared among channels. Because of resource sharing, capacity limited processing implies the initial rate of processing for each item would be determined by obtaining the 'standalone' rate and multiplying it by a factor less than one. The specific sharing rules would determine what factor to use for each item in the display. As a result, the completion times of individual items/channels in the display would differ directly as a result of the number of items among which the resources must be shared.

#### Independent Processing

In the case of visual processing, by independent processing we mean that processing at each location unfolds independently of what other visual information may or may not be present in the display and independently of what other processing may or may not be going on. For example, in a scene containing only two elements being processed independently of one another, once started, the processing of each item does not depend on (a) the presence of the other element nor (b) whether the other element is being processed or even (c) whether the other element has already been fully processed.

#### Reallocatable Capacity

One example of nonindependent processing is when processing capacity is allowed to be redistributed upon channel completion. Imagine a case where three items are present in the display. Reallocatable capacity would mean that as soon as one channel completes its processing, the capacity freed by this termination can be reallocated to the two unfinished channels, thereby speeding processing in the two still active channels. And, when the second channel completes its processing, its processing capacity would in turn be reallocated to the lone active channel, further speeding processing at that channel.



Notice that *capacity* and *independence* are two orthogonal aspects of visual processing: a system can be capacity-limited, but processes items independently, or it can be unlimited capacity with reallocatable capacity. In fact all four of the concepts (serial vs. parallel, capacity, independence, stopping rule) defined here are orthogonal to each other and thus can be freely combined to describe a specific processing system.

## Stopping Rule

The stopping rule refers to the rule that determines when the processing at any particular stage will end or be aborted. Having an *exhaustive* stopping rule means the processing for a particular stage will only end once the processing of *all* channels has completed. In contrast, in a search task with a single target, having a *self-terminating* stopping rule for a particular stage means that processing for that stage will end *as soon as the channel containing the target* completes processing (i.e., as soon as the target is fully processed). For instance, many people associate the self-terminating stopping rule with the stage-two attentional scrutiny of hard to find targets: in a “Where’s Waldo?” type of search, stage-two processing of the display will end as soon as one is lucky enough to stumble on Waldo (even if it means not having scrutinized all elements in the image). Similarly, a famous example of an exhaustive stopping rule being at play was demonstrated by Sternberg (1966, 1969), who found that search through a memory set is in fact exhaustive.

Finally, we will be using two terms to define distractor stimuli in our experiments. *Candidates* refers to items that are sufficiently similar to the target as to require close scrutiny by focused attention to determine whether or not the item is the target itself or a distractor. The term was first proposed by Neider and Zelinsky (2008, see also Lleras, Buetti & Mordkoff, 2013) to refer to the subset of search items that compose the “functional set size” of a display. This proposal was based on the observation that in natural scenes the actual number of objects in the world does not equal the total set size for a search task, given that attention will likely skip many (the most dissimilar) items (see also Wolfe, Alvarez, et al., 2011). Thus, we propose to use the term “*lures*” to refer to those items that are in the scene but are sufficiently different from the target as to not require the sort of close attentional scrutiny that candidates require. Lures are not part of the “functional set size.” Adding the total number of lures plus the total number of candidates in a scene should equal the total number of objects in the scene. In terms of the two-stage models of attention, this distinction between lures and candidates amounts to say that lures are those stimuli processed by stage one, but not stage two, whereas candidates are fully processed by both stages one and two.

Interestingly, researchers in visual cognition do not agree on the nature of the resources (unlimited vs. limited capacity) nor on the stopping rule (exhaustive vs. self-terminating) in the most basic of visual tasks: an efficient, feature-search task. We asked 56 anonymous attendants at the 2015 Meeting of the Vision Sciences Society (graduate students, postdocs and investigators interested in visual attention and visual search) to report whether they thought this type of search task was performed with limited or unlimited capacity: 38% reported unlimited capacity, 52% reported limited capacity, and 10% reported “don’t know/prefer not to answer.” Thus, it appears that there is no clear agreement among this expert

group on the type of capacity limitations (or lack thereof) that govern this very simple search task, though there was a tendency to report processing is capacity limited. We also asked them whether this type of search was self-terminating or exhaustive in nature: 52% reported that feature search is self-terminating, 25% thought it would be exhaustive and 23% reported “don’t know/prefer not to answer.” Thus, among those familiar with these concepts, by a 2:1 margin, most investigators believe efficient search tasks terminate as soon as the target in the display is processed/found. In sum, there is no clear agreement among experts regarding these basic concepts and we hope the current paper will help shed some light on these issues.

## Rationale of the Current Study

This paper aims at providing a first characterization of stage-one processing. First, we studied several examples of “efficient” search tasks (Experiments 1A and 1B) to illustrate stage-one variability. Our first attempt was based on the idea that even in efficient search, target/nontarget similarity *ought to* influence stage-one processing times. In fact, findings from monkey neurophysiological studies have long suggested as much. In a recent paper, Purcell and colleagues (2010) showed that single-cell activity in the FEF was determined by the similarity of the distractor to the target in its receptive field. In their study, more similar distractors drove higher responses in the cell than more dissimilar distractors, even though the initial spiking response (the accumulation of evidence) occurred at the same pace for both types of distractors. In other words, perceptual evidence regarding the identity of an object in the receptive field of the neuron is initially accumulated at the same rate, irrespective of the object’s similarity to the target. However, more similar distractors were processed for longer periods and produced heightened sustained responses compared to less similar distractors. This is true despite the fact that for both similar and dissimilar distractors, the target was sufficiently different from the distractors so that the search could be characterized as a pop-out. This was also evidenced by very fast and accurate saccades directed to the target. Similarly, Wan and Lleras (2010) showed that target-distractor similarity could produce massive differences in overall RTs in a pop-out search task, whereas the search slopes continued to be clearly efficient.

Second, in Experiments 3A through 3D, we tried to experimentally dissociate stage-one variability from stage-two variability by using displays that contain two types of elements: candidates, which should require close scrutiny to be differentiated from the target, and lures, which being so dissimilar from the target, should not. These displays were inspired by the concept of “functional set size” (e.g., Neider & Zelinsky, 2008) that has been applied to real world scenes. The idea is that, in contrast to traditional lab displays, in the real world, observers only scrutinize the subset of the elements in the scene that are most similar to the target. The question is simple: if we keep the same “functional set” of distractors constant across experiments and only manipulate those elements outside of this set (the highly dissimilar lures), will the number and type of lures present in the display impact response times?

Third, we tested whether stage-one variability is influenced by top-down factors, such as the experimental task (Experiment 4). Thus, using the exact same displays as in Experiment 1A (as well

as the same type of response), we changed only the nature of the search task, going from a fixed-target search (Experiment 1A) to a feature-oddball task (Experiment 4). The goal was to test whether such high-level constraints could fundamentally alter the manner in which stage one processes visual input.

Finally, the paper also includes one control experiment (Experiment 2) that was designed to demonstrate that the set of stimuli we designated as “candidates” do produce linear reaction time (RT)  $\times$  set size functions.

### Experiments 1A and 1B: “Efficient” Search Is Logarithmic and Is Sensitive to Lure-Target Similarity

Typically, researchers seeking to elicit efficient search use stimuli with starkly different features, surely inspired by the early work of Treisman and Gelade (1980). Their goal was to have one stimulus (the target) represented alone in one feature map by making it the only red or only vertical or only square item in a display. This idea of using a unique feature to distinguish the target from the distractors overshadowed the possibility of investigating how efficient search might be affected by more fine-grained differences in low-level features. For example, might efficient search for a red target be easier in a display filled with blue distractors than in a display filled with green distractors? In Treisman and Gelade’s framework (and much of the work that followed), this is an ill-posed question because, as long as each color feature had its own feature map, there ought not be any differences. In contrast, target/nontarget similarity is known to give rise to differences in stage-two processing (e.g., Duncan & Humphreys, 1989; Zhang et al., 2015). In fact, Attentive Engagement Theory was proposed based on experiments that only used Ls and Ts as stimuli (Duncan & Humphreys, 1989). This experimental design was elegant and much needed at the time, but it resulted in a skewed distribution of investigations of similarity space: on one end, items very-to-extremely similar to one another (rotated Ts and Ls) and on the other extreme, items very-to-extremely different from each other (red vs. green). Thus, we have moved forward as a field without having empirically populated the space between these two extremes.

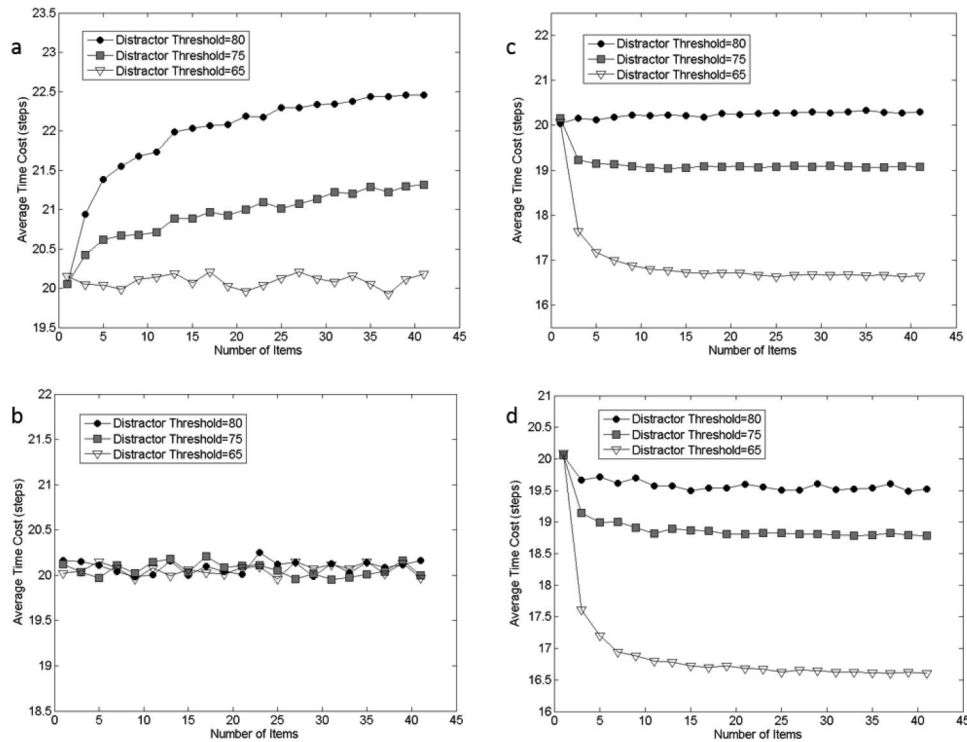
Here, we propose that a better understanding of visual search (and therefore attentive vision) requires a bigger experimental sampling of similarity space. The goal of Experiment 1A is to examine how presumably efficient visual search differs when performance on trials containing diverse a priori levels of target/nontarget similarity is compared. In this experiment, the target was a red triangle and participants were asked to report whether the target was pointing to the right or left. The target was presented among a field of identical lures. Set size and lure type were varied across trials. There were three types of lures: blue disks, yellow triangles, and orange diamonds. The orange diamonds had a red-dish hue and were composed by abutting the right and left pointing triangles that were used as the target in 1A. These stimuli had the highest level of lure-target similarity, as they shared both shape and color with the target. Yellow triangle lures had an intermediate level of similarity, as they shared shape but not color with the target. Blue disks were the most dissimilar to the target as they shared neither shape nor color with the target.

The predictions for this experiment according to most models of visual attention are straightforward: there should not be many (if

any) differences in performance across the three different lure types and RTs should be fairly unaffected by set-size. This follows because if the target is always the first item to be inspected by attention (i.e., it pops out), RTs should always be fairly constant. Perhaps the unspoken assumption in many current attention models is stage one would work as a parallel, unlimited capacity but *self-terminating* process, such that once the pop-out item in a scene is found (something made possible by the low-level contrast computations), nothing else would be attended in the scene. And further, if target/nontarget similarity were to affect performance, these models predict the modulation would occur in the “inefficient” range of search slopes (i.e., slopes larger than 10 ms/item or so), with the linear slope increasing with increasing target/nontarget similarity (e.g., Duncan & Humphreys, 1989; Zhang et al., 2015).

In contrast, there is reason to believe that stage-one parallel processing should impact performance, even when processing occurs with unlimited capacity. What patterns of results may be expected from this type of processing? In 1983, Townsend and Ashby systematically characterized the shape of the RT by set-size functions that one would expect for different forms of parallel, unlimited capacity processing (see Figure 3). There are at least four different types of unlimited capacity parallel functions. The first has independent capacity and is exhaustive in processing, where each item is processed with a separate pool of resources and processing continues until all items have been processed (this is what we propose stage-one processing is like). The second has independent capacity but is a self-terminating process, where once the target is found, processing of other items stops. As mentioned above, this is the type of processing that is often assumed to be at play in visual search because it produces RT functions that are unaffected by set size. The third type has reallocatable capacity, with exhaustive processing, where processing ends when all items are processed, but once the processing of any item finishes, the resources freed up by the completion of that item are reallocated to the processing of unfinished items. The fourth type has reallocatable capacity and self-terminating processing, where processing ends as soon as the target is found and any capacity freed up by completing nontarget items is reallocated to unfinished items. Townsend and Ashby presented the underlying mathematical models describing each of these forms of parallel processing. In Appendix A, we present their mathematical formulas and how we have applied them to predicting RT in visual search task.

The critical insight here is that we propose *the required amount of processing for any given item is determined by that item’s similarity to the target template*. In other words, there is a top-down contribution to stage one, such that by defining a target template, processing during stage one evolves *in relation to* that template. We can simply imagine that at each location, information is accumulating toward a decision: is this item likely to be the target or not? More target-similar items will require more evidence accumulation before a decision is reached compared to more target-dissimilar items. Note that an alternative could be that evidence accumulates at different rates, as a function of lure-target similarity. But, we propose instead that, all else being equal, perceptual accumulation occurs at the same pace and that similarity, or rather *dissimilarity* determines the decision threshold for each item. For example, when looking for an elephant in a zoo, very little evidence is required to reach the decision that an



**Figure 3.** Simulated search time as a function of number of items to be processed for four models of unlimited capacity, independent parallel processing. In all four models, the processing of each item involved an independent, noisy accumulation of evidence, which completes when accumulated evidence reaches its threshold. There was always one target item with a fixed threshold of 80 (arbitrary units), and only one type of distractor item, with a lower threshold. The four models differed in when processing of the entire display terminates (exhaustive or self-terminating) and in whether the resources of completed channels get reallocated to working channels. In this simulation, evidence gathered at each step for each channel is randomly drawn from a Gaussian distribution of  $\mu = 4$ ,  $\sigma = 3$ . All four variations were simulated with distractor thresholds at 65, 75 and 80. (a) *Independent and exhaustive processing*: processing was terminated when all channels complete and no reallocation occurs. For all distractor threshold levels a distinct logarithmic pattern can be observed. Curve fitting to logarithmic functions for the upper two cases (distractor thresholds at 80 and 75) had R squares greater than 0.96. (b) *Independent, self-terminating processing*: no reallocation and processing terminated as soon as the target channel completed. Naturally, this model rendered distractor channels irrelevant, so that processing time cost was unaffected by either number or threshold of distractors. (c) *Reallocated and exhaustive processing*: reallocation occurred such that working channels speed up as other channels complete. For example if half of the channels had completed, the evidence gathered at each step would be drawn from a Gaussian with  $\mu = 8$ ,  $\sigma = 6$ . (d) *Reallocated, self-terminating processing*. Reallocation models (c and d) predicted that processing time decreases with addition of more distractor items (except for the exhaustive model when distractors have the same threshold as the target). These simulation results are mostly consistent with what Townsend and Ashby (1983) had originally shown by analytical solution (under a different assumption about individual item processing), with the addition of the effect of difference in distractor thresholds. See Appendix A for more detail.

ice-cream booth (a very dissimilar lure) is not an elephant. However, more evidence accumulation might be required to decide a hippopotamus (a less dissimilar lure) is not an elephant. This is consistent with single-cell recordings in FEF from Purcell et al. (2010) showing that a cell's response is determined by the similarity of the distractor to the target in its receptive field. In their study, more similar distractors drove higher responses in the cell than more dissimilar distractors, even though the initial spiking response (the accumulation of evidence) occurs at the same pace for both types of distractors. This is true despite the fact that in both cases the target is sufficiently different from the distractors to be characterized as a pop-out.

In Townsend and Ashby's terms, Purcell et al.'s findings mean that expected completion times for a similar item are going to be longer than for a dissimilar item. Because stage one is unlimited in capacity, information accumulates at all locations at the same speed, so the rate of evidence accumulation itself is not a limiting factor in performance. Attention's top-down effect determines the goal of the accumulation (based on the target template). This proposal is also consistent with single-cell recordings from monkeys showing that attention's top-down expectations change the receptive field properties of neurons as early as V1 ahead of the presentation of the stimuli (Ito & Gilbert, 1999; Li, Piëch, & Gilbert, 2004, 2006; Roelfsema, Lamme, & Spekreijse, 1998).



To summarize, the prediction of a parallel, unlimited capacity stage that processes displays elements in independent and exhaustive fashion is that RTs should increase in logarithmic fashion as a function of the number of lures in the display (Figure 3A). Further, as demonstrated in Appendix A, if lure-target similarity affects the decision thresholds, then one should expect that more similar lures will produce *steeper* logarithmic curves than less similar lures. In contrast, if processing is self-terminating, then RTs will not be affected by set size nor by lure type (see Figure 3B). This follows because processing always ends when the target is finished processing, thus neither the number of other elements in the display, nor their similarity to the target template matter.

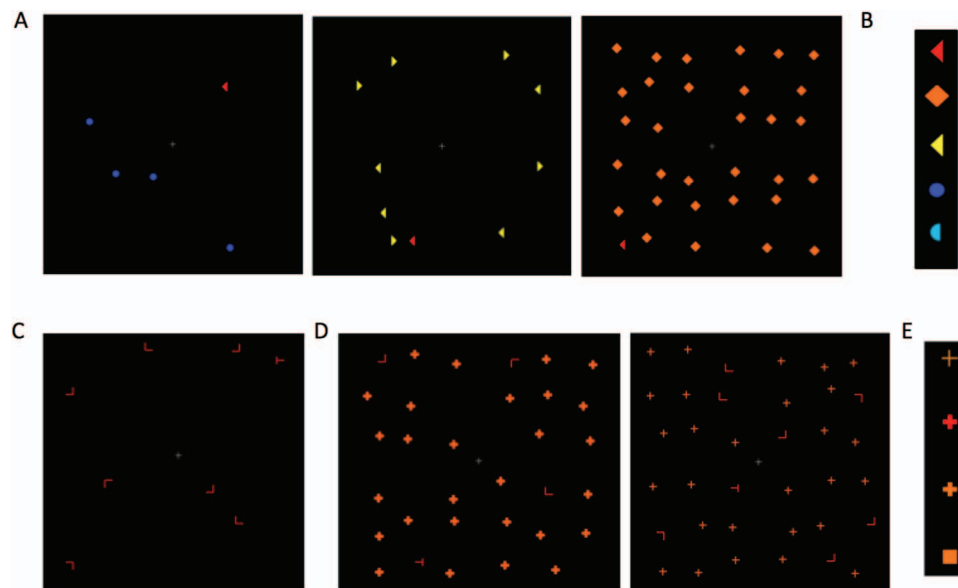
### Method: Experiment 1A

**Subjects.** Participants completed the experiments in exchange for course credit in a Psychology class. Target sample size was determined based on a power analysis performed on a pilot study using the same pairing of target and distractors used in Experiment 1. The power analysis showed that to find an effect of our most dissimilar distractors on performance (the blue circles) with 80% power, a minimum sample size of 18 subjects was needed. This was based on a manipulation that included a maximum set size of 26 blue items. The observed effect size was  $f = 0.44$ . We decided to increase the sample size to 20 participants in anticipation that (a) the effect of the most dissimilar items (blue) may be noisier than

our original pilot data suggested and (b) because we were interested in being able to detect differences between lure levels. The first 20 subjects meeting the inclusion criteria were used in each experiment's analyses for both Experiments 1A and B: we replaced participants who performed below 90% accuracy and who had mean response times higher than 2 times the standard deviation of the mean response time for the group. This resulted in one subject being replaced in Experiment 1A. One subject was replaced in Experiment 1B because of an experimenter error. Participants in Experiment 1B were tested for color blindness using the Ishihara color plates.

**Apparatus, stimuli and procedure.** In Experiment 1A the target was a red triangle and in Experiment 1B the target was a blue half disk. Participants were asked to find the target and indicate whether the target was pointing to the left or right. Eye movements were not monitored and participants were free to move their eyes as they wished. There were three different types of lures: blue circles, yellow triangles pointing right or left and orange diamonds. The hue of the orange diamonds was similar to that of the target in Experiment 1A (Figure 4A) and the hue of the blue disks was similar to that of the target in Experiment 1B. The number of possible lures was 1, 4, 9, 19, and 31, and a target-alone condition was included. A target was present on every trial.

There were 50 observations per condition. All items subtended .833 degrees of visual angle and were presented (with



**Figure 4.** (A) Sample displays from Experiment 1A showing the three different lures used in the Experiment and three different levels of set size. The target was the red triangle and participants indicated whether it pointed right or left. (B) Target and lure stimuli from Experiments 1A and 1B. The top stimulus (red triangle) was the target in Experiment 1A and the bottom stimuli (blue half disk) was the target in 1B. The lure stimuli are presented in between and are organized in descending order of similarity to the red triangle: orange diamond, yellow triangle and blue disk. They are also ordered in increasing similarity to the blue half disk. (C) Sample display used in Experiment 2. Participants indicated the presence or absence of the oriented red T. (D) Sample displays used in 3A and 3B, respectively from left to right, and showing two different levels of candidates (4 on the left, 8 on the right). Participants reported the direction of the oriented red T. (E) Lures used in Experiments 3A through 3D, arranged from most similar (top) to least similar (bottom) to the target template. See the online article for the color version of this figure.

random jitter) inside an invisible square grid with 36 cells ( $6 \times 6$ ), that occupied the entire 20 in. display (20 degrees of visual angle). When two stimuli appeared in adjacent cells, the average spacing between elements was 1.5 degrees of visual angle. Search displays were randomly generated for each trial. Displays remained on the screen until response or until 5 seconds had passed. All experiments were run on a PC, using Matlab, and programmed in the Psychophysics Toolbox, version 3 (Brainard, 1997; Pelli, 1997).

### Results of Experiment 1A: Red Triangle Targets

The results from Experiment 1A (Figure 5A) indicate that efficient search for a fixed target can be characterized as having a logarithmic relationship between set size and RT. Indeed, a logarithmic fit produced very high R-squares for all conditions. For orange diamonds, yellow triangles, and blue circles the R-squares were  $R^2 = 0.9938$ ,  $R^2 = 0.9625$ , and  $R^2 = 0.9386$ , respectively. The logarithmic fit was in all cases higher than the linear fit,  $R^2 = 0.8575$ ,  $R^2 = 0.6947$ , and  $R^2 = 0.7789$ , respectively. We used the natural logarithm when fitting the data (see Figure 5) and in our modeling equations (see Appendix A), but it is worth noting that all log functions are proportional to each other, so the curves in Figure 5B would have been linear in log space, irrespective of the log base used (log 2 or log 10).

Furthermore, lure-target similarity modulated the steepness of the logarithmic function, with steeper functions for more similar lures. We computed a within-subject ANOVA with lure type and set size as factors on correct RTs (excluding the target only condition). RTs increased monotonically with set size,  $F(4, 76) = 56.13$ ,  $p < .001$ ,  $f = 1.718$  (see Figure 5A), and with lure-target similarity,  $F(2, 38) = 18.94$ ,  $p < .001$ ,  $f = 0.998$ . More critically, the interaction between lure type and set size was significant,  $F(8, 152) = 45.87$ ,  $p < .001$ ,  $f = 1.553$ , indicating that different levels of lure-target similarity produced different levels of processing efficiency. Paired  $t$  tests indicated that all logarithmic search slopes differed from each other: the

log slope for orange diamonds (41 ms per log unit) was larger than the log slope for yellow triangles (24 ms per log unit,  $t(19) = 4.74$ ,  $p < .001$ ), which itself was larger than the log slope for blue circles (11 ms per log unit,  $t(19) = 2.98$ ,  $p = .0076$ ).

In these three lure conditions, the observed “linear” slopes were very small and fall squarely in the *efficient* search end of the spectrum, as traditionally described (Figure 5B). For instance, the linear slope over the entire range of set sizes (i.e., including the target only condition as a datum in all three search functions) was 4.2 ms per item for orange diamonds, 2.6 ms per item for yellow triangles, and 1.1 ms per item for blue circles. An argument could be made that the functions should not include the target-only condition, as it creates a large discontinuity in most of the supposedly linear search functions. The same analysis excluding the target-only condition produced smaller linear search slopes: 3.6 ms per item for orange diamonds, 1.9 ms per item for yellow triangles, and 0.9 ms per item for blue circles. The linear fit improved somewhat, with  $R^2 = 0.9168$ ,  $R^2 = 0.7714$ , and  $R^2 = 0.7606$ , respectively, but remained lower than the log fits including the target-only condition. In sum, the stimuli we chose in all three conditions gave rise to what are traditionally considered to be “efficient” or “pop-out” search functions.

Finally, it is worth pointing out that the target-only condition (set size one) appears to be a natural continuation for two of the three search functions when set size is plotted in a logarithmic scale. The observed lure logarithmic intercepts were computed (without the target only condition) for each lure type and a Bayesian Information Criterion (BIC) was used to test the hypothesis that the observed logarithmic functions for orange and blue lures actually do not differ in the point of convergence ( $H_0$ ) against the hypothesis that the functions do differ ( $H_a$ ; Masson, 2011). The BIC results suggest positive evidence for the hypothesis that these two log intercepts do not differ from the target only Condition  $P(H_0|D) = 0.811$  and  $P(H_a|D) = 0.189$ , and  $P(H_0|D) = 0.811$  and  $P(H_a|D) = 0.189$  for the

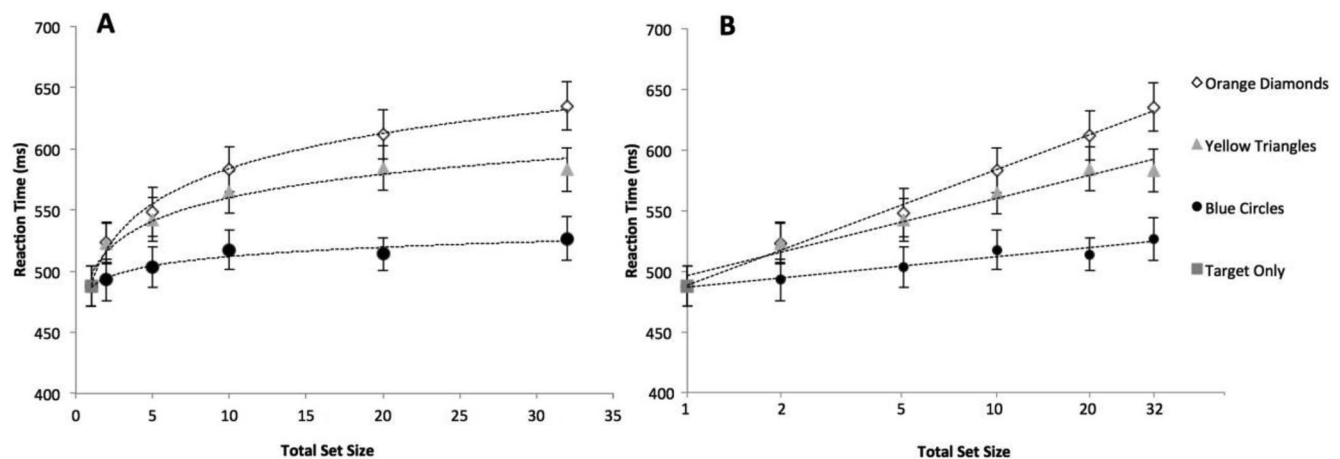


Figure 5. (A) Response times in Experiment 1A for correct trials only, plotting set size on a linear scale, showing a logarithmic relationship between number of items and RT (in ms). The  $R^2$  for the logarithmic fit for orange diamonds, yellow triangles, and blue squares were  $R^2 = 0.9938$ ,  $R^2 = 0.9625$ , and  $R^2 = 0.9386$ , respectively. (B) Same data, plotting set size on a logarithmic scale. Error bars indicate the standard error of the mean.

orange and blue lures, respectively. In contrast, an ANOVA for yellow lures did show a significant difference,  $F(1, 19) = 5.19$ ,  $p = .034$ ,  $f = 0.214$ . In sum, there is evidence in two of the three search functions that the logarithmic functions converge on the “target-only” datum.

### Method: Experiment 1B

Experiment 1B was run as a follow-up experiment using the same set of lure stimuli but changing the target stimulus. The goal was to demonstrate that it is the lure-target similarity relationship that drives the steepness of the logarithmic slopes observed in Experiment 1A, not the lure stimuli themselves. We selected a target that was closer in similarity to the blue disks, which were the most target-dissimilar lures in 1A. Thus, the target in 1B was a blue half disk, slightly lighter than the blue lures. This half-disk target was obtained by cutting in half a blue disk lure along its vertical axis, to obtain two possible target disks, one with its curved edge to the left, one to the right, so that we could use the same type of response discrimination task in 1B as we did in 1A: reporting which direction the target “pointed.” As a result, the blue disks had the highest a priori level of lure-target similarity (shared color and shape with the target). Yellow triangles had, at most, mild similarity to the half-disk targets since they were both left/right oriented shapes. Finally, orange diamonds were the most dissimilar to the target since neither shape nor color were shared with the target. In sum, the identity of the target stimulus is the only thing that changed between Experiments 1A and 1B. Everything else was identical across experiments. We predicted the blue disks should now produce the steepest logarithmic slope, whereas orange diamonds and yellow triangles should produce much shallower slopes.

### Results of Experiment 1B: Blue Half-Disk Targets

The results from Experiment 1B (Figure 6A) confirm the findings from Experiment 1A: search times increase logarithmically with set size and the slope of the log function is sensitive to lure-target similarity. A logarithmic fit produced higher R-squares for all conditions than the corresponding linear fit. For blue disks, yellow triangles, and orange diamonds the R-squares the logarithmic fits were  $R^2 = 0.9681$ ,  $R^2 = 0.7504$ , and  $R^2 = 0.9456$ , respectively, and the linear fits were  $R^2 = 0.8819$ ,  $R^2 = 0.4138$ , and  $R^2 = 0.6488$ , respectively. As with Experiment 1A, the lure-target similarity modulated the steepness of the logarithmic function, with, as expected, the blue lures producing the steepest logarithmic function. A within-subject ANOVA was computed with lure type and set size as factors on correct RTs (excluding the target only condition). RTs increased monotonically with set size,  $F(4, 76) = 54.46$ ,  $p < .001$ ,  $f = 1.691$  (see Figure 6A), and with lure-target similarity,  $F(2, 38) = 57.88$ ,  $p < .001$ , Cohen's  $f = 1.746$ . More critically, the interaction between lure type and set size was significant,  $F(8, 152) = 19.40$ ,  $p < .001$ ,  $f = 1.010$ , indicating different levels of lure-target similarity produced different levels of lure processing efficiency. Paired  $t$  tests indicated the logarithmic search slope produced from the blue circle lure items (65.4 ms per log unit) differed from the two other lure log slopes. The blue circles log slope was larger than the log slope for yellow triangles (20.1 ms per log unit,  $t(19) = 7.40$ ,  $p < .001$ ) and orange

diamonds (16.7 ms per log unit,  $t(19) = 7.19$ ,  $p < .001$ ), but the yellow triangles and orange diamonds log slopes were not reliably different from each other,  $t(19) = 1.43$ ,  $p = .168$ . All these slopes were computed including the target-only condition.

As was found in Experiment 1A, the observed linear slopes for the three lure conditions were very small and are consistent with conventionally defined efficient search. The linear slope over the entire range of set sizes was 6.87 ms per item for blue circles, 1.65 ms per item for yellow triangles, and 1.52 ms per item for orange diamonds.<sup>1</sup>

We completed a similar analysis of the observed log intercepts as in Experiment 1A. The observed lure logarithmic intercepts for each lure type (computed without the target only datum) were compared to the target-only condition using a BIC analysis on the blue and orange lures conditions. The BIC results suggested weak and positive evidence, respectively, for the hypothesis that the log intercepts for the blue and orange lure functions did not differ from the target-only condition  $P(H_0|D) = 0.6559$  and  $P(H_a|D) = 0.3441$ , and  $P(H_0|D) = 0.7816$  and  $P(H_a|D) = 0.2184$ . As in Experiment 1A, in the yellow lure case, the intercept did significantly differ from the target-only conditions,  $F(1, 19) = 13.492$ ,  $p = .01616$ ,  $f = 0.8423$ . In sum, Experiment 1B again produced converging evidence that two of the three lure search functions, when plotted on a logarithmic scale, converged on the “target-only” datum. Interestingly, in both experiments, the yellow lures produced intercepts that were slightly elevated compared to the target-only condition. This may be because the yellow lures are the only stimuli that have directional information (i.e., they point rightward and leftward) just like the target, so perhaps the higher elevation of the yellow lures function may be due to a small response-interference effect.

Finally, we should note that we analyzed RTs by position in the display in both Experiments 1A and 1B. These analyses (and similar analysis for every experiment) are presented in Appendix B. The pattern showed an effect of eccentricity, with longer RTs at more eccentric locations.

### Discussion

Three important conclusions can be drawn from Experiments 1A and 1B. First, we found that performance in efficient visual search tasks when the target is fixed can be best described in terms of a logarithmic relationship between set size and RTs. Thus, an important methodological implication is that studying efficient visual search requires a different type of sampling of set sizes than is usually done. Typically, people use a small number of set sizes (as few as two or three levels) that are evenly spaced in linear coordinates (e.g., 3, 6, and 9). Sampling a small number of set sizes may lead to the mistaken impression that there is a linear dependency between lure set size and RT, and it may produce very noisy estimates of search efficiency that will vary depending on the range where set size was sampled. For instance, if one had only computed the search slope over the first three levels of set sizes in Experiment 1A (set sizes 2, 5 and 10), one would have found for

<sup>1</sup> Excluding the target only datum, the linear search slopes were smaller: 5.74 ms per item for blue circles, 0.764 ms per item for yellow triangles, and 1.13 ms per item for orange diamonds, and the linear fit improves some, with  $R^2 = 0.9771$ ,  $R^2 = 0.5363$ , and  $R^2 = 0.6460$ , respectively.

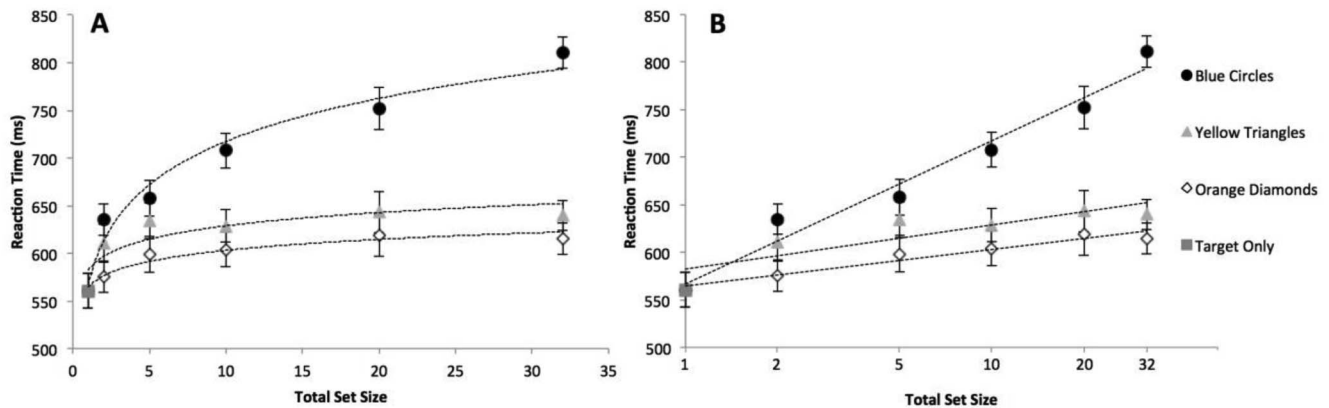


Figure 6. (A) Results of Experiment 1B, plotting set size on a linear scale, showing a logarithmic relationship between number of items and RT (in ms). The  $R^2$  for the logarithmic fit for blue circles, yellow triangles, and orange diamonds were  $R^2 = 0.9681$ ,  $R^2 = 0.7504$ , and  $R^2 = 0.9456$ , respectively. (B) Same data, plotting set size on a logarithmic scale. Error bars indicate the standard error of the mean.

the orange lures a slope of 8.5 ms/item (with an  $R^2 = 0.99$ ). The slope computed over the large set sizes (10, 20 and 32) would have given a very different impression, with a slope of 2.3 ms/item (with an  $R^2 = 0.99$ ). One could have concluded that search had become more efficient at larger set sizes (perhaps reflecting a different type of visual processing or the advent of a new factor, such as display density, influencing the results), while in fact search efficiency had remained identical throughout the entire set size range in logarithmic scale, implying identical processing was at work throughout the entire range of set sizes. Indeed, the log slope computed over the three small set sizes is 41.95 ms per log unit and it is 41.7 ms per log unit for the large set sizes. Identical observations can be made for the other two types of lures.

Second, we observed that logarithmic efficiency changed as a function of lure to target similarity in both Experiments 1A and B: increasing lure-target similarity increased the logarithmic slope of the search function. This reflects the idea that the more similar the lure is to the target, the more evidence needs to be accumulated about a lure to disqualify it as a possible target. This observation is consistent with our simulations based on Townsend and Ashby's (1983) equations for parallel unlimited capacity exhaustive processing. One implication of this finding is that it emphasizes the need to investigate a large range of distractor-target similarity space and it shows that not all of the variation along this similarity space is reflected in changes to stage-two processing (as proposed by Attentive Engagement Theory—Duncan & Humphreys, 1989—see also Alvarez & Cavanagh, 2004). A second implication of this result relates to how we interpret visual search findings for particular feature categories, in particular, the extent that a particular visual attribute can “guide attention” (e.g., Wolfe & Horowitz, 2004). The reason is that our results suggest that we should place less emphasis on visual attributes by themselves and perhaps more on visual similarity as a whole. Even though we did not have a direct measure of similarity, the stimuli of Experiment 1A were designed a priori with increasing levels of subjective similarity (i.e., we made them look more and more similar). The target of Experiment 1B was selected a priori to reverse those subjective similarity relationships, while keeping the

lures the same (see Figure 4B). It is worth noting that this emphasis on overall object-to-object similarity is consistent with the theoretical framework laid out by Ahissar and Hochstein in Reverse Hierarchy Theory (RHT, 2004). Indeed, RHT proposes that when interacting with scenes, we ‘deal’ first with complete, fully parsed, scenes and object representations, rather than with specific low-level features. This is also in line with recent work by Becker, who has showed that feature values by themselves are less important than their relation to the target in various visual search tasks (Becker, 2010; Becker, Harris, Venini & Retell, 2014; Becker, Valuch, & Ansorge, 2014).

A final important aspect of our results is that it highlights the large role that stage-one variability can play in visual search performance. Indeed, looking at Figures 5A and 6A, one can see that RTs differ by almost 150 ms (in 1A) and 250ms (in 1B) between the RTs at the low end of the set size spectrum (with low similarity) and those at the high end of the set size spectrum (with high similarity). This is important because it suggests that, just because a search is efficient, one should not discard or ignore the variability produced by changes in set size and lure-target similarity. For example, imagine trying to interpret search performance when observers are looking for a real world object embedded in a real-world scene (e.g., Wolfe, Alvarez, et al., 2011). Even if the functional set size is one (i.e., participants always only end up scrutinizing a single item—the target), RTs in the task can vary widely depending on the lure-target similarity in that scene: if the target is a single red flower among 20 blue flowers, RTs are likely to be substantially faster than if the single red flower is embedded among 20 orange-reddish flowers. This type of variability will reduce substantially the  $R^2$  of a single function trying to relate RT to set size (see Figure 2). In contrast, Wolfe and colleagues proposed that the lowered  $R^2$  probably reflected the impact of nonvisual guiding factors on performance (i.e., how knowledge and context about a scene tells us where the target might be in that particular scene). Such factors are surely affecting performance. But our results show that visual factors impacting stage-one processing are likely also playing a large role, perhaps one just as important. We will demonstrate in Experiments 3A through 3D



that this variability is observed even when stage-two processing is required to find the target among a subset of candidate targets. But before doing so, in Experiment 2 we present evidence that when distractor-target similarity is relatively high, visual search does become linear as a function of set size, as supported by current theory.

### Experiment 2: Inefficient Search Is Linear and Self-Terminating

In Experiment 1, we used distractor stimuli that were sufficiently different from the target such that stage-one processing could efficiently exclude them as nontargets, producing logarithmic search efficiency. Experiment 2 functions as a control experiment, simply to demonstrate that when target-distractor similarity is sufficiently high, stage-one processing can no longer dismiss distractors as nontargets. As a result, all items require scrutiny, which produces a linear dependency between number of to-be-scrutinized elements and RT. Processing also becomes self-terminating. This is well in accordance to extant theories of visual attention, positing a scrutiny stage at the second stage of processing, where items are compared with a target template (e.g., Alvarez & Cavanagh, 2004; Duncan & Humphreys, 1989; Fisher, 1982; Wolfe, 2005). As stimuli, we used a letter T for the target (pointing right or left) with rotated Ls as distractors. All stimuli were the same color (red).

Experiment 2 differed from Experiment 1 in two ways. Experiment 1 demonstrated the importance of sampling set size with enough levels to be able to assess differences between linear and logarithmic functions. Here, we also sampled set size at six different levels to assess as precisely as possible the linearity of the observed functions. However, pilot work in our lab had shown us that for displays sizes as small as 10 items, with the Ls and T stimuli and the displays we were using, target-absent trials were already getting overly long (mean RT of about 1500ms). In comparison, the longest mean RT in Experiment 1A was 645ms, and in 1B 811ms. Thus, for practical reasons, to be able to fit the experiment in one hour, we settled on a maximum set size of 10. Second, and most importantly, we used a present/absent task. Although using these Ls and T stimuli we anticipated that the RT by set size function would be linear, obtaining a linear fit alone would not be sufficient evidence for arguing that a qualitatively different type of processing was taking place, simply because, mathematically speaking, very steep linear functions can still be fairly well approximated by logarithmic functions over truncated ranges (as with set sizes varying between 1 and 10). The benefit of testing both target absent and target present conditions is that these conditions would allow us to use the slope-ratio logic to conclude that a different stopping rule is at play with these stimuli, than the one we observed in Experiments 1A and 1B. When a self-terminating rule is used, the ratio of the target-absent to target-present search slopes should be at least 2:1, although in practice it is often larger than that. The logic is as follows. If processing is self-terminating, on target-present trials, at any given set size  $N$ , participants will on average inspect  $(N + 1)/2$  items in the display before finding the target and stopping. In contrast, on target-absent trials, participants will not be able to stop until all  $N$  items have been inspected (ergo the  $2*N/(N + 1)$  minimum slope ratio, which is close to 2/1). The ratio is in practice larger than that because on target-absent trials participants often revisit already inspected locations before quitting. In contrast, if processing is exhaus-

tive, target-present and target-absent conditions should produce identical slopes because in both cases, all items in the display are processed. Thus, obtaining both good linear fits for both the target-absent and target-present functions, as well as a slope ratio of at least 2:1 between these two functions would allow us to conclude that, indeed, the type of processing required to find a target among these highly similar distractors (i.e., candidates) is fundamentally different from the one required to find a target among lures.

### Method

**Subjects.** Data from 20 new participants from the same pool as in Experiment 1 were analyzed in Experiment 2. Four participants were replaced because of low accuracy (<90% correct).

**Apparatus and stimuli.** Each display contained either 1, 2, 4, 6, 8, or 10 candidate elements. The nontarget candidates were thin, randomly rotated red Ls. The target candidate was a thin, red T pointing either right or left (see Figure 4B). On half of the trials, the target was present, on the other half it was absent and participants reported the presence/absence of the target T. There were 50 observations per condition. All search displays were randomly generated for each trial, with the only condition being that no more than 2 candidates would appear on any quadrant in the display when the set size was smaller than eight or no more than 3 per quadrant for set size 10. Everything else was identical to Experiment 1.

### Results

In contrast to the results of Experiment 1, the functions for target present and target absent were highly linear (see Figure 7). The  $R^2$  for the linear trend for target-present trials and target-absent trials were 0.995 and 0.991, respectively. This is well in accordance with previous literature using these types of stimuli.

### Discussion

As expected, when all items in the display are sufficiently similar to the target, they all require scrutiny, as evidenced by the linear increase in RT as a function of the number of items. Consistent with current theories, this linear trend is a signature of stage-two processing. Further, the data is also consistent with processing being of a *self-terminating* nature: the target absent to target present slope ratio in this experiment was 2.6. Observing a slope ratio larger than 2 is one of the theoretical signatures of stage-two processing.

Now that we have demonstrated that our Ls and T stimuli require stage-two processing to distinguish target (sideways T) from candidates (oriented Ls), in Experiments 3A through 3D, we will mix these candidate stimuli with different types of lure stimuli in the same displays. Again, the expectation is that lures should not require scrutiny, so their contribution to RT should be logarithmic instead of linear. Thus, the goal will be to document the separate contributions of stage one and stage two on performance, when displays contain both types of stimuli.

### Experiments 3A Through 3D: Evidence for Two Separate Coexisting Stages in Visual Search Revealed by Separately Manipulating Lure and Candidate Elements in a Scene

Experiments 1 and 2 have presented evidence for two different behavioral signatures associated with the processing of two different

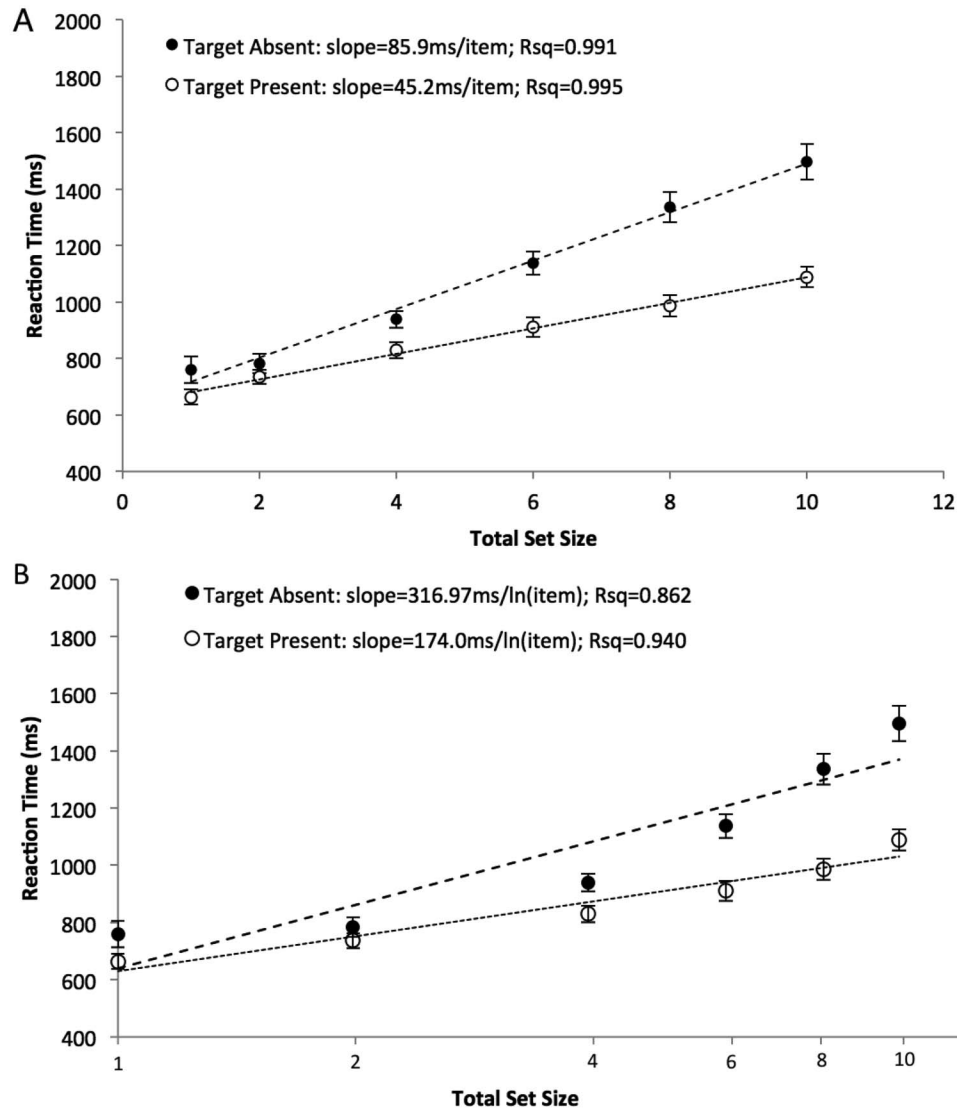


Figure 7. Results from Experiment 2, plotted on a linear (A) and logarithmic (B) scale for set size, showing a linear relationship between number of items and RT (in ms) over a large range of performance (from about 800 ms to almost 1800 ms). The  $R^2$  for the linear fit for target present trials was 0.995 and 0.991 for target-absent trials. Error bars indicate the standard error of the mean.

types of distractors, when studied in isolation. We propose that these behavioral signatures attest to the existence of at least two separate stages in visual search. In the next four experiments, we will investigate how the two stages work together in scenes that simultaneously contain two types of distractor stimuli, some dissimilar to the target (lures) and some very similar to the target (candidates).

As mentioned before, we propose that the first stage of processing is parallel and unlimited in capacity. When used in visual search for a specific item, the first type of processing can simultaneously process *all* elements in a scene and does so in *exhaustive* fashion. The output of this function is a classification of elements in a binary fashion: elements that are sufficiently different from the target are rejected from further consideration, whereas items that are sufficiently similar to the target are passed onto the next stage of processing for further

scrutiny. We refer to this sorting function of stage-one processing as *Screening*. This function is instantiated in parallel across the entire visual field and has two important properties:

1. It has logarithmic efficiency: response times increase as a logarithmic function of the number of items in the display.
2. The steepness of the logarithmic function is directly sensitive to the similarity between the to-be-rejected items and the target template.

The second stage of processing, in tasks like visual search, then implements a *Scrutiny* function, where individual items are closely inspected. That is, the items are precisely compared with the target

template. If the item ends up being the target, during scrutiny, the specific identity of the target is also determined (in our case, which of the two possible Ts is the actual target in the current display). This information is necessary to program a response in the task. This second type of processing is the one that most theories of visual attention have studied the most and is characterized by search functions that show linear increases in RTs as a function of set size and self-terminating processing.

One corollary of the first property of Screening is that *all* items in a display will be processed, and thus all impact RTs. This is a unique contribution of our proposal. Indeed, previous models of attention have all explicitly stated that when the visual difference between a target and the distractors is sufficiently large (efficient search), the number of items in the display does not affect search times (e.g., Bundesen, 1990; Treisman & Gelade, 1980; Wolfe, 1994), or does so in a nonmeaningful manner. In Appendix A, we present a mathematical argument that the logarithmic increase in RTs is evidence that each element in the display is processed independently of one another by an unlimited capacity processor that exhaustively processes all items in the display.

One corollary of the second property, that Screening is sensitive to distractor-target similarity, is that not all efficient searches are equal. Even minute differences in efficient search slopes, when replotted on logarithmic scales, may reflect meaningful similarity dependencies. This is also a departure from current views that propose all efficient searches are largely identical and differences in target-distractor similarity are only reflected in changes to the linear slope relating number of items to search times (e.g., Wolfe & Horowitz, 2004; Duncan & Humphreys, 1989). To be clear, we are not arguing that others did not find effects of set size on RTs in efficient searches, but that no one so far has attempted to interpret those very small slope values because they either thought the slope was not meaningfully different from zero or because it was below the efficient search threshold of 10 ms/item. In fact, a careful look at Treisman and Gelade's (1980) efficient search plots suggests the presence of logarithmic functions like the ones we found in our experiments. In Appendix A, we demonstrate that differences in the steepness of these logarithmic curves can be interpreted as differences in the amount of evidence (or decision threshold) that is required to classify an item as likely or unlikely to be the target (i.e., to *screen* that item). More dissimilar items require little evidence and will produce shallow logarithmic curves. Less dissimilar items will require more evidence before they can be rejected and will thus have steeper logarithmic curves. And, finally, the reason all items always influence RT is that the first stage of processing is exhaustive, at least when participants are looking for a specific target in a scene.

The fact that search does not end after Screening is completed suggests this stage is resolution limited (for a similar proposal see Doherty & Foster, 1999). If there were no resolution limits, Screening would always discard all nontargets and only the target would remain. All searches would be efficient, all search functions logarithmic, and absent/present slope ratios would be 1:1. Because of this limitation in resolution, items that are too similar to the target require focused attention to be disambiguated as target or nontarget elements. The time to complete this second stage of attentive processing, *Scrutiny*, is linearly dependent on the number of to-be-inspected items, as has been repeatedly demonstrated in

the literature (Treisman & Gelade, 1980; Bundesen, 1990; Wolfe, 1994).

Finally, it should be noted that if elements in a display are too close to each other, they will produce visual (Pelli, 2008; Pelli, Palomares & Majaj, 2004) or attentional (Chun & Cavanagh, 1997; Intriligator & Cavanagh, 2001) crowding, situations that can only be resolved by directing overt or covert attention to the crowded location (Neider & Zelinsky, 2008; Reynolds & Chelazzi, 2004). Absent this need for serial inspection, information should accumulate at all locations simultaneously and independently, allowing for Screening to take place. Also, in some extreme circumstances, displays that are overly dense may produce texture segmentation effects where the distractors collectively function as a background to a target that breaks the texture by way of contrast (e.g., Rangelov, Müller & Zehetleitner, 2013). Our displays were designed to minimize such contributions to response times.

Experiments 3A through 3D were designed to characterize the Screening function with different levels of lure-target dissimilarity, when multiple possible targets are present in a scene. We expected logarithmic screening functions, with steeper curves for more similar items, and independence between the observed screening (logarithmic) and scrutiny (linear) functions. Given that we wanted to use the Ls and T stimuli as the candidates for the search task, we needed a different set of lures than the ones used in Experiments 1A and 1B, because none of those resemble a T. Using entirely different lures also allowed us to test the generalizability of our logic to novel stimuli. In Experiment 3A, we used thick orange crosses which resembled the red T target somewhat in color (the orange used had a rich red tint in it) and somewhat in shape (it had straight lines, intersecting at T junctions), but had thicker strokes than the red T (which had single pixel lines). Thus, the lures remained obviously different from red Ts (see Figure 4E). In Experiments 3B and 3C, we aimed at increasing the similarity of the lures to the target red T in two different fashions. In Experiment 3B, we used the same reddish orange crosses, but we drew them with the same thickness as the red Ts (single pixel width), increasing target-lure similarity along the shape dimension. In contrast, in Experiment 3C, we used the same thicker cross shapes as in Experiment 3A, but this time we used the exact same red color as the red T, increasing the target-lure similarity along the color dimension. Finally, in Experiment 3D, we decreased the similarity of the lures to the target by using the same color (reddish orange) of Experiment 3A, but decreasing the similarity along the shape dimension by using square shapes. Using Experiment 3A as a benchmark, we expected to find steeper logarithmic curves for screening lures in Experiments 3B and 3C than in 3A, and shallower logarithmic curves in Experiment 3D than in 3A. Importantly, we predicted that, keeping the number of candidates constant, we should find logarithmic increases in RT as set size increases (i.e., as the number of lures increase). This should be true both for displays containing 4 candidates and those containing 8 candidates.

## Method

**Subjects.** Data from 80 new participants from the same pool as in Experiments 1 and 2 were gathered in Experiment 3A through 3D (20 participants per experiment). No subjects needed replacement in Experiment 3A, four were replaced in Experiment

3B, three in Experiment 3C, and one in Experiment 3D because of low accuracy performance ( $<90\%$ ).

**Apparatus and stimuli.** The target candidate was the same thin red T pointing right or left used in Experiment 2. It was presented with either 3 or 7 nontarget candidates, the same thin red oriented Ls used in Experiment 2. However, as in Experiments 1A and 1B, we asked participants to report the direction the target T was pointing toward (right or left) and all trials contained a target. All search displays were randomly generated for each trial, with the only condition that no more than 2 candidates would appear on any quadrant in the display. Zero, 4, 8, 16, or 28 lures were presented in addition to the candidate elements (see Figure 4C for an example of display). Lure items were, respectively, thick orange crosses (Experiment 3A), thin orange crosses (Experiment 3B), thick red crosses (Experiment 3C), and orange squares (Experiment 3D). Figure 4D provides an illustration of the four types of lures. There were 60 observations per condition for a total of 600 trials per experiment.

## Results

**Experiment 3A: Thick orange cross lures.** The mean error rate was 2% and did not vary by search condition (all  $F$ s  $<1.2$ ). A two-way ANOVA with Candidate number (4 vs. 8) and Lure number (0, 4, 8, 16 or 28) as within-subject factors on correct RTs showed a significant main effect for number of candidates,  $F(1, 19) = 422.43$ ,  $p < .001$ ,  $f = 4.718$ , and for number of lures,  $F(4, 76) = 26.24$ ,  $p < .001$ ,  $f = 1.175$ . The interaction was not significant,  $F(4, 76) = 0.38$ ,  $p = .82$ , reflecting that RTs increased as the number of candidates increased as well as the number of lures increased.

To better understand these effects we looked at RT as a function of the number of lures, separately for each level of candidates (see Figure 8A). Consistent with Experiment 1 we found that logarithmic functions fitted the RT curves when displays contained four candidates ( $R^2 = 0.93$ ) and when displays contained eight candidates ( $R^2 = 0.89$ ), as the number of lures increased from 0 to 28. Henceforth we will refer to these functions as Screening functions.

Critically, the absence of an interaction between number of candidates and number of lures reflects the fact that the slope for searching through the candidates was unaffected by the number of lures. As can be seen in Figure 8A, the slope of the scrutiny functions (indicated by the dotted lines joining 4 and 8 candidates for any given number of lures) remained parallel throughout the screening function. In other words, the efficiency with which participants scrutinized the candidates when there were *no lures* in the display was the same to the efficiency with which participants scrutinized candidates when there were 4, 8, 16, or 28 lures present in the display. Henceforth we will refer to these functions as Scrutiny functions, and we assumed that they are linear based on the results from Experiment 2, even though here we only tested two levels of candidates. The average slope of the Scrutiny functions for 0, 4, 8, 16, and 28 lures was 65 ms/item.

To conclude, according to additive factors logic (Sternberg, 1969, 1998), the results of this experiment suggest that Screening and Scrutiny are different functions occurring at separate stages of processing, with screening occurring before scrutiny.

**Experiment 3B: Thin orange cross lures.** A two-way ANOVA with Candidate number (4 vs. 8) and Lure number (4, 8,

16 or 28) as within-subject factors on correct RTs showed a significant main effect for number of candidates,  $F(1, 19) = 343.3$ ,  $p < .001$ ,  $f = 4.270$ , and for number of lures,  $F(1, 19) = 31.863$ ,  $p < .001$ ,  $f = 1.294$ , but no significant interaction,  $F(4, 76) = 2.10$ ,  $p = .089$ . The logarithmic fit ( $R^2$ ) for the 4-candidate and 8-candidate Screening functions were both 0.97. The average slope of the Scrutiny functions for 0, 4, 8, 16, and 28 lures was 69 ms/item.

**Experiment 3C: Thick red cross lures.** The two-way ANOVA on correct RTs showed a significant main effect for number of candidates,  $F(1, 19) = 589.13$ ,  $p < .001$ ,  $f = 5.591$ , and for number of lures,  $F(1, 19) = 55.97$ ,  $p < .001$ ,  $f = 1.718$ , but no significant interaction,  $F(4, 76) = 1.29$ ,  $p = .28$ . The logarithmic fit ( $R^2$ ) for the 4-candidate and 8-candidate Screening functions were 0.97 and 0.83, respectively. The average slope of the Scrutiny functions for 0, 4, 8, 16, and 28 lures was 76 ms/item.

**Experiment 3D: Orange square lures.** The two-way ANOVA on correct RTs showed a significant main effect for number of candidates,  $F(1, 19) = 680.04$ ,  $p < .001$ ,  $f = 6.003$ , and for number of lures,  $F(1, 19) = 8.961$ ,  $p < .001$ ,  $f = 0.6766$ , but no significant interaction,  $F(4, 76) = 2.386$ ,  $p = .059$ . The logarithmic fit ( $R^2$ ) for the 4-candidate and 8-candidate Screening functions were 0.93 and 0.68, respectively. The average slope of the Scrutiny functions for 0, 4, 8, 16, and 28 lures was 56 ms/item.

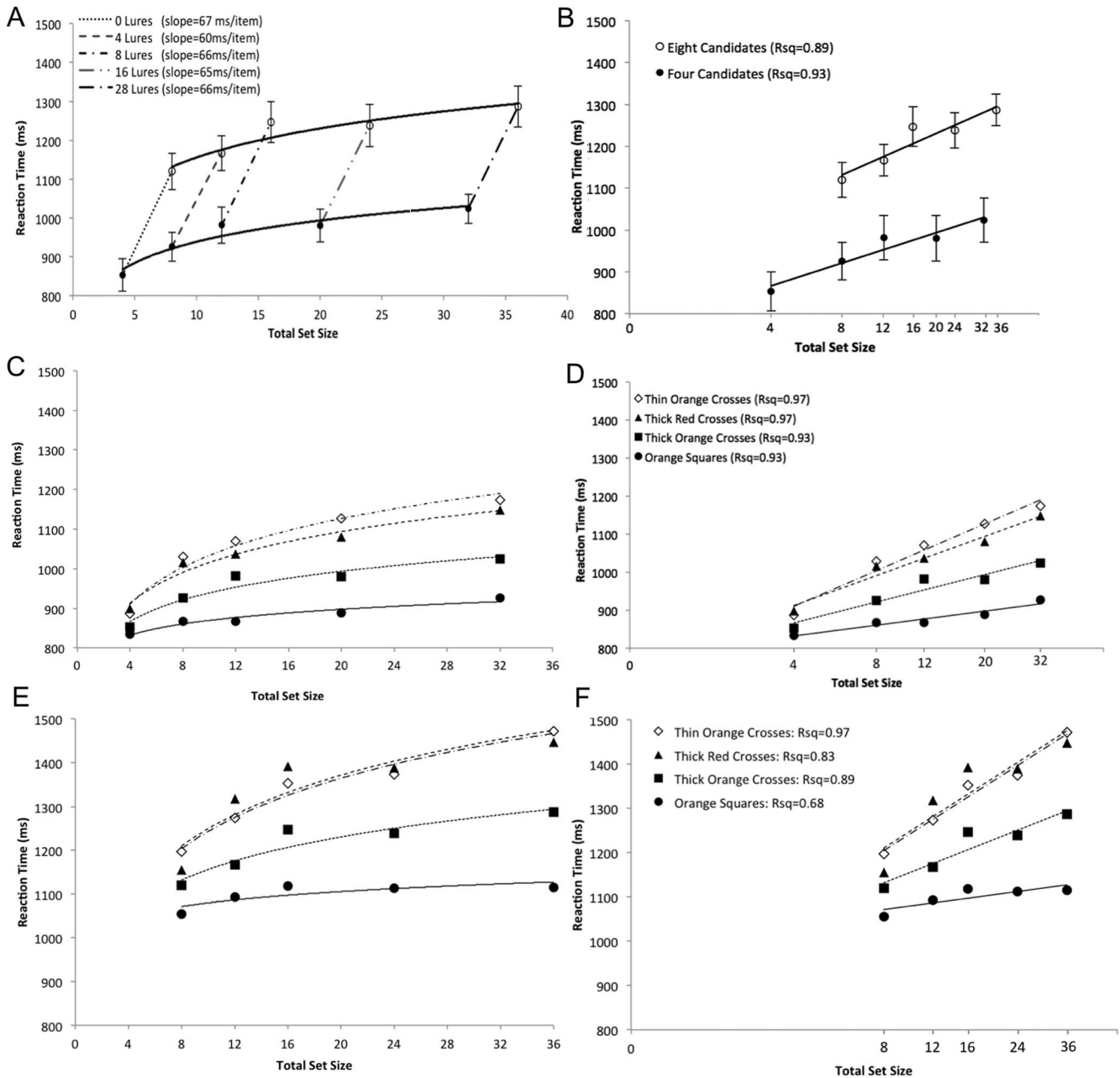
Figure 8C and 8E shows screening functions for 4 and 8 candidates, respectively, obtained across Experiments 3A through 3D where lure-candidate similarity was varied. As predicted, an orderly arrangement of screening functions by similarity was found both when there were four and eight candidates in the display. For the four candidates condition, the slope of the log functions for orange squares (40 ms/log unit) was significantly smaller from the slope for thick orange crosses (79 ms/log unit,  $t(38) = 2.84$ ,  $p = .007$ ), which was itself smaller than the slope for thick red crosses (112 ms/log unit,  $t(38) = 2.16$ ,  $p = .037$ ). The slope for thick red crosses did not differ from the slope found for thin orange crosses (134 ms/log unit,  $t(38) = 1.11$ ,  $p = .276$ ). Analogous results were found for the eight candidate log slopes. The slope of the log functions for orange squares (37 ms/log unit,  $t(38) = 3.17$ ,  $p = .003$ ) was significantly smaller from the slope for thick orange crosses (97 ms/log unit,  $t(38) = 2.84$ ,  $p = .007$ ), which was itself smaller than the slope for thick red crosses (176 ms/log unit,  $t(38) = 2.51$ ,  $p = .016$ ). The slope for thick red crosses did not differ from the slope found for thin orange crosses (175 ms/log unit,  $t(38) = 0.023$ ,  $p = .982$ ). In sum, consistent with the results of Experiments 1A and 1B, lure-target similarity determined the slope of logarithmic function relating number of lures to overall RTs.

A position analysis of the data from Experiments 3A through 3D showed no effect of eccentricity on RTs (see Appendix B).

## Discussion

When using heterogeneous displays containing both lures and candidates, we observed that the number of total items in a display impacted RTs logarithmically. We interpret this pattern as the signature of the stage-one screening function, where elements are





**Figure 8.** Results from Experiments 3A through 3D showing time (in ms) to find a target (an oriented red T) as a function of the number of elements in the display, shown separately for displays containing 4 and 8 candidates (red oriented Ls), among a varying number of lures. The full lines show the best fitting logarithmic trend for each series, and the corresponding measure of fit ( $R^2$ ). (A) Data from Experiment 3A: the dotted lines visualize the scrutiny functions for each level number of lures. The slopes for the scrutiny function when no lures were presented (0 lures) was 67 ms/item and was no different than when 4, 8, 16, or 28 lures were present. Error bars indicate the between subject standard error of the means. (B) Data from Experiment 3A presented on a logarithmic scale. (C through F) Combined data from Experiments 3A-D showing logarithmic screening functions when 4 (panel B) or 8 (panel C) candidates are present in the display in linear (panels C and E) and logarithmic (panels D and F) space, indicating orderly logarithmic sensitivity to target-lure dissimilarity.

categorized as lures or candidates. This effect on RTs appears to be separate from the effect of the number of candidates<sup>2</sup> present in a scene, which is characterized by a linear increase in RTs. This indicates that when searching, first, locations unlikely to contain the target are discarded, and then, locations that are likely to contain one are scrutinized. Furthermore, the logarithmic effect of lures on RT was once again modulated by the similarity between lures and the target template, with the more similar items producing steeper processing costs. Similar effects on performance can be achieved through different avenues, as similarity to the target template can be increased along different channels. Starting with a baseline similarity established in Experiment 3A, we were able to increase similarity to the target template to equal effect on performance by either increasing similarity along the shape dimension (Experiment 3B) or along the color dimension (Experiment 3C).

Furthermore, the results of Experiments 3A through 3D demonstrated two clearly different families of visual distractors in visual search (see also Lleras, Buetti & Mordkoff, 2013), which affect behavior on fundamentally different manner: lures and candidates. In fact, the results suggest that lacking a correct classification of visual stimuli into lures and candidates, it should prove almost impossible to understand, or even empirically evaluate, search performance in a scene containing both types of elements (see Wolfe, Alvarez, et al., 2011a). We will return to this point in the General Discussion.

Returning to the observed data, it should be noted that the logarithmic fits for the 8 candidate condition in Experiments 3C and 3D were lower than in all the other screening functions reported here. However, these logarithmic fits ( $R^2 = 0.83$  in 3C and  $R^2 = 0.68$  in 3D) were still substantially higher than the linear function fits through those same data points ( $R^2 = 0.67$  in 3C and  $R^2 = 0.48$  in 3D). It is unclear to us why these log functions were flattened at higher lure set sizes. Particularly surprising is that these deviations from logarithmic fits were observed at opposite ends of the similarity spectrum: judging by the steepness of the screening functions, 3C and 3D had the lures with the highest and lowest similarity to the red target in this stimulus set, respectively. Thus, deviations from the expected logarithmic functions do not appear to be systematically related to lure-target similarity. On the other hand, simulations in our lab indicate that the more candidates there are in a scene, the noisier the accumulation process becomes, so these deviations might be a result of the relatively high number of candidates in those displays, rather than something systematic about the rejected lures.

We should reemphasize that according to traditional theories of visual search where search unfolds in decreasing order of similarity to the target (see Itti & Koch, 1999; Wolfe, 1994), neither the number of lures nor the type of lures present in the display should have affected RTs in Experiments 3A through 3D. This is because in any given display, the items most similar to the target template are the candidates, and a target is always contained within this set. Thus, observers should always find a target prior to having to inspect any of the more dissimilar items, like the orange crosses or squares. What is remarkable is that the number and type of lures in Experiments 3A through 3D account for differences as large as 300 ms between conditions (see Figure 8), a substantial amount of variance that, up to now, has been neglected by current theories. We will further discuss the implications of this finding in the General Discussion.

In sum, unlike what is proposed by current theory, our data suggest visual search unfolds first by discarding unlikely targets via Screening. Any item that did not get screened out is 'selected,' and by selected, we mean that its location will be tagged as requiring scrutiny. Our data is inconsistent with attentional selection in search being an *active process* whereby object salience (or activation or similarity to a template) determines the order in which items will be scrutinized in a scene, as most visibly proposed by the attention theories of Bundesen (1990); Duncan and Humphreys (1989), and Wolfe (1994). That said, because we have not yet varied stimulus similarity within the candidate set, we cannot rule out that *scrutiny among candidates* will follow a priority as proposed by these theories. At this point, we prefer to think that scrutiny of likely locations occurs in random fashion, as a more parsimonious account, but more testing is required to conclusively rule on this issue.

There is one final caveat to our results that is worth mentioning. Because we did not measure eye movements in these experiments, we do not know yet what the contribution of eye movements is to the pattern of results observed here. In particular, we do not know *when* the eyes start to move with respect to the screening process. It could be that the eyes move after the screening is completed. Or it could also be that the eyes begin moving before screening is complete. This would necessitate some screening of the display be done after each eye movement, over the entire display or perhaps over a small subsection of the display. Finally, it is also possible that the eye scanning of the scene and the screening process unfold in more or less simultaneous fashion and interact in more complex ways than we have foreseen at this point. More research is needed to know the answer to these questions.

#### Experiment 4: Oddball Search Produces Negative Slopes

So far, in our experiments, we have studied search tasks where the target is known and fixed throughout the experiment. In Experiments 1A, 1B, and 3A through 3D, participants were required to find the target and report an incidental property about it. While maintaining the target constant, we then varied other aspects of the scene such as what type of distractors were present, how similar they were to the target and how many of them there were. We found evidence for a parallel architecture that initially processes the visual input with unlimited capacity and in exhaustive manner, with the *function* of screening the scene for potential target locations. Here, we ask a different question: does changing the observers' task change the function performed by this architecture? In other words, is the visual analysis of the scene in stage one

<sup>2</sup> We want to note that we designed our experiments to best isolate and characterize the existence of the screening function and its theoretical independence from the scrutiny function. Although the statistics here suggested independent stages, it is quite possible that these two stages are not strictly sequential nor strictly independent. Different display arrangements and experimental conditions, we believe, may produce situations where the first stage may be truncated or even eliminated, if for instance one actively encourages participants to move their eyes (see Lleras & von Muhlenen, 2004), or when candidates appear in clusters or when one asks participants to implement systematic search strategies (such as searching left-to-right, up-to-down).

fundamentally altered by the type of task observers perform on the scene?

To tackle this issue, we focused on studying displays that *ought* to produce efficient visual search. When performing an efficient search, does knowing the precise target (as in Experiment 1) produce a fundamentally different form of stage-one processing than not knowing the precise identity of the target a priori? A version of the latter task is known as the “oddball search task,” where one item differs visually from all others along one specific feature (say, color), so much so as to produce the phenomenon known as visual pop-out. To our knowledge, there has actually been no direct comparison in the literature between the two tasks. The closest literature is the one on contingent capture, where the question has been asked in the following manner (Lamy, Carmel, Egeth & Leber, 2006; Leber & Egeth, 2006). How does the likelihood that a salient or unexpected object will capture attention change as a function of whether observers are looking for a specific target (a.k.a., feature-search mode) or for a unique target in the display (a.k.a., singleton-search mode). This is a different question from the one we pose here. And the answer to that question does not necessarily address whether the display, as a whole, was processed fundamentally differently when observers used a feature-search mode compared to a singleton-search mode.

Many computationally oriented approaches to vision set aside the issue of implementing or considering an observer’s task, and focus instead on implementing computations such as feature contrasts (e.g., Itti & Koch, 2000, 2001), or feature-suppression mechanisms (e.g., Li, 1999), or object segmentation (Walther & Koch, 2006), or summary statistics of the display (e.g., Rosenholtz et al., 2012). And, in most cases, these computations are agnostic as to what an observer might actually be doing with the visual information in the scene. The perhaps implicit assumption then, is that changing the task would change little to this computation. To be clear, we are not saying that the computation would not change at all. Clearly, since the influential theory of Guided Search was first proposed (Wolfe, 1998), the idea that top-down goals alter some aspect of stage-one processing has been implemented by most theories of visual search (e.g., Eckstein et al., 2000; Najemnik & Geisler, 2005; Palmer, Ames & Lindsey, 1993). However, for these theories, the influence of task on stage one is limited to the influence a specific target template has on the processing of a display. For example, knowing that the target is red does something like boost the gain for processing locations that contain the feature red in the display (see Verghese, 2001 for a review). As a result, this type of influence on visual processing produces quantitative changes in processing, *not qualitative ones*. Further, when visual cognition investigators have used oddball search tasks (as in the large literature on priming of pop-out, e.g., Lamy & Kristjánsson, 2013; Maljkovic & Nakayama, 1994, 1996), the assumption has often been that the underlying stage-one computations remain fundamentally the same, and that the presence of effects like intertrial priming reflect modulations of otherwise identical low-level computations: the boosting of a feature associated with target on a preceding trial (Wolfe, Butcher, Lee, & Hyle, 2003) or the inhibition of a feature associated with distractor status (Lamy, Yashar, & Ruderman, 2013). In sum, although one might expect that changing the experimental task will fundamentally alter how stage one processes visual stimuli, there is no actual direct evidence in the literature that it does.

Experiments using oddball-search tasks have documented a rather unique finding: that RTs decrease as set size increases, a so-called negative search slope. The modal interpretation of those findings is that specific display properties are responsible for this phenomenon. For example, that increasing set size may increase target contrast (Meeter & Olivers, 2006): the more distractors there are, the closer some of them will be to the target, and thus, the higher the target-distractor contrast signal at that location would be (see also, Bravo & Nakayama, 1992). Such proposals are also consistent with computational theories where, indeed, increasing set size *might* influence feature contrast, either because of feature-suppression mechanisms (Li, 1999, 2002) or because feature-contrasts are computed over different spatial scales, iteratively until a target is found (Busch & VanRullen, 2010; Itti & Koch, 2000; VanRullen, Carlson, & Cavanagh, 2007; VanRullen & Koch, 2003). In sum, the implication is that if these oddball-search studies had used the same displays but changed the task to be a fixed-target search, the RT functions would have still shown a negative search slope.

One could argue, then, that the existence of negative search slopes cast a shadow of doubt upon our discovery of logarithmic efficiency in visual search because it is mathematically impossible for a logarithmic function (which is a monotonically increasing function) to produce a downward trend. For example, Bravo and Nakayama (1992) used displays in which the stimuli were simple diamonds (missing either a right or left corner) that differed in color (e.g., a red target diamond among green distractor diamonds). The authors observed decreasing RTs when the number of distractors increased (see also, Meeter & Olivers, 2006). On the surface, the stimulus set is very similar to one used in Experiment 1 of the present study (red triangle among yellow triangles). However, in Bravo and Nakayama (as well as many of the priming of pop-out studies that followed), the task was *to find the color oddball* in the display. Participants did not know before the trial which color was associated with target status and which color with distractor status. Thus, they did not have a target template in mind against which to compare the incoming sensory information. This was not the case in our Experiment 1 where the target was fixed throughout the experiment, and subjects therefore had a well-defined target template.

We propose that the negative search slopes obtained with Bravo and Nakayama’s type of displays do not reflect a property of the stimulus (as many have generally assumed), but rather a property of the task (oddball search). To test this idea we used the red and yellow triangle stimuli from Experiment 1 and simply changed the task into an oddball search task. Displays varied in set size and target color changed randomly from trial to trial. If our logarithmic functions are attributable to the stimuli and display configurations we used in our experiments, changing the task should have little effect on the overall pattern of results and we should still observe logarithmically increasing RTs when set size increases. On the other hand, if oddball-search tasks tend to produce negative search slopes (and this pattern of results is independent from the stimulus set used), we should be able to change the logarithmic pattern of results observed in Experiment 1 into a pattern with negative search slopes (consistent with the previous literature on oddball tasks).

## Method

**Subjects.** Data from 20 new participants from the same pool as in Experiments 1 through 3 were analyzed in Experiment 4. All subjects were tested for colorblindness. One subject's data was replaced due to having response Times 2 standard deviations above the group mean. No subjects were excluded because of low accuracy.

**Apparatus and stimuli.** The target was either a red triangle among yellow triangle distractors, or a yellow triangle among red triangle distractors. The stimuli (yellow and red triangles) were the same as in Experiment 1. The spatial arrangement of stimuli in this experiment was the same as Experiment 1. Participants were asked to find the oddball and report the direction the target was pointing. There were six set sizes: 1, 3, 5, 10, 20, and 32. We used a set size of three (rather than two) because an oddball cannot be defined when there are just two items in the display. Also like in Experiment 1A, a target was present on every trial. The target was equally to be red or yellow on any given trial.

## Results

A two-way ANOVA was run with Target color (red vs. yellow) and Set size (3, 5, 10, 20, and 32) as within-subject factors on correct RTs (see Figure 9). We did not include the data for set size 1 in the analysis as it represents a case where there is no uncertainty as to which item in the display is the target. We included this condition in the experiment only to have a point of comparison to Experiment 1A. The results showed that average RTs decreased with increasing set size,  $F(4, 76) = 34.71, p < .001, f = 1.351$ . There was no reliable effect of Target color, RTs were similar for yellow oddballs and red oddballs,  $F(1, 19) = 3.83, p = .065, f = 0.449$ . Additionally, there was no significant color by set size interaction,  $F(2.40, 76) = 2.033, p = .129, f = 0.327$  (using a Greenhouse-Geisser correction). What is critical to our predictions is that for both types of oddballs, RTs decreased with increasing set size: a negative search slope effect. More specifically, the exact

same displays (red target among yellow lures) that produced a monotonically increasing logarithmic function in Experiment 1A, now produced a monotonically decreasing function as set size increased for set sizes larger than 3. This decreasing function appeared to be equally well fitted by a negative power and by an exponential function (R squares for linear fit: 0.579; exponential fit: 0.983 power fit: 0.997).

In contrast with Experiments 1A and 1B, the position analysis showed no effect of eccentricity on RTs in Experiment 4 (see Appendix B).

## Discussion

The results of Experiment 4 confirm our hypothesis that negative slopes are not produced by low-level differences between target and distractor stimuli, but rather that negative slopes are obtained in oddball search tasks. The results of Experiment 4 suggest that stage-one processing can alter its functions/computations to fit the experimental task while maintaining the same architecture (parallel, unlimited capacity, resolution-limited processing). In contexts where the target template is known and fixed across trials, this stage exhaustively processes all elements comparing them to the template in order to identify likely from unlikely target locations. In contexts where the target and distractor colors can swap roles between trials, the first stage of processing must first gather evidence regarding which color is repeated in the display so as to screen out those locations from further scrutiny. The latter process actually takes much longer than the former (as shown in Figure 9), particularly at small set sizes, where RTs in target-fixed condition can be well over 100 ms faster than under oddball conditions.

It is important to remember that there is substantial evidence that oddball search tasks are stage-one tasks, that is, that only one item (the oddball) is scrutinized in these displays. A number of electrophysiological experiments conducted on color oddball search tasks show that only one movement of attention (one scrutiny, as indexed by the N2pc) takes place in these displays

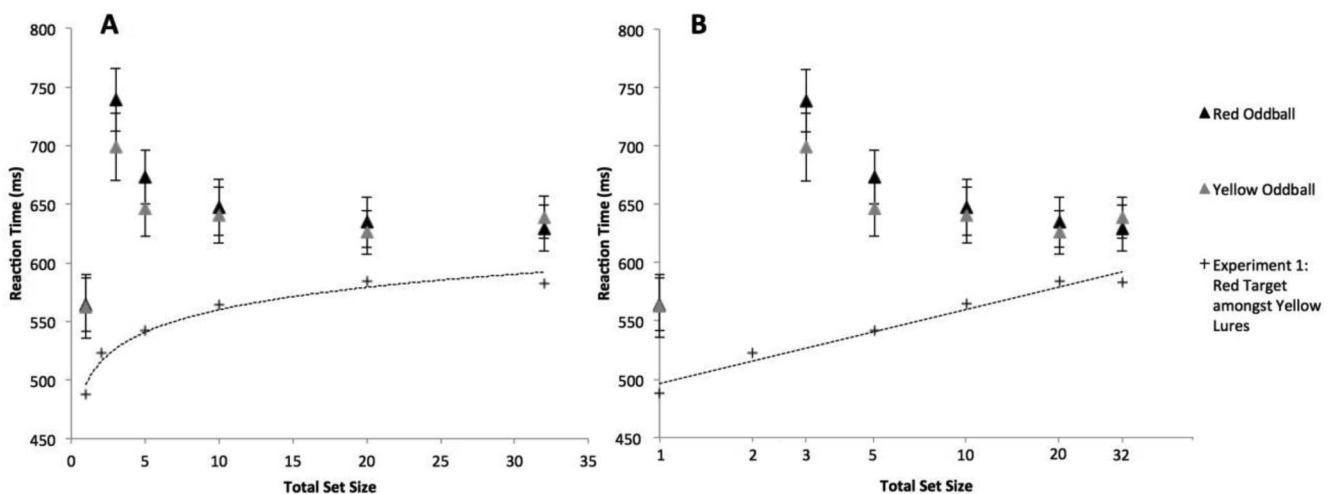


Figure 9. Results from Experiment 4 showing mean RTs as a function of total set size in linear (A) and logarithmic (B) coordinates. The dotted lines represent the data from the red target among yellow lures condition of Experiment 1, replotted for comparison purposes.



(e.g., Eimer, Kiss & Cheung, 2010; Shin, Wan, Fabiani, Gratton, & Lleras, 2008). Additionally, the first (and often only) saccade produced in these tasks is directed toward the target, with few miss-fixations directed at distractors (see Becker, 2008; Caddigan & Lleras, 2010). In sum, taken together, the electrophysiological and eye-tracking evidence indicates that stage-two processing is not driving the negative search slopes observed in oddball paradigms. That is, in these oddball displays, participants always end up scrutinizing one item (the target). The negative slope is therefore a result of processing that occurred prior to the scrutiny stage (a stage-one process, as we propose here). Our lab recently proposed that oddball search tasks can be well understood as attention decision-making tasks, where the goal of the first stage of attention is to determine which color contains the target and which one is shared by distractors (e.g., Lleras & Buetti, 2014; Tseng, Glaser, Caddigan & Lleras, 2014). The first stage does this by accumulating evidence in parallel (and with unlimited capacity) to decide which color is repeated in more than one item, and thereby which color is associated with distractors. Once it reaches a conclusion, it can eliminate those distractor elements from consideration and pass on the (lone) location of the unrepeatable color to the second stage of processing (see Tseng et al., 2014 for a computational model of this decision task and the behavioral benefits that ensue when features repeat or change across trials). The target emerges as the only scrutinized item in the display because of all this stage-one processing.

Our account stands in contrast with previous conceptions of these pop-out tasks, where it was thought that because the target was a feature singleton, it had to “pop-out” from the distractors. In this context, if RTs decreased with increasing set size, it probably meant the local contrast of the target increased as the number of distractors increased. Consequently, a negative search slope was a property of the stimulus set and arrangements. Our data do not support such accounts or intuitions, given that the same target-distractor contrast was present in the target-fixed condition in Experiment 1 and the color oddball task in Experiment 4. In fact, Bravo and Nakayama (1992) proposed that the negative search slope they observed in their oddball search task was the direct consequence of participants not knowing the identity of the target ahead of the presentation of the display (i.e., stemming from the fact that the task was an oddball task). In the absence of a fixed target template to guide attention in a top-down fashion, observers could only find the target with bottom-up information, such as local feature contrasts. Overall, their account suggests a bottom-up account of negative search slopes, though they failed to discuss why this bottom up information would not have also contributed to negative search slopes in fixed-target searches, nor did they discuss why RTs in oddball search tasks are overly long in comparison to those elicited by other pop-out search tasks. One would assume that if the bottom-up contrast is an available source of information for guiding attention, then this information would be extracted much faster and help produce RTs much smaller than the typical 600ms (or more) often observed in these oddball tasks. In sum, we question the idea that the determining factor in the type of oddball search displays used in Bravo and Nakayama and here in Experiments 1 and 4 is a change in bottom-up signal as set size increases. We prefer to think that these tasks, particularly at small set sizes, are more akin to the attention decision-making tasks described above (see also, Tseng et al., 2014), where a top-down

evaluation of all (or at least a subset of) stimuli is very much needed to produce a response in the task. This proposal is also consistent with fMRI data showing activation in the superior attentional control network (areas like FEF and IPS) during this color oddball task (Kristjánsson, Vuilleumier, Schwartz, Macaluso & Driver, 2007).

## General Discussion

The results of the current experiments are consistent with the view that what is considered stage one in most two-stage models of visual search is implemented through a parallel, unlimited capacity architecture. We inferred that the resolution of this first stage is somewhat limited when items fall outside of the fovea. If this were not the case, all searches, irrespective of target-distractor similarity, would show logarithmic efficiency and the functional set size would always be one (the target). These conclusions are largely consistent with most attention theories of visual search. One critical point where our findings significantly advance the literature is in the demonstration that the first stage is exhaustive in nature when the task of the observer is to find a specific target in a scene (Experiments 1A and 1B). Furthermore, our experiments allowed us to differentiate behaviorally the separate contributions of stage-one and stage-two variability to RTs (Experiments 3A through 3D). Future visual search theories should attempt to incorporate into their models the separate contributions to RT of number of lures (logarithmic) and number of candidates (linear). Similarly, we believe the results of Experiments 3A through 3D highlight the importance of moving away from homogeneous search displays (with all lure or all candidate elements) and toward more heterogeneous displays that include both types of stimuli. It goes without saying that in the real world, the latter case is likely to be encountered more frequently than the former. Finally, we believe our results are the first to directly demonstrate that the task observers perform on very simple displays can fundamentally alter the manner in which those displays are processed (comparing Experiments 1 and 4).

## Prior Logarithmic Findings in the Literature

We are not the first to document the existence of logarithmic functions in the search literature. In the seminal work by Treisman and Gelade (1980), evidence pointed toward efficient search being logarithmic and *against* linearity: in their Experiment 1, the authors noted the RT for set size one deviated significantly from linearity with respect to the remaining RTs for larger set sizes. Further, a quick visual examination of the figures in that paper suggests a logarithmic function was at play in the efficient search conditions tested. Several other contemporary studies also found evidence of *negatively accelerating* RT functions (e.g., Briggs & Swanson, 1970; Corballis, Kirby, & Miller, 1972; Kristofferson, 1972; Simpson, 1972; Swanson & Briggs, 1969). Briggs (1974) argued that these functions were likely better understood as logarithmic functions than linear (or other) functions. More recently, Palmer, Ames and Lindsey (1993, see also, Palmer, 1995 and Carrasco, McLean, Katz, & Frieder, 1998) reported logarithmic effects of set size on RT under conditions where crowding was minimized. The authors proposed a signal-detection inspired decision model for visual search that is similar in some respects to

what is proposed here. In particular, Palmer et al. proposed that attention determines what stimuli are excluded from further attentive processing. However, Palmer and colleagues proposed a form of preattentive visual processing, much like a passive filter that simply blocks a subset of stimuli from further processing. Our current data suggest that the time spent on filtering is dependent on lure-target similarity as well as on the number of items in a scene. In fact, we do not envision this filtering as a passive process, but rather as an active visual process (see more below). Furthermore, we propose here that the stage-one architecture can implement different processing *functions* depending on the task (see Experiments 1 and 4), so it does not always work as a single-threshold filter. In addition, we recently presented data showing that in visual search displays containing different types of lures simultaneously (each with its own lure-target similarity coefficient), each item in the display is processed up to its own threshold, irrespective of the identity of other items in the display (Madison, Buetti, & Lleras, 2015). In other words, each item is processed to the point where a decision can be made about it (the decision being: is this a likely target or not) independent of what other items are present in the display. Lures more dissimilar to the target will have a decision process that requires less evidence accumulation than more similar lures. Thus, stage-one processing is not a fixed-threshold stage, but rather, it is a stage where a set of multiple decision thresholds can simultaneously exist. These thresholds reflect the set of similarity relations between the lures present in that display and the target template. Finally, Carrasco and Yeshurun (1998, see also Carrasco, Evert, Chang, & Katz, 1995) also reported logarithmic dependencies on set size, though, these logarithmic dependencies were not documented concurrently with linear dependencies, as we found in Experiment 3. Carrasco and colleagues' results accentuate that manipulations of set size in visual search experiments have effects other than uniquely changing the degree of scrutiny required by the display: set size manipulations can also lead to differences in the ability to accurately represent the information in the display, a conclusion with which we agree.

### Impact of Observer's Goals on Stage-One Processing

One of the goals of the present paper is to present evidence that there is an important and substantial source of variability in visual processing times that has been overlooked thus far in the attention literature. As our data demonstrated, during the first stage of processing there are at least three sources of variability. The first is set size. Although most visual cognition theories assume that set size does not impact stage-one completion times, here we demonstrated that in fact it does in tasks where the target is fixed (Experiments 1A and 1B, 3A through 3D) or when the target is an oddball (Experiment 4), which presumably make up a majority of search tasks.

A second source of variability is lure-target similarity. Perhaps the most well-known theory arguing for a role of distractor-target similarity on performance is Attentional Engagement Theory (Duncan & Humphreys, 1989, 1992), which claims lure-target similarity affects the scrutiny time of items in the display: the more similar an item was to the target template, the longer it would take to decide whether the item currently being scrutinized was the target or not, leading to steeper search functions for more similar items. Other theories like TVA (Bundesen, 1990) and Guided Search (Wolfe, 1994) propose feature-based decompositions of the

stimuli that could in theory produce similarity effects. For instance, imagine that observers look for a red target in a scene. Elements that have a color closer to red (like orange) would be more likely to benefit from the top-down "boost" meant to prioritize red items in the scene, than elements that have colors farther from red (like green, blue, or yellow). In practice, however, such stage-one similarity effects were not thought to produce any meaningful variability in search times. In contrast to these feature-based decomposition theories, we think of similarity more along the lines of Reverse Hierarchy Theory (Ahissar & Hochstein, 2004). In RHT, the pertinent level of analysis is not at the isolated feature level, but rather at the more complete object representation stage (see also, Henderson, Malcolm & Schandl, 2009 for a similar proposal). Similarity is computed not necessarily along specific feature channels, but at a more global or holistic level. The results of Experiments 3B and 3C might be telling on this regard: lures in these experiments produced the same steepness of logarithmic screening functions, though each type of lure was similar to the target for different reasons (3B lures were more similar in the shape dimension and less so along the color dimension than 3C lures were to the target). More research is needed to fully understand whether overall similarity can be thought of as the simple combination of similarity along separate feature channels (as proposed for instance in the Target Acquisition Model of Zelinsky, 2008), or whether similarity should be thought of more holistically. Finally, our data show that both set-size and target-lure similarity are sources of variability impacting the same stage of processing, as they interact with one another (Donders, 1968/1969; Sternberg, 1969).

A third source of variability in processing times is the task goal of the observer. With identical displays, we demonstrated that response times can have qualitatively different dependencies on set size when the task is varied. In Experiments 1A and 1B, target-fixed searches produce monotonically increasing RTs as a function of set size, whereas an oddball-search task produces monotonically decreasing RTs (Experiment 4). This is in itself an important finding because many in the literature had assumed that stage-one processing would be qualitatively similar across both types of tasks (e.g., Wolfe et al., 2003). In other words, having a fixed target would produce a behavioral benefit over not having a fixed target template because, for example, the top-down boosting for a specific target template could build up across multiple trials and confer larger benefits (see also, Maljkovic & Nakayama, 1994; Wolfe et al., 2003). Such an effect would imply a quantitative modulation of stage-one processing, not a qualitative one. Thus, such accounts could never explain why monotonically increasing functions turn into monotonically decreasing ones.

In sum, in this study we found three different sources of stage-one variability. Two of these sources seem to fall squarely in the category of "top-down" effects: lure-target similarity and task goal. First, the lure-target similarity effect should be interpreted as a top-down effect because it demonstrates that the target template is the yardstick against which all elements in the display are measured. Given different target templates, identical objects will have very different effects on performance (Experiments 1A vs. 1B). In terms of our modeling, this means that elements will be processed to different levels: the ones more similar to the template will require longer processing times, and in turn, they will produce steeper logarithmic functions. Second, the presence of task goal modulation on RT is, of course, the very definition of a top-down

effect because it amounts to changing what the observer *does* with the stimuli *without* changing the stimuli.

### Do These Effects Reflect Attentive Processing?

To answer this question, we must attempt to define visual attention. The term attention is generally used to describe processes that are selective in nature and/or driven by top-down goals and expectations. By selective, it is typically understood that processing is unequal and favors some stimuli over others. Note that this definition is agnostic as to whether processing is serial or parallel. For instance with *spatial* selective attention, observers can decide to preferentially process (attend to) one location over others (Ling & Carrasco, 2006). Spatial attention tends to process locations serially (e.g., Dugue, Marque, & VanRullen, 2011; Dugue, McLelland, Lajous, & VanRullen, 2014; VanRullen & Koch, 2003), although there is evidence that a small set of locations may be preferentially processed at the same time (e.g., Jefferies, Enns, & Di Lollo, in press). Whether at one or a few locations, the point is spatial attention will preferentially process information at those locations over information at other locations. Something similar occurs in the literature of feature-based selective attention: attending to a specific feature in the scene has been shown to not only change the processing of objects/locations containing that feature, but can also change the baseline firing rate of neurons that code for that feature in empty locations (Serences & Boynton, 2007; Lustig & Beck, 2012) presumably in anticipation of those features appearing at those locations. Thus, attending to downward movement, for instance, allows observers to preferentially process downward moving stimuli over upward moving stimuli. This concept of selectivity is exactly what Wolfe proposed as the mechanism behind *guidance* in Guided Search (Wolfe, 1994), and Bundesen (1990) in terms of *pertinence values*. We believe the effects of lure-target similarity demonstrated here reflect selective processing: less similar lures are clearly processed to a lesser extent than the more similar lures. In our modeling, this is reflected by more similar lures requiring higher decision thresholds than less similar lures. In other words, similarity relations impose unequal processing of the stimuli, with more processing required for more similar lures. Stage-one thresholds are determined in graded fashion, reflecting lure-target similarity, and differ from previous proposals regarding early attention filters (Neisser & Becklen, 1975; Palmer et al., 1993; Treisman, 1969; Treisman & Geffen, 1967).

As mentioned above, the concept of attention is also associated with the concepts of top-down goals and expectations (also referred to as *current goals* in Awh, Belopolsky & Theeuwes, 2012, taxonomy of attentional control). In the current study, we found evidence for both types of these effects on stage-one processing. There can be different kinds of expectations. To us, one such expectation is that of finding a specific target in the scene. Having a target template in mind is a form of expectation regarding what may or may not be present in the display. Thus, demonstrating that processing of identical stimuli is changed when a target template in the observer's mind is changed is evidence of expectations affecting processing. This is what we find when comparing Experiments 1A and 1B: the lures that underwent the least processing in 1A (blue circles) were the ones that required the most processing in 1B. This was a direct consequence of having changed the target template. Regarding top-down goals, we compared how identical

displays would be processed under two different top-down goals: the goal to find a specific target (Experiment 1A) and the goal to find a unique item (Experiment 4). The results showed clear evidence that altering such top-down goals can fundamentally alter the manner in which observers process information, even when target stimulus on each display differs substantially from its surrounding elements (i.e., search is efficient).

Using this set of properties of what it means to have "attentive processing," our results are consistent with stage-one processing being attentive. In Awh et al.'s (2012) terminology, our results are evidence that voluntary attentional control (i.e., the current goals of the observer) dramatically impact the unfolding of stage-one processing. In the next section, we will review other studies that observed very early attentional modulations of processing consistent with stage-one being influenced by attention. What is novel in our findings is the discovery that, when people are looking for a specific target in a scene, stage-one processing is "exhaustive" in nature (Townsend & Ashby, 1983): in other words, *we see (and we attend to) everything in a scene*. This is not to say that we process everything to the same extent as the degree of processing of any individual element is directly related to its match to the observers' goals. But to a certain extent, everything in the scene is attentively processed (see also, Tsal & Makovski, 2006 for evidence that we attend all possible target and distractor locations, even before the onset of the display). This form of parallel attentive processing might be the fabric for the rich sense of awareness that we have of the world as being constituted by many, perhaps all, the elements in a scene, and not just the one or few items that are the focus of stage-two attentive processing.

### Converging Evidence That Early Perceptual Processing Is Attentive

There is growing evidence that attention may affect all items in a display simultaneously. Indeed, in cognitive neuroscience, there is evidence that certain types of attentional selection (like feature-based prioritization) can happen in parallel across the visual scene (e.g., Bichot, Rossi, & Desimone, 2005; Lustig & Beck, 2012; Serences & Boynton, 2007). Such forms of parallel attention ought to produce differences in processing across the display, thus impacting processing prior to the scrutiny of individual items. From neuro-functional studies, there is also growing evidence that top-down attention modulates activity in even the earliest brain structures involved in processing visual information (e.g., Fischer & Whitney, 2012; Kastner et al., 2004). These results cast doubt on the notion of "pre-attentive" processes in vision (as proposed by many cognitive theorists like Rensink, 2000; Treisman & Gelade, 1980; Wolfe, 1994, among others) and support the idea of extremely early and parallel attentive filtering or boosting of visual information according to observers' goals. Others have similarly challenged the use of the concept "pre-attentive" in vision research (e.g., Di Lollo, Kawahara, Zuvic & Visser, 2001; Lleras, Buetti, & Mordkoff, 2013; Quinlan, 2003). Further, there is existing evidence from "pop-out" search that differences in intercepts can uniquely track selection difficulty, even when search slopes fall clearly in the efficient range (e.g., Wan & Lleras, 2010).

There is also recent evidence from the flanker effect literature that flankers are selected (or rejected) extremely early in processing: the identity of a letter during the first 17 ms after its onset



determines the congruency effect it will have in a flanker task (Max & Tsal, 2015). Also, when target-like stimuli are used as the unexpected stimuli in an Inattentional Blindness task, Inattentional Blindness is abolished even when those stimuli are smaller and less luminant than other unexpected stimuli to which observers are inattentionally blind (Buetti, Lleras & Moore, 2014). This is evidence that the top-down setting to find a specific target in a scene modulates the processing of unexpected stimuli wherever and whenever they might appear, if they match the target template.

Electrophysiological evidence from monkeys shows that, even in efficient search, target-distractor similarity modulates single-cell responses in the Frontal Eye Fields (FEF; e.g., Purcell et al., 2010), therefore modulating the onset time of eye movements to the target. Such results are difficult to square with prevailing views because manipulations of target-distractor similarity are thought to be uniquely linked to changes in search slopes (e.g., Duncan & Humphreys, 1989), and not to the putatively preattentive stage-one processing. In addition, there is now substantial evidence that the top-down goals of an awake behaving monkey alter the receptive field properties of neurons as early as V1 (Ito & Gilbert, 1999; Li et al., 2004, 2006; Roelfsema, Lamme, & Sepekreijse, 1998), even *before* the presentation of stimuli. Finally, and more directly supportive of our proposal, Bichot, Rossi and Desimone (2005) found evidence in Macaques that target-related features are represented in parallel across the visual field, thereby biasing the processing of visual information to find locations likely to contain the target. This is consistent with our screening function. Bichot and colleagues also found evidence of specific item-by-item biasing, consistent with a specific comparison of a candidate to a target template, which is consistent with stage-two scrutiny. Together, these findings all suggest there is an attentive stage of parallel processing in vision that likely precedes the scrutiny of specific candidate items.

Our findings that a subset of distractors (the lures) can be excluded from attentive scrutiny is also very much in line with recent findings in cognitive neuroscience showing that distractor exclusion may be a very early process in the visual stream both anatomically (e.g., in the pulvinar, Fischer & Whitney, 2012) and temporally (in the P1 wave of the ERP, Moher, Lakshmanan, Egeth & Ewen, 2014). Hopf, Boelmans, Schoenfeld, Luck and Heinze (2004) also showed ERP differences as early as 140ms, prior to the onset of the N2pc (in the 200-ms range). They interpreted their results as evidence of two stages of attentive processing, one early and parallel (because it was not spatially specific) that was responsible for the early difference at 140 ms, and a second component, that was more directly associated with the scrutiny of the target (responsible for later differences in the N2pc). Eimer and Grubert (2014) recently published evidence supporting the idea that attention can work in parallel fashion and process multiple items independently and simultaneously, even when one of them is a nontarget, as long as it is similar to the target. This is consistent with our proposal for parallel stage-one processing that is sensitive to lure-target similarity.

We propose that visual and attentive processing unfolds mostly over two stages. The first is characterized by a parallel, unlimited capacity architecture with limited resolution that can, when looking for a specific target, process all items independently, screening likely targets from unlikely targets (see Appendix A for more specifics). The output of this function is a map with likely loca-

tions that require more focused attentive processing.<sup>3</sup> The second stage is capacity limited but has functionally unlimited resolution insofar as it can tell a target from a nontarget element. Generally, these two stages are separate and likely sequential, but that does not mean they are always separate or do not interact with each other. As we mentioned earlier, observers may want to impose top-down scanning strategies that might negate the impact of screening on RTs. Also, while we tried to minimize “candidate crowding” in our displays (i.e., the presence of clumps of candidates in a particular region of the display), we assume this form of crowding would likely be best resolved (when detected) by triggering eye movements to the crowded location. We do not know yet how nor when the need for such saccades will interplay with screening. Would it halt the otherwise exhaustive processing of a display? More studies will be needed to address such questions. Finally, it is worth noting that many of the specifics of our theory can be found across different theories, most notably in theories based on signal detection theory (Eckstein et al., 2000; Palmer et al., 1993; Verghese, 2001) and on eye-movement based theories of search (such as those of Najemnik & Geisler, 2005; Zelinsky, 2008 and Rosenholtz et al., 2012).

Eye-movement based theories of search propose that RTs’ linear dependence on set size reflects not the need to scrutinize candidates necessarily, but rather a need to produce more eye movements when more candidates are present (Najemnik & Geisler, 2005; Rosenholtz et al., 2012; Zelinsky, 2008). This follows because the representation of peripheral items tends to be lossy and very sensitive to crowding (Rosenholtz, 2011; Rosenholtz et al., 2012). Increasing the number of candidates aggravates crowding, which is best resolved by moving fixation closer to those elements. However, we do not think visual crowding alone can explain all search behavior. Our preliminary investigations (not reported here) indicate that “lure-crowding” does not qualitatively change the pattern of results we observed in Experiment 1. Future investigations will need to systematically study the effects of lure- and candidate-crowding on screening functions. Finally, it should be noted that the mechanisms proposed by some of these theories tend to have a difficult time dealing with displays containing simple stimuli, like the ones we used here (e.g., Najemnik & Geisler, 2005; Zelinsky, 2008). This comes from the fact that these models are in a way *too good*: when analyzing scenes made up of clearly segregated objects on uniform backgrounds (like the stimuli we used here), the models often find the target at once and with little variability. Thus, investigators often have to artificially introduce background noise to these displays to reduce the target signal (e.g., Itti & Koch, 2000; Michel & Geisler, 2011; Zelinsky, 2008). We take this as an indication that perhaps those models may be working on different factors that affect search behavior (i.e., situations of low signal-to-noise ratios) than the ones we have focused

<sup>3</sup> Note that, in our architecture, screening lures out does not necessitate an inhibitory component to “mark” the rejected lures. That is, once a lure is identified as an unlikely target, we see no need for that location to be actively inhibited. Instead, the mere act of not passing (i.e., not actively representing) the coordinates of this location to the next stage of processing is sufficient for scrutiny to never be directed to this location. That being said, whether stage-one screening uses a form of inhibitory representation (for instance, across eye movements) is something that requires further investigation.



on here (i.e., presumably a higher signal to noise ratio than the ones used in those studies). It is likely that a comprehensive understanding of visual search behavior in the real world requires an integration of these two approaches.

### Implications for the Attention Literature

There are several implications of our findings for current research and theory on attention. First, our results show a richer understanding of visual processing may come about by looking at a finer analysis of target-distractor similarity than the current feature-discontinuity framework (see Wolfe & Horowitz, 2004 for a review). Indeed, stimuli are often created using large feature discontinuities (e.g., red vs. blue) that may obscure the graded processing that occurs when stimuli vary along that specific similarity axis (as blue colors become closer in hue to red). Examples of this form of stimulus selection can be found in classic studies of efficient search, including Bravo and Nakayama (1992); Treisman and Gelade (1980) and Wolfe (1994). Signal detection models of visual search, such as Verghese (2001), have similarly emphasized how the range of target-distractor discriminability affects search performance, though these models typically focus on predicting search accuracy, not RTs. Similarity is not limited to variation along one feature dimension, but instead varies continuously along multiple visual features. Thus, as revealed by Experiments 3B and 3C, two visual objects can differ from each other in multiple ways, yet be equally similar to a third stimulus: thin orange crosses were different in line weight and color from thick red crosses, yet they both produce the same screening cost when the target was a thin red T. Similarity effects in search have also been well documented in several experiments on triple conjunctions. In these experiments, it is found that search efficiency improves when a target differs from distractors along two dimensions (of three possible ones), and worsens (becomes inefficient) when the target and distractors share two of the three possible features (Nordfang & Wolfe, 2014; Williams & Reingold, 2001; Wolfe, Cave, & Franzel, 1989). This is easily accounted for in terms of overall similarity: sharing two out of three possible feature values is more likely to produce distractors that are actually candidates in terms of their similarity to the target, whereas sharing only one of three possible feature values is more likely to produce distractors that are lures. This account of triple conjunction in terms of overall similarity may be more parsimonious than accounts based on guidance by one or two specific feature channels (Nordfang & Wolfe, 2014; Wolfe, 1994).

Second, we contend that our view of visual processing and attention is ecologically valid (perhaps more so than current theories). Given the visual heterogeneity of objects around us, it is likely the case that *most items* in any given real-world scene will be vastly different from the object that we are looking for (something that has been suggested before, Neider & Zelinsky, 2008; Wolfe, Alvarez, et al., 2011). Our results show that sufficiently dissimilar items require very little processing before a decision can be made to exclude them from further processing. Thus, having a two-stage architecture (first rough, but quick decisions on lures followed by slow, but precise judgments about candidates) is computationally efficient and advantageous. The screening process will quickly reduce the location uncertainty about where the target in the scene might be, vastly reducing the number of to-be-

scrutinized locations. This process occurs before detailed representations of the candidate items are available for scrutiny. If we are correct, this means that visual salience may play a smaller role in visual search than previously thought (see also Clarke, Dziemianko & Keller, 2014; Najemnik & Geisler, 2005).

Returning to the proposal by Neider and Zelinsky (2008) that in real-world scenes there exists a *functional set size* (i.e., what we call the candidate set), our results show that computing the functional set size takes time and the time it takes depends on the specifics of each visual scene: how many lures there are and how similar each one is to the target template. Therefore, an important source of variance in real world search tasks must be determined by these visual factors, a conclusion that up to now has not been fully acknowledged in the literature. Current proposals have instead stressed the importance of knowledge and context in determining the functional set size (Wolfe, Vo, et al., 2011b). We agree that these factors likely play a role. However, in addition to such nonvisual sources of variance, our results show that it is important to study and better understand the contribution of visual factors, as evidenced by the results of Experiments 1 and 3.

A proof of concept is presented in Figure 10 where we show what expected response times would be if we had created scenes with large set sizes that differed in terms of the overall target-lure similarity from very low, to medium to high, and for three different levels of functional set size (1, 2, and 4). Simply taking into account variations in these two factors produces very large variability at any given set size, without invoking additional contributions of knowledge and context. It is likely that in real-world scenes, functional set sizes and lure heterogeneity probably vary more than in this demonstration, but the overall message would remain the same: variations in target-lure similarity across scenes (i.e., how much like the target do the background objects look) as well as variations in functional set size (i.e., how many likely candidates are in a scene) can produce substantial variations in RT, similar to those observed in Wolfe, Alvarez, et al. (2011).

Third, it is also interesting to note that the type of parallel processing we propose for stage one can help attention deal with crowded scenes and visual crowding under certain circumstances. Imagine a scene where there is a region of very crowded stimuli (a group of letters, as in a road sign), but the target you are looking for is very dissimilar from the items composing the crowded set (say the target is a car on the highway). We know from the literature in crowding that we can extract general visual features about the set of crowded items (i.e., we may know that the crowded items are letters or have letter-like features), without having access to the specific bindings (which letters, see Pelli, 2008). It is very likely that this general information about the set of items is sufficient for stage-one processing to reach a decision that a target is unlikely to be at the location of the crowded items (in this example, that a car is unlikely to be at the location of the letters), without the need for scrutinizing that location. In other words, crowding in the world might not always produce large effects on visual search (more specifically, crowded locations might not require stage-two processing), if the information at the crowded location is sufficiently dissimilar to the target template (see also, Rosenholtz et al., 2012).

Fourth, our results help to put into perspective the role of visual STM on visual search. As mentioned before, we suggest that visual search can be separated into two attentive stages that rely most

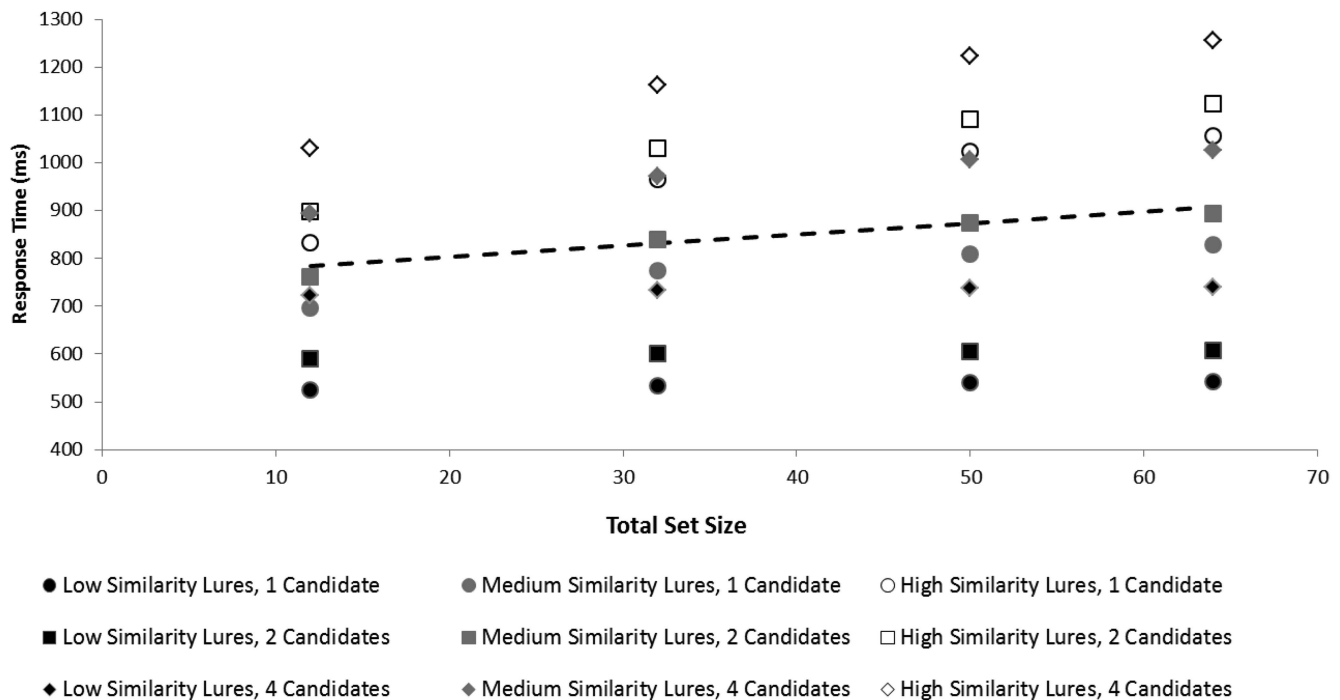


Figure 10. Simulated data showing expected response time as a function of total set size with three different types of lures (low, medium and high similarity) and three levels of functional set size (number of candidates equals 1, 2 or 4). The values for the logarithmic screening slopes were taken from observed values in Experiments 1 and 3 (10, 79 and 134 ms per log unit). The scrutiny cost was set at 66 ms/item (Experiment 3A), in additive fashion to the screening cost. The resulting data shows a cloud of data that is poorly explained by a linear function ( $R$ -square = 0.05) and shows a very flat linear slope (2 ms/item), just as the data in Figure 2. Note that a single logarithmic function would also fit the data very poorly (also 5% of the variance). However, each data point represents the separate contribution of both attentive screening (a logarithmic cost as a function of the number of lures of a given similarity) and attentive scrutiny (a linear cost as a function of the number of candidates). Thus, a two-stage model reflecting variation in visual factors alone can explain 100% of this variance, despite the very large variability observed in the data. Indeed, RTs can vary by as much as 700 ms at a given set size, when both screening and scrutiny contributions are taken into consideration.

critically on different types of information. The first stage, Screening, is the “visually-dependent” stage, and by that we mean the stage where visual similarity between objects and a target template drive the decision making (does this location require further processing?). The second stage, Scrutiny, is a “spatially-dependent” stage, meaning that this stage relies on location maps to drive its processing (deploying detailed processing at the coordinates of the candidates that survived screening). It has been shown that while spatial working memory loads slow down scrutiny in visual search (Oh & Kim, 2004; Woodman & Luck, 2004), feature working memory load does not (Oh & Kim, 2004; Woodman, Vogel & Luck, 2001). It was initially surprising to find that feature working memory loads did not impact search efficiency. However, given the present results, if feature working memory load impacts stage-one processing, then it would do so by modulating the length of time to complete screening, yet it would not alter the number of locations to be scrutinized. As such, it would not have an impact on the slope of the linear search function computed to index the scrutiny stage. In contrast, loading spatial working memory with location information will surely interfere with attentive scrutiny as this type of processing relies on maintaining location information

regarding likely target locations. Thus, loading spatial working memory should interfere, more specifically increase, the slope of the search function.

Based on our findings, we can speculate that there is a powerful, parallel analysis of a scene, even before precise scrutiny of elements in that scene begins. Our rich sense of visual awareness might actually be related to that first stage of parallel processing of a scene, whereas evidence for failures to report aspects of that experience (as documented in the Inattentional Blindness literature, e.g.) might be related to the capacity-limited nature of stage-two processing, in particular to limitations in our ability to store the objects or features that we want to remember (Chen & Wyble, 2015). Here, we have found behavioral evidence that in fact, when we look for something in the world, we attentively process every single item in front of our eyes (though not all to the same extent). Thus, contrary to many current beliefs, we propose our results suggest that we see everything in front of our eyes, in fact, that *we attend to everything* in front of our eyes, even if our memory of having done so might be quite frail and unreliable (Chen & Wyble, 2015; Mack & Rock, 1998; Moore & Egeth, 1997).

## Limitations

There are many limitations to the current study. First of all, it is important to remember the pattern of additivity we observed between the two stages is more reflective of our experimental design than it might be of the manner in which the stages work in general. There are reasons why one might expect stage one would not influence search times as much as they do here. For instance, if subjects decide to scan a scene with their eyes in systematic fashion, the effects of stage-one processing times may not be observed. Thus, further research will be needed to map out the boundary conditions of our findings. In fact, a second limitation stems from the fact that we did not monitor eye movements in our experiments, thus we do not know what role they play in our results.

In our displays we tried to minimize what can be referred to as “candidate crowding”: the likelihood that two (or more) candidates will be placed very near each other and thereby crowd each other. However, the grid-like arrangement of stimuli in our displays made it so that the likelihood of crowding of the target by lure stimuli was high in large set size conditions. Indeed, most grid-like arrangements of stimuli violate Bouma’s law (Bouma, 1970; Pelli & Tillman, 2008) which states spacing ought to be proportional to eccentricity (by at least a factor of 0.5) to avoid crowding. Crowding-free arrangements are rarely the case in evenly spaced grid displays, and it was certainly not the case in our displays. But, was lure crowding responsible for the logarithmic effects obtained? Likely not. At low set sizes, the likelihood that a lure would fall on any of the locations adjacent to the target was quite low (17% at set size of 2), whereas at large set sizes (like 32), that probability was close to 100%. As a result, one can see that lure crowding was fairly rare at low set sizes, and quite prevalent at large set sizes. Yet, looking at our data, lures had the largest effects at low set sizes, and the smallest effects at large set sizes. Thus, we can conclude that lure crowding was likely not playing a major role in our experimental results, and certainly not responsible for the logarithmic nature of the observed search functions. However, we cannot say the same thing about “candidate crowding” (the presence of a like-target stimuli near the target itself). There has been substantial research on crowding that shows that when alike stimuli are presented within Bouma’s window, crowding has very deleterious effects on visibility and search performance and often forces eye movements (Vlaskamp, Over & Hooge, 2005). Thus, to minimize candidate crowding in Experiments 3A through 3C, we attempted to distribute the candidates as evenly as possible throughout the display. That said, we did not completely eliminate candidate crowding from our displays. Thus, the question of how candidate crowding might have impacted our results remains open. It will be also important to systematically study how increased levels of candidate crowding (beyond the ones used in our experiments) impact the screening function.

In our paper, we describe the second stage of processing (scrutiny) as being random (unbiased). In other words, our proposal states that the actual similarity of specific candidates to the target or their priority score on a priority/saliency map type of representation does not impact the order in which candidates are scrutinized. Instead, we propose that the output of stage one is not a priority/saliency score, but rather a list of randomly ordered addresses representing the locations of candidates in the scene, with

perhaps a bias reflecting distance from fixation. We describe stage two as operating in this fashion simply because it appears the most parsimonious possibility. The present data are moot on this point because we only used one type of candidate. Future research should look at different levels of candidate heterogeneity to determine whether there is an ordered preference by similarity (e.g., as in Guided Search; Wolfe, 1994) or not, at the time of stage-two scrutiny.

Another limitation of our paper is the reliance of grouped data to infer within-subject mechanisms (see Borsboom et al., 2009; Cronbach, 1957; Hull, 1945). We will not say much on this point, except to say that given the stochastic nature of our models, to obtain stable observations for each experimental condition one requires many observations and averaging across a multitude of unrelated variations (like target location). Having 50 observations per condition and 20 subjects per experiment produced close to 1000 observations per condition (once errors were discarded). The observed grouped data did produce stable RT patterns. To obtain similar levels of stability at the subject level would require a more psychophysical approach, with multiple-session studies and expert observers. This is something we have not attempted yet, but something worth pursuing.

We believe that, in the context of our displays, our data are supportive of a two-stage architecture underlying visual search performance, something that many previous investigators have proposed but few have directly tested. One direct test of the architecture of visual processing during visual search was done by Fific, Townsend and Eidels (2008), the only one to our knowledge. The authors used System Factorial Technology to uncover whether visual processing during visual search is serial or parallel in nature. Their analysis suggested a single parallel stage architecture (as have done several signal detection theory inspired models of search, such as, Eckstein et al., 2000; Palmer, Ames, & Lindsey, 1993; Verghese, 2001), with positively interacting channels. Though compelling, this evidence is somewhat restricted by the fact that the displays only contained two items and they both were presented at the fovea. Thus, we believe the architecture proposed by these authors may account for foveal processing, whereas our architecture may be better suited for understanding processing of larger visual scenes and peripherally presented stimuli. Note that Wolfe (2005, Guided Search 4.0) more recently proposed a parallel-serial hybrid architecture for visual search, where a certain number of items could be processed in parallel, and a new item could start its processing when one item completes.

In terms of other future directions, it will be important to extend the current findings to real life objects and scene stimuli. This will allow a more direct comparison to the search in real-scenes literature can be made (e.g., Henderson, Brockmole, Castelano, & Mack, 2007; Henderson, Weeks, & Hollingworth, 1999; Torralba, Oliva, Castelano, & Henderson, 2006; Malcolm & Henderson, 2009). Second, all the stimuli in our displays had relatively high signal-to-noise ratios (SNRs). That is, it is not visually difficult to discriminate the stimuli from the background in the visual field. In contrast, there are studies of visual search that use low signal-to-noise stimuli (Henderson, Malcolm, & Schandl, 2009; Neider & Zelinsky, 2006), or that introduce artificial noise to simple stimuli (Itti & Koch, 2000; Michel & Geisler, 2011; Zelinsky, 2008). It is possible, then, that our current proposal may be complementary to the findings from studies with low SNRs displays. That is, how we



propose visual processing unfolds may be representative of scenes and stimuli with high SNRs, with a shift toward more eye-movement driven search as SNRs decrease in scenes.

## Conclusions

We documented for the first time the existence of a new form of variability (a logarithmic contribution) in search times resulting from manipulations that affect stage-one processing that can be experimentally differentiated from stage-two variability (a linear contribution). We propose that stage-one variability is consistent with a parallel, unlimited capacity and resolution limited architecture. When observers are asked to find a specific object in a scene, this architecture is used to compare *all* locations in the world in parallel to a target template, allowing fast rejection of locations that are unlikely to contain the target. We showed that the processing is sensitive to the task goals and to the similarity of items to the target template.

## References

- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, 8, 457–464. <http://dx.doi.org/10.1016/j.tics.2004.08.011>
- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, 15, 106–111. <http://dx.doi.org/10.1111/j.0963-7214.2004.01502006.x>
- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, 16, 437–443. <http://dx.doi.org/10.1016/j.tics.2012.06.010>
- Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision*, 9, 1–18.
- Baldauf, D., & Deubel, H. (2009). Attentional selection of multiple goal positions before rapid hand movement sequences: An event-related potential study. *Journal of Cognitive Neuroscience*, 21, 18–29. <http://dx.doi.org/10.1162/jocn.2008.21021>
- Becker, S. I. (2008). The mechanism of priming: Episodic retrieval or priming of pop-out? *Acta Psychologica*, 127, 324–339. <http://dx.doi.org/10.1016/j.actpsy.2007.07.005>
- Becker, S. I. (2010). The role of target-distractor relationships in guiding attention and the eyes in visual search. *Journal of Experimental Psychology: General*, 139, 247–265. <http://dx.doi.org/10.1037/a0018808>
- Becker, S. I., Harris, A. M., Venini, D., & Retell, J. D. (2014). Visual search for color and shape: When is the gaze guided by feature relationships, when by feature values? *Journal of Experimental Psychology: Human Perception and Performance*, 40, 264–291. <http://dx.doi.org/10.1037/a0033489>
- Becker, S. I., Valuch, C., & Ansorge, U. (2014). Color priming in pop-out search depends on the relative color of the target. *Frontiers in Psychology*, 5, 289.
- Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science*, 308, 529–534. <http://dx.doi.org/10.1126/science.1109676>
- Blackmore, S. J., Brelstaff, G., Nelson, K., & Troscianko, T. (1995). Is the richness of our visual world an illusion? Transsaccadic memory for complex scenes. *Perception*, 24, 1075–1081. <http://dx.doi.org/10.1068/p241075>
- Borsboom, D., Kievit, R. A., Cervone, D., & Hood, S. B. (2009). The two disciplines of scientific psychology, or: The disunity of psychology as a working hypothesis. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudhary (Eds.), *Dynamic process methodology in the social and developmental sciences*. Heidelberg, Germany: Springer. [http://dx.doi.org/10.1007/978-0-387-95922-1\\_4](http://dx.doi.org/10.1007/978-0-387-95922-1_4)
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226, 177–178. <http://dx.doi.org/10.1038/226177a0>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436. <http://dx.doi.org/10.1163/156856897X00357>
- Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics*, 51, 465–472. <http://dx.doi.org/10.3758/BF03211642>
- Briggs, G. E. (1974). On the predictor variable for choice reaction time. *Memory & Cognition*, 2, 575–580.
- Briggs, G. E., & Swanson, J. M. (1970). Encoding, decoding, and central functions in human information processing. *Journal of Experimental Psychology*, 86, 296–308. <http://dx.doi.org/10.1037/h0029979>
- Buetti, S., Lleras, A., & Moore, C. M. (2014). The flanker effect does not reflect the processing of “task-irrelevant” stimuli: Evidence from inattentive blindness. *Psychonomic Bulletin & Review*, 21, 1231–1237. <http://dx.doi.org/10.3758/s13423-014-0602-9>
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97, 523–547. <http://dx.doi.org/10.1037/0033-295X.97.4.523>
- Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 107, 16048–16053. <http://dx.doi.org/10.1073/pnas.1004801107>
- Caddigan, E., & Lleras, A. (2010). Saccadic repulsion in pop-out search: How a target’s dodgy history can push the eyes away from it. *Journal of Vision*, 10, 9. <http://dx.doi.org/10.1167/10.14.9>
- Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, 57, 1241–1261. <http://dx.doi.org/10.3758/BF03208380>
- Carrasco, M., McLean, T. L., Katz, S. M., & Frieder, K. S. (1998). Feature asymmetries in visual search: Effects of display duration, target eccentricity, orientation and spatial frequency. *Vision Research*, 38, 347–374. [http://dx.doi.org/10.1016/S0042-6989\(97\)00152-1](http://dx.doi.org/10.1016/S0042-6989(97)00152-1)
- Carrasco, M., & Yeshurun, Y. (1998). The contribution of covert attention to the set-size and eccentricity effects in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 673–692. <http://dx.doi.org/10.1037/0096-1523.24.2.673>
- Chen, H., & Wyble, B. (2015). Amnesia for object attributes: Failure to report attended information that had just reached conscious awareness. *Psychological Science*, 26, 203–210. <http://dx.doi.org/10.1177/0956797614560648>
- Chong, S. C., & Treisman, A. (2005a). Statistical processing: Computing the average size in perceptual groups. *Vision Research*, 45, 891–900. <http://dx.doi.org/10.1016/j.visres.2004.10.004>
- Chong, S. C., & Treisman, A. (2005b). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, 67, 1–13. <http://dx.doi.org/10.3758/BF03195009>
- Chun, M. M., & Cavanagh, P. (1997). Seeing two as one: Linking apparent motion and repetition blindness. *Psychological Science*, 8, 74–79. <http://dx.doi.org/10.1111/j.1467-9280.1997.tb00686.x>
- Clarke, A., Dziemianko, M., & Keller, F. (2014, May). Measuring the salience of an object in a scene. Poster presented at the meeting of the Vision Sciences Society, St. Pete Beach, FL.
- Corballis, M. C., Kirby, J., & Miller, A. (1972). Access to elements of a memorized list. *Journal of Experimental Psychology*, 94, 185–190. <http://dx.doi.org/10.1037/h0032780>
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671–684. <http://dx.doi.org/10.1037/h0043943>
- Dennett, D. C. (1991). *Consciousness explained*. Boston, MA: Little, Brown.
- Dennett, D. C. (1992). “Filling in” versus finding out: A ubiquitous



- confusion in cognitive science. In H. L. Pick, Jr., P. van den Broek, & D. C. Knill (Eds.), *Cognition: Conceptual and methodological issues* (pp. 33–49). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10564-002>
- Dennett, D. C. (1998). No bridge over the stream of consciousness. *Behavioral and Brain Sciences*, 21, 753–754. <http://dx.doi.org/10.1017/S0140525X98281751>
- Di Lollo, V., Kawahara, J., Zuvic, S. M., & Visser, T. A. W. (2001). The preattentive emperor has no clothes: A dynamic redressing. *Journal of Experimental Psychology: General*, 130, 479–492. <http://dx.doi.org/10.1037/0096-3445.130.3.479>
- Doherty, L. M., & Foster, D. H. (1999). Limitations of rapid parallel processing in the detection of long and short oriented line targets. *Spatial Vision*, 12, 485–497. <http://dx.doi.org/10.1163/156856899X00274>
- Donders, F. C. (1968/1969). On the speed of mental processing. In W. G. Koster (Trans. & Ed.), *Attention and Performance II: Acta Psychologica*, 30, 412–431.
- Dugué, L., Marque, P., & VanRullen, R. (2011). Transcranial magnetic stimulation reveals attentional feedback to area V1 during serial visual search. *PLoS ONE*, 6, e19712. <http://dx.doi.org/10.1371/journal.pone.0019712>
- Dugue, L., McLelland, D., Lajous, M., & VanRullen, R. (2014). Spatial deployment of attention in visual search: New evidence against a strict parallel model. Proceedings of the 14th Vision Sciences Society annual meeting, St. Pete Beach, Florida.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458. <http://dx.doi.org/10.1037/0033-295X.96.3.433>
- Duncan, J., & Humphreys, G. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 578–588. <http://dx.doi.org/10.1037/0096-1523.18.2.578>
- Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics*, 62, 425–451. <http://dx.doi.org/10.3758/BF03212096>
- Eimer, M., & Grubert, A. (2014). Spatial attention can be allocated rapidly and in parallel to new visual objects. *Current Biology*, 24, 193–198. <http://dx.doi.org/10.1016/j.cub.2013.12.001>
- Eimer, M., Kiss, M., & Cheung, T. (2010). Priming of pop-out modulates attentional target selection in visual search: Behavioural and electrophysiological evidence. *Vision Research*, 50, 1353–1361. <http://dx.doi.org/10.1016/j.visres.2009.11.001>
- Fifić, M., Townsend, J. T., & Eidels, A. (2008). Studying visual search using systems factorial methodology with target-distractor similarity as the factor. *Perception & Psychophysics*, 70, 583–603. <http://dx.doi.org/10.3758/PP.70.4.583>
- Fischer, J., & Whitney, D. (2012). Attention gates visual coding in the human pulvinar. *Nature Communications*, 3, 1051. <http://dx.doi.org/10.1038/ncomms2054>
- Fisher, D. L. (1982). Limited-channel models of automatic detection: Capacity and scanning in visual search. *Psychological Review*, 89, 662–692. <http://dx.doi.org/10.1037/0033-295X.89.6.662>
- Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 718–734. <http://dx.doi.org/10.1037/a0013899>
- Haslam, N., Porter, M., & Rothschild, L. (2001). Visual search: Efficiency continuum or distinct processes? *Psychonomic Bulletin & Review*, 8, 742–746. <http://dx.doi.org/10.3758/BF03196212>
- Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. *Eye movements: A window on mind and brain*, 537–562.
- Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review*, 16, 850–856. <http://dx.doi.org/10.3758/PBR.16.5.850>
- Henderson, J. M., Weeks, P. A., Jr., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 210–228. <http://dx.doi.org/10.1037/0096-1523.25.1.210>
- Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 598–615. <http://dx.doi.org/10.1037/0278-7393.10.4.598>
- Hoffman, J. (1979). A two-stage model of visual search. *Perception and Psychophysics*, 25, 319–327.
- Hopf, J. M., Boelmans, K., Schoenfeld, M. A., Luck, S. J., & Heinze, H. J. (2004). Attention to features precedes attention to locations in visual search: Evidence from electromagnetic brain responses in humans. *The Journal of Neuroscience*, 24, 1822–1832. <http://dx.doi.org/10.1523/JNEUROSCI.3564-03.2004>
- Hull, C. L. (1945). The place of innate individual and species differences in a natural-science theory of behavior. *Psychological Review*, 52, 55–60. <http://dx.doi.org/10.1037/h0056383>
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, 43, 171–216. <http://dx.doi.org/10.1006/cogp.2001.0755>
- Ito, M., & Gilbert, C. D. (1999). Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron*, 22, 593–604. [http://dx.doi.org/10.1016/S0896-6273\(00\)80713-8](http://dx.doi.org/10.1016/S0896-6273(00)80713-8)
- Itti, L., & Koch, C. (1999). Comparison of feature combination strategies for saliency-based visual attention systems. In *Proceedings SPIE Human Vision and Electronic Imaging IV (HVEI'99)*, San Jose, CA (Vol. 3644, pp. 473–82). Bellingham, WA: SPIE Press.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40, 1489–1506. [http://dx.doi.org/10.1016/S0042-6989\(99\)00163-7](http://dx.doi.org/10.1016/S0042-6989(99)00163-7)
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2, 194–203. <http://dx.doi.org/10.1038/35058500>
- Jefferies, L. N., Enns, J. T., & Di Lollo, V. (in press). The flexible focus: Whether spatial attention is unitary or divided depends on observer goals. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 465–470.
- Kastner, S., O'Connor, D. H., Fukui, M. M., Fehd, H. M., Herwig, U., & Pinsk, M. A. (2004). Functional imaging of the human lateral geniculate nucleus and pulvinar. *Journal of Neurophysiology*, 91, 438–448. <http://dx.doi.org/10.1152/jn.00553.2003>
- Kristjánsson, A., Vuilleumier, P., Schwartz, S., Macaluso, E., & Driver, J. (2007). Neural basis for priming of pop-out during visual search revealed with fMRI. <http://dx.doi.org/10.1093/cercor/bhl072>
- Kristofferson, M. W. (1972). When item recognition and visual search functions are similar. *Perception & Psychophysics*, 12, 379–384. <http://dx.doi.org/10.3758/BF03207225>
- Lamme, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *The Journal of Neuroscience*, 15, 1605–1615.
- Lamme, V. A. F., Supér, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, 8, 529–535. [http://dx.doi.org/10.1016/S0959-4388\(98\)80042-1](http://dx.doi.org/10.1016/S0959-4388(98)80042-1)
- Lamy, D., Carmel, T., Egeth, H. E., & Leber, A. B. (2006). Effects of search mode and intertrial priming on singleton search. *Perception & Psychophysics*, 68, 919–932. <http://dx.doi.org/10.3758/BF03193355>
- Lamy, D. F., & Kristjánsson, A. (2013). Is goal-directed attentional guidance just intertribal priming? A review. *Journal of Vision*, 13, 1–19.

- Lamy, D., Yashar, A., & Ruderman, L. (2013). Orientation search is mediated by distractor suppression: Evidence from priming of pop-out. *Vision Research*, 81, 29–35. <http://dx.doi.org/10.1016/j.visres.2013.01.008>
- Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, 43, 149–164. [http://dx.doi.org/10.1016/S0042-6989\(02\)00402-9](http://dx.doi.org/10.1016/S0042-6989(02)00402-9)
- Landman, R., Spekreijse, H., & Lamme, V. A. F. (2004). Relationship between change detection and post-change activity in visual area V1. *NeuroReport*, 15, 2211–2214. <http://dx.doi.org/10.1097/00001756-200410050-00013>
- Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, 13, 132–138. <http://dx.doi.org/10.3758/BF03193824>
- Li, W., Piëch, V., & Gilbert, C. D. (2004). Perceptual learning and top-down influences in primary visual cortex. *Nature Neuroscience*, 7, 651–657. <http://dx.doi.org/10.1038/nn1255>
- Li, W., Piëch, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron*, 50, 951–962. <http://dx.doi.org/10.1016/j.neuron.2006.04.035>
- Li, Z. (1999). Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network*, 10, 187–212. [http://dx.doi.org/10.1088/0954-898X\\_10\\_2\\_305](http://dx.doi.org/10.1088/0954-898X_10_2_305)
- Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, 6, 9–16. [http://dx.doi.org/10.1016/S1364-6613\(00\)01817-9](http://dx.doi.org/10.1016/S1364-6613(00)01817-9)
- Ling, S., & Carrasco, M. (2006). When sustained attention impairs perception. *Nature Neuroscience*, 9, 1243–1245. <http://dx.doi.org/10.1038/nn1761>
- Lleras, A., & Buetti, S. (2014). Not all “distractor” tags are created equal: Using a search asymmetry to dissociate the inter-trial effects caused by different forms of distractors. *Frontiers in Psychology*, 5, 669. <http://dx.doi.org/10.3389/fpsyg.2014.00669>
- Lleras, A., Buetti, S., & Mordkoff, J. T. (2013). When do the effects of distractors provide a measure of distractibility? *Psychology of Learning and Motivation*, 59, 261–316. <http://dx.doi.org/10.1016/B978-0-12-407187-2.00007-1>
- Lleras, A., & Von Mühlenen, A. (2004). Spatial context and top-down strategies in visual search. *Spatial Vision*, 17, 465–482. <http://dx.doi.org/10.1163/1568568041920113>
- Lustig, A. G., & Beck, D. M. (2012). Task-relevant and task-irrelevant dimensions are modulated independently at a task-irrelevant location. *Journal of Cognitive Neuroscience*, 24, 1884–1895. [http://dx.doi.org/10.1162/jocn\\_a\\_00249](http://dx.doi.org/10.1162/jocn_a_00249)
- MacK, A., & Rock, I. (1998). *Inattention blindness*. Cambridge, MA: MIT Press.
- Madison, A., Buetti, S., & Lleras, A. (2015, May). *Singleton search performance predicts performance on heterogeneous displays: Evidence in support of the information theory of vision*. St. Pete, FL: Vision Sciences Society. <http://dx.doi.org/10.1167/15.12.1359>
- Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, 9, 1–13.
- Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition*, 22, 657–672. <http://dx.doi.org/10.3758/BF03209251>
- Maljkovic, V., & Nakayama, K. (1996). Priming of pop-out: II. The role of position. *Perception & Psychophysics*, 58, 977–991. <http://dx.doi.org/10.3758/BF03206826>
- Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, 43, 679–690. <http://dx.doi.org/10.3758/s13428-010-0049-5>
- Max, R., & Tsai, Y. (2015). Selection in flanker tasks is governed by identities and not by locations. *Journal of Vision*, 15, 1069. <http://dx.doi.org/10.1167/15.12.1069>
- Meeter, M., & Olivers, C. N. L. (2006). Intertrial priming stemming from ambiguity: A new account of priming in visual search. *Visual Cognition*, 13, 202–222. <http://dx.doi.org/10.1080/13506280500277488>
- Michel, M., & Geisler, W. S. (2011). Intrinsic position uncertainty explains detection and localization performance in peripheral vision. *Journal of Vision*, 11, 18. <http://dx.doi.org/10.1167/11.1.18>
- Moher, J., Lakshmanan, B. M., Egeth, H. E., & Ewen, J. B. (2014). Inhibition drives early feature-based attention. *Psychological Science*, 25, 315–324. <http://dx.doi.org/10.1177/0956797613511257>
- Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 339–352. <http://dx.doi.org/10.1037/0096-1523.23.2.339>
- Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, 434, 387–391. <http://dx.doi.org/10.1038/nature03390>
- Neider, M. B., & Zelinsky, G. J. (2006). Searching for camouflaged targets: Effects of target-background similarity on visual search. *Vision Research*, 46, 2217–2235. <http://dx.doi.org/10.1016/j.visres.2006.01.006>
- Neider, M. B., & Zelinsky, G. J. (2008). Exploring set size effects in scenes: Identifying the objects of search. *Visual Cognition*, 16, 1–10. <http://dx.doi.org/10.1080/13506280701381691>
- Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, 7, 480–494. [http://dx.doi.org/10.1016/0010-0285\(75\)90019-5](http://dx.doi.org/10.1016/0010-0285(75)90019-5)
- Nordfang, M., & Wolfe, J. M. (2014). Guided search for triple conjunctions. *Attention, Perception, & Psychophysics*, 76, 1535–1559. <http://dx.doi.org/10.3758/s13414-014-0715-2>
- Oh, S. H., & Kim, M. S. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin & Review*, 11, 275–281. <http://dx.doi.org/10.3758/BF03196570>
- Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *The encyclopedia of neurobiology of attention* (pp. 251–256). San Diego, CA: Elsevier. <http://dx.doi.org/10.1016/B978-012375731-9/50045-8>
- Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research: Visual perception*, 155, 23–36.
- O'Regan, J. K. (1992). Solving the “real” mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 46, 461–488. <http://dx.doi.org/10.1037/h0084327>
- Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science*, 4, 118–123. <http://dx.doi.org/10.1111/1467-8721.ep10772534>
- Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 108–130. <http://dx.doi.org/10.1037/0096-1523.19.1.108>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442. <http://dx.doi.org/10.1163/156856897X00366>
- Pelli, D. G. (2008). Crowding: A cortical constraint on object recognition. *Current Opinion in Neurobiology*, 18, 445–451. <http://dx.doi.org/10.1016/j.conb.2008.09.008>
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, 4, 12. <http://dx.doi.org/10.1167/4.12.12>
- Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 11, 1129–1135. <http://dx.doi.org/10.1038/nn.2187>

- Peters, R. J., Iyer, A., Itti, L., & Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Research*, 45, 2397–2416. <http://dx.doi.org/10.1016/j.visres.2005.03.019>
- Potter, M. C. (1975). Meaning in visual search. *Science*, 187, 965–966. <http://dx.doi.org/10.1126/science.1145183>
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 509–522. <http://dx.doi.org/10.1037/0278-7393.2.5.509>
- Potter, M. C. (2012). Recognition and memory for briefly presented scenes. *Frontiers in Psychology*, 3, 32. <http://dx.doi.org/10.3389/fpsyg.2012.00032>
- 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*, 117, 1113–1143. <http://dx.doi.org/10.1037/a0020311>
- Quinlan, P. T. (2003). Visual feature integration theory: Past, present, and future. *Psychological Bulletin*, 129, 643–673. <http://dx.doi.org/10.1037/0033-2909.129.5.643>
- Rangelov, D., Muller, H. J., & Zehetleitner, M. (2013). Visual search for feature singletons: Multiple mechanisms produce sequence effects in visual search. *Journal of Vision*, 13(3), 22, 1–16.
- Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research*, 40, 1469–1487. [http://dx.doi.org/10.1016/S0042-6989\(00\)00003-1](http://dx.doi.org/10.1016/S0042-6989(00)00003-1)
- Rensink, R. A., O'Regan, K., & Clark, J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8, 368–373. <http://dx.doi.org/10.1111/j.1467-9280.1997.tb00427.x>
- Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience*, 27, 611–647. <http://dx.doi.org/10.1146/annurev.neuro.26.041002.131039>
- Roelfsema, P. R., Lamme, V. A., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, 395, 376–381. <http://dx.doi.org/10.1038/26475>
- Rosenholtz, R. (2011). What your visual system sees where you are not looking. [Human Vision and Electronic Imaging XVI]. *Proceedings of the Society for Photo-Instrumentation Engineers*, 7865, 786510, 786510–14. <http://dx.doi.org/10.1117/12.876659>
- Rosenholtz, R., Huang, J., & Ehinger, K. A. (2012). Rethinking the role of top-down attention in vision: Effects attributable to a lossy representation in peripheral vision. *Frontiers in Psychology*, 3, 13. <http://dx.doi.org/10.3389/fpsyg.2012.00013>
- Rosenholtz, R., Huang, J., Raj, A., Balas, B. J., & Ilie, L. (2012). A summary statistic representation in peripheral vision explains visual search. *Journal of Vision*, 12, 14. <http://dx.doi.org/10.1167/12.4.14>
- Serences, J. T., & Boynton, G. M. (2007). Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron*, 55, 301–312. <http://dx.doi.org/10.1016/j.neuron.2007.06.015>
- Shin, E., Wan, X. I., Fabiani, M., Gratton, G., & Lleras, A. (2008). Electrophysiological evidence of feature-based inhibition of focused attention across consecutive trials. *Psychophysiology*, 45, 804–811. <http://dx.doi.org/10.1111/j.1469-8986.2008.00679.x>
- Simpson, P. J. (1972). High-speed memory scanning: Stability and generality. *Journal of Experimental Psychology*, 96, 239–246. <http://dx.doi.org/10.1037/h0033672>
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PLoS ONE*, 3, e1699. Retrieved from <http://www.plosone.org/article/info:doi%2F10.1311%2Fjournal.pone.0001699>
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2009). V4 activity predicts the strength of visual short-term memory representations. *The Journal of Neuroscience*, 29, 7432–7438. <http://dx.doi.org/10.1523/JNEUROSCI.0784-09.2009>
- Sligte, I. G., Vandenbroucke, A. R. E., Scholte, H. S., & Lamme, V. A. F. (2010). Detailed sensory memory, sloppy working memory. *Frontiers in Psychology*, 1, 175. <http://dx.doi.org/10.3389/fpsyg.2010.00175>
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153, 652–654. <http://dx.doi.org/10.1126/science.153.3736.652>
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II. Acta Psychologica*, 30, 276–315.
- Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science, volume 4: Methods, models, and conceptual issues* (pp. 703–863). Cambridge, MA: MIT Press.
- Supér, H., Spekreijse, H., & Lamme, V. A. F. (2001). A neural correlate of working memory in the monkey primary visual cortex. *Science*, 293, 120–124. <http://dx.doi.org/10.1126/science.1060496>
- Supér, H., Spekreijse, H., & Lamme, V. A. F. (2003). Figure-ground activity in primary visual cortex (V1) of the monkey matches the speed of behavioral response. *Neuroscience Letters*, 344, 75–78. [http://dx.doi.org/10.1016/S0304-3940\(03\)00360-4](http://dx.doi.org/10.1016/S0304-3940(03)00360-4)
- Swanson, J. M., & Briggs, G. E. (1969). Information processing as a function of speed versus accuracy. *Journal of Experimental Psychology*, 81, 223–229. <http://dx.doi.org/10.1037/h0027774>
- Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network (Bristol, England)*, 14, 391–412. <http://dx.doi.org/10.1088/0954-898X.14.3.302>
- Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, 113, 766–786. <http://dx.doi.org/10.1037/0033-295X.113.4.766>
- Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.
- Townsend, J. T., & Roos, R. N. (1973). Search reaction time for single targets in multiletter stimuli with brief visual displays. *Memory & Cognition*, 1, 319–332. <http://dx.doi.org/10.3758/BF03198116>
- Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, 76, 282–299. <http://dx.doi.org/10.1037/h0027242>
- Treisman, A., & Geffen, G. (1967). Selective attention: Perception or response? *The Quarterly Journal of Experimental Psychology*, 19, 1–17. <http://dx.doi.org/10.1080/14640746708400062>
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136. [http://dx.doi.org/10.1016/0010-0285\(80\)90005-5](http://dx.doi.org/10.1016/0010-0285(80)90005-5)
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107–141. [http://dx.doi.org/10.1016/0010-0285\(82\)90006-8](http://dx.doi.org/10.1016/0010-0285(82)90006-8)
- Tsal, Y., & Makovski, T. (2006). The attentional white bear phenomenon: The mandatory allocation of attention to expected distractor locations. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 351–363. <http://dx.doi.org/10.1037/0096-1523.32.2.351>
- Tseng, Y. C., Glaser, J. I., Caddigan, E., & Lleras, A. (2014). Modeling the effect of selection history on pop-out visual search. *PLoS ONE*, 9, e89996.
- Vandenbroucke, A. R. E., Sligte, I. G., Fahrenfort, J. J., Ambroziak, K. B., & Lamme, V. A. F. (2012). Non-attended representations are perceptual rather than unconscious in nature. *PLoS ONE*, 7, e50042. <http://dx.doi.org/10.1371/journal.pone.0050042>
- VanRullen, R., Carlson, T., & Cavanagh, P. (2007). The blinking spotlight of attention. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 104, 19204–19209. <http://dx.doi.org/10.1073/pnas.0707316104>
- VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, 7, 207–213. [http://dx.doi.org/10.1016/S1364-6613\(03\)00095-0](http://dx.doi.org/10.1016/S1364-6613(03)00095-0)
- Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, 31, 523–535. [http://dx.doi.org/10.1016/S0896-6273\(01\)00392-0](http://dx.doi.org/10.1016/S0896-6273(01)00392-0)



- Vlaskamp, B. N., Over, E. A., & Hooge, I. T. C. (2005). Saccadic search performance: The effect of element spacing. *Experimental Brain Research*, 167, 246–259. <http://dx.doi.org/10.1007/s00221-005-0032-z>
- Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, 19, 1395–1407. <http://dx.doi.org/10.1016/j.neunet.2006.10.001>
- Wan, X., & Lleras, A. (2010). The effect of feature discriminability on the inter-trial inhibition of focused attention. *Visual Cognition*, 18, 920–944. <http://dx.doi.org/10.1080/13506280903507143>
- Ward, E. J., & Scholl, B. J. (2015). Inattention blindness reflects limitations on perception, not memory: Evidence from repeated failures of awareness. *Psychonomic Bulletin & Review*, 22, 722–727. <http://dx.doi.org/10.3758/s13423-014-0745-8>
- Williams, D. E., & Reingold, E. M. (2001). Preattentive guidance of eye movements during triple conjunction search tasks: The effects of feature discriminability and saccadic amplitude. *Psychonomic Bulletin & Review*, 8, 476–488. <http://dx.doi.org/10.3758/BF03196182>
- Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, 1, 202–238. <http://dx.doi.org/10.3758/BF03200774>
- Wolfe, J. M. (1998). What can 1,000,000 trials tell us about visual search? *Psychological Science*, 9, 33–39. <http://dx.doi.org/10.1111/1467-9280.00006>
- Wolfe, J. M. (2005). Guidance of visual search by preattentive information. In L. Itti, G. Rees, & J. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 101–104). San Diego, CA: Academic Press. <http://dx.doi.org/10.1016/B978-012375731-9/50021-5>
- Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., Kuzmova, Y. I., & Sherman, A. M. (2011). Visual search for arbitrary objects in real scenes. *Attention, Perception, & Psychophysics*, 73, 1650–1671. <http://dx.doi.org/10.3758/s13414-011-0153-3>
- Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 483–502.
- 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. <http://dx.doi.org/10.1037/0096-1523.15.3.419>
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5, 495–501. <http://dx.doi.org/10.1038/nrn1411>
- Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J., & Bompas, A. (2002). Segmentation of objects from backgrounds in visual search tasks. *Vision Research*, 42, 2985–3004. [http://dx.doi.org/10.1016/S0042-6989\(02\)00388-7](http://dx.doi.org/10.1016/S0042-6989(02)00388-7)
- Wolfe, J. M., Võ, M. L. H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, 15, 77–84. <http://dx.doi.org/10.1016/j.tics.2010.12.001>
- Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, 11, 269–274. <http://dx.doi.org/10.3758/BF03196569>
- Woodman, G. F., Vogel, E. K., & Luck, S. J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science*, 12, 219–224. <http://dx.doi.org/10.1111/1467-9280.00339>
- Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, 115, 787–835. <http://dx.doi.org/10.1037/a0013118>
- Zhang, X., Huang, J., Yigit-Elliott, S., & Rosenholtz, R. (2015). Cube search, revisited. *Journal of Vision*, 15, 9. <http://dx.doi.org/10.1167/15.3.9>

## Appendix A

### Mathematical Modeling of Stochastic Processing Times

Townsend and Ashby (1983) proposed that cognitive processing of a set of discrete items, such as searching through an array of visual stimuli, can be modeled by a set of stochastic processes. Processing of each item in the set is considered an individual stochastic process. How the individual processes interact with each other and determine the process as a whole can be mathematically defined and differentiated. Thus different stochastic models can vary on several dimensions and have different predictions about processing time. Based on the consistency of patterns of predicted and observed data, some models can then be recognized as formally equivalent to the actual underlying cognitive mechanism while others can be ruled out.

Specifically, one can distinguish serial from parallel processing, self-terminating from exhaustive processing, and models having either fixed or unlimited capacity, as well as with or without

reallocation of resources. This framework therefore can host a variety of theoretical accounts of many cognitive functions. With appropriate assumptions about individual stochastic processing of each item, one could proceed to predict each model's performance by either mathematical computation or numerical simulation.

Townsend & Ashby assumed that each individual process should have a constant hazard function for processing completion. Hazard function  $h(t)$  is defined as the conditional probability of an event (in this case the completion of this item) occurs at time  $t$  given that it has not occurred before. Therefore,

$$h(t) = \frac{f(t)}{1 - F(t)} = w \quad (1)$$

It follows that the completion time of one individual process should be a random variable with an exponential distribution

(Appendices continue)



$$f(t) = we^{-wt} \quad (2)$$

Now one can find the expected completion time for one individual process with rate  $w$ :

$$\mathbb{E}(t) = \int_0^\infty t f(t) dt = \frac{1}{w} \quad (3)$$

With this knowledge, an expected time cost for processing a set of such items can be predicted for various models. For example, consider the simplest, serial model, where the system serially processes one item after another. If the model is assumed to be exhaustive, that is, processing terminates only after all items are completed, then the expected time cost would be

$$\mathbb{E}_{EX}(T) = \frac{n}{w} \quad (4)$$

where  $T$  denotes the completion time for the entire system,  $n$  is the number of items, and  $w$  the rate of processing (assumed to be identical for all items, or, items are assumed to be identical). Now if the model is thought to be self-terminating, that is, the system is looking to process one critical item (for instance, the target in visual search display), and the order of processing is random, then expected time cost would be

$$\mathbb{E}_{ST}(T) = \frac{n+1}{2w} \quad (5)$$

In both cases expected time cost increases linearly with the number of items (or set-size)  $n$ . One model of special interest to us is the parallel, unlimited capacity, exhaustive processing model. Maintaining the assumption of an exponential distribution, and adding the assumption that completion of any one item's processing does not affect other items' processing (i.e., interchannel independence, no reallocation of resources), it can be shown that the expected time cost is

$$\mathbb{E}_{EX}(T) = \sum_{i=1}^n \frac{1}{iw} = \frac{1}{w} \sum_{i=1}^n \frac{1}{i} \quad (6)$$

It is worth noting that the total expected time cost to process  $n$  items can also be understood as the time cost to process the longest element in the display because the process is exhaustive. In Equation (6), this expected cost is actually decomposed into the sum of the marginal costs that arise each time an additional element is added to the display. Thus, because the architecture is parallel, it is important to note that the  $i$ th element in the sum is *not* equal to the time to process the  $i$ th element. Rather, it reflects the *additional* cost to finish a display with  $i$  elements compared to a display containing only  $(i-1)$  elements.

A simple proof for (6) is as follows:

For  $n$  items being processed whose completion time follows the exponential distribution as in equation (2), the probability for one item still being processed at time  $t$  is

$$R(t) = 1 - F(t) = 1 - \int_0^t f(t) dt = e^{-wt} \quad (7)$$

Because processing of items is independent, the probability that at time  $t$ , all  $n$  items are still being processed would simply be

$$\mathcal{R}(t) = (R(t))^n = e^{-nwt} \quad (8)$$

Therefore the cumulative probability that one of the  $n$  items gets completed at time  $t$  (cumulative density function, CDF) is

$$\mathcal{F}(t) = 1 - \mathcal{R}(t) = 1 - e^{-nwt} \quad (9)$$

Recall that for an individual exponential process, CDF  $F(t) = 1 - e^{-wt}$  and expected completion time is  $\frac{1}{w}$ , and observe that the event that one out of  $n$  items being processed gets completed also has an exponential distribution, thus we deduce by formal equivalence that its expected time would simply be  $\frac{1}{nw}$ .

Once one item gets completed, we only need to consider the rest  $n-1$  items, and expected time cost for this stage would naturally be  $\frac{1}{(n-1)w}$ .

Recursively, the following stages would have  $\mathbb{E}(t)$  at  $\frac{1}{(n-2)w}$ ,  $\frac{1}{(n-3)w}$ ,  $\dots$ ,  $\frac{1}{w}$ .

Hence the expected time cost for the entire process should be

$$\mathbb{E}_{EX}(T) = \frac{1}{nw} + \frac{1}{(n-1)w} + \dots + \frac{1}{w} = \sum_{i=1}^n \frac{1}{iw} \quad (10)$$

Here we would like to observe that equation (6) is essentially the harmonic series up to  $n$  multiplied by a constant  $\frac{1}{w}$ . Recall that for a harmonic series,

$$\sum_{i=1}^n \frac{1}{i} = \ln(n) + \gamma + \varepsilon_n \quad (11)$$

where  $\gamma$  is the Euler-Mascheroni constant and  $\varepsilon_n \sim \frac{1}{2n}$ , which approaches zero as  $n$  approaches infinity. Therefore the prediction of this model fits closely with our observed data.

Another model whose predictions are relevant is the parallel, unlimited capacity model with reallocation. It was shown by Townsend and Ashby (1983) that for this model,

$$\mathbb{E}_{EX}(T) = \frac{1}{w} \quad (12)$$

$$\mathbb{E}_{ST}(T) = \frac{n+1}{2nw} \quad (13)$$

Inspired by this finding, we proceeded to further explore different models' predictions relevant to visual search research findings. We used a different conceptualization of individual item processing from Townsend & Ashby's exponential process. Each item's processing was modeled as a noisy accumulator that completes when the accumulated amount of evidence reaches a certain threshold. An analytic prediction of expected completion time of

this process is difficult if not impossible, but this way of modeling is very convenient for numerical simulation. It also allowed us to set up different thresholds to simulate different items (i.e., targets and distractors) being processed while keeping the evidence accumulation rate equally constant for all channels, whereas for Townsend & Ashby's work, only one parameter  $w$  controlled the speed of individual processing, and under the case where all  $w$  were equal, they cannot incorporate differences in item identity.

We simulated four models as explained in Figure 3. The set-up of one target item among several identical distractor items is meant to imitate the case of simple feature search or pop-out search, like our Experiment 1. This type of search task is thought to require only stage-one processing, and since people have agreed that stage-one should be parallel with unlimited capacity, we ran four possible variations of this class of model. Only the parallel, independent, exhaustive model (Figure 3a) exhibited the logarithmic RT-set size pattern that we found in our experiments. Briefly, this is because with more distractor items, there is a greater chance that some distractor items haven't completed when target processing completes, and the increment caused by each additional distractor decreases as number of distractor increases. For reallocation models (Figure 3c and 3d), because the system is assumed to be unlimited capacity, once channels start completing, each individual channel's processing speed accelerates, so that on average each

channel is accumulating evidence at a greater rate for a larger number of items. Therefore these models would predict a decreasing time cost for greater number of items.

For all of the four variations, we included a condition in which the distractors have equal threshold as the target, this was meant to provide a comparison point to Townsend & Ashby's original results, since they were working with functionally identical items. Most of our results match, except for the reallocated, exhaustive model. In our simulation when distractors and target had the same threshold, time cost increased a little with increasing number of distractors. Townsend and Ashby (1983) had predicted that this function should be a flat straight line. We reasoned that this difference arose from the difference in our modeling of individual item's processing: ours as a noisy accumulator as opposed to their exponential process.

The key insight from our work is that different lure-target similarity levels ought to require different amounts of evidence accumulation (i.e., different thresholds) before a decision can be made about a lure. Further, differences in decision thresholds produce differences in the slope of the logarithmic curves (the constant that is being multiplied to the harmonic series). In fact, our simulation indicates that there is a linear dependency between distractor thresholds and logarithmic coefficient given a fixed target threshold (see Figure A1 below).

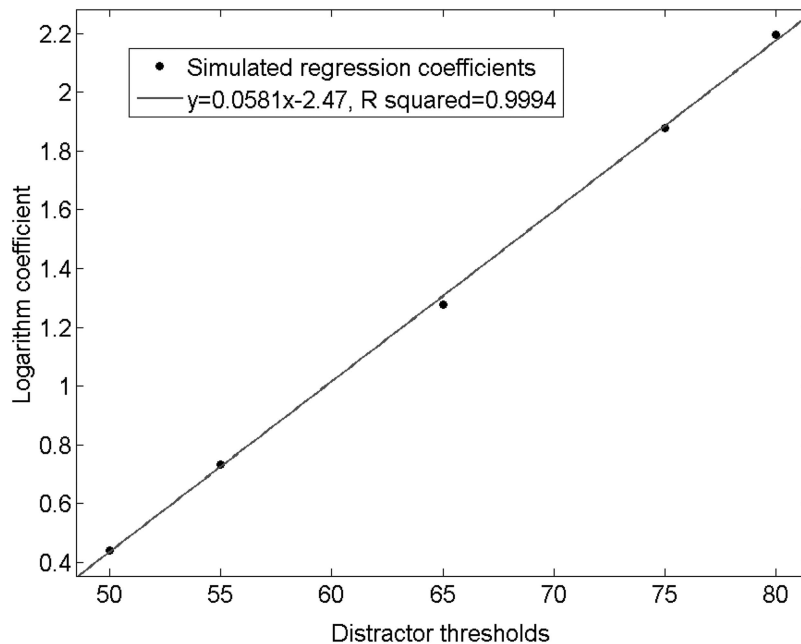


Figure A1. Linear relationship between distractor thresholds and simulated regression coefficient for the logarithmic term, when target threshold is fixed at 80. We used simulation results from Figure 3 (a) plus an additional distractor threshold value at 50. We fitted our simulation results to the following model equation:  $RT = a \ln(\text{set size}) + b$ , and plotted each estimated value of coefficient  $a$  against distractor threshold values. A strong linear trend can be observed as evidenced by  $R^2$  at 0.9994.

(Appendices continue)

## Appendix B

### Position Analyses

#### Experiments 1A and 1B

We present one final post hoc analysis of the data from Experiments 1A and 1B. In the introduction, we argued that there ought to be factors like eccentricity that impact stage-one processing times. Unfortunately, we do not have enough observations to break up our data by set size (five levels), lure type (two levels), and position in the display (36 possible target locations in our display grid). That said, we can ask whether eccentricity impacts processing speeds by collapsing across both set size and lure type and plotting overall RTs by target position in the display. Doing so we obtain an average of about 22 observations per target position per subject, which allows us to compute noisy, but likely meaningful averages. The RT-by-position data can be seen in Figure A2 (note: in subsequent RT-by-position figures, the positions of every spot are adjusted from actual layout in order to best visualize the eccentricity effects, when present). As can be seen in the figure, there is a clear eccentricity effect in the data for both Experiments 1A and 1B. To test the statistical significance of this effect, we averaged the data at three different eccentricities from fixation: the inner-most four locations around fixation (near eccentricity), the outer-most four locations (far eccentricity) and finally, we aver-

aged the four locations along the diagonal that fell in between the inner-most and outer-most corners of the display (intermediate eccentricity). Computed as such, eccentricity had a significant effect on RT in both Experiment 1A,  $F(2, 38) = 54.424$ ,  $p < .001$ , Cohen's  $f = 1.692$ , 95% Confidence Interval = [1.339, 2.394]; and Experiment 1B,  $F(2, 38) = 95.791$ ,  $p < .001$ , Cohen's  $f = 2.246$ , 95% Confidence Interval = [1.788, 3.494]. Note that the presence of an eccentricity effect is also indirect evidence that eye movements probably did not contribute much to search performance in this experiment and that participants gathered significant amounts of information while their eyes were at fixation before moving attention to the target. The Figure also makes it clear that well-known search biases that are often present in difficult searches, such as systematically searching for a target from left to right and top to bottom, are not present in the data (Townsend & Roos, 1973; Hockley, 1984).

The presence of this eccentricity effect does not undermine the general claim that stage one proceeds in a parallel, unlimited capacity fashion, with exhaustive processing. Rather, we believe it illustrates that the speed of evidence accumulation at any given location is modulated by eccentricity, and may be affected by other

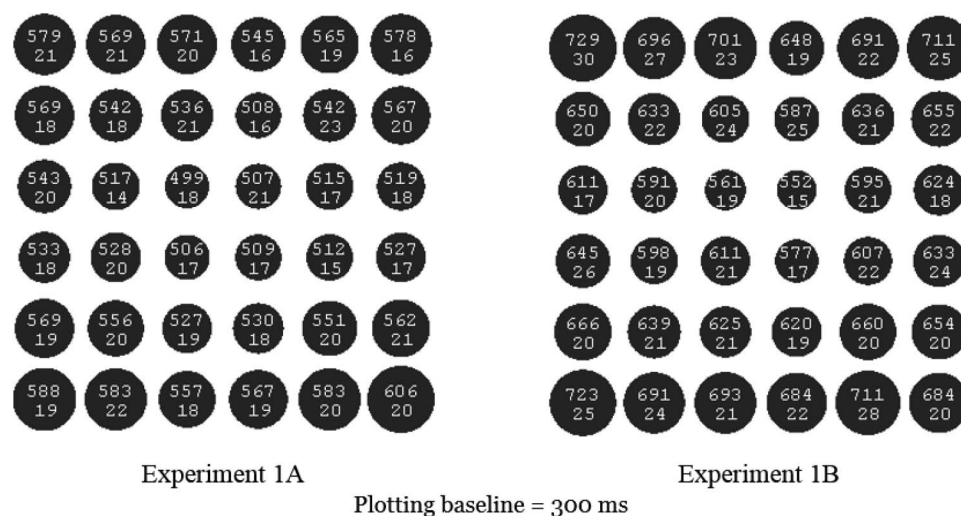


Figure A2. Mean response time in Experiments 1A (left panel) and 1B (right panel) as a function of target position in our displays. Each circle indicates the mean response time (top number) observed at that location in the grid, as well standard error of the mean (bottom number). The grid locations in the image are directly proportional to the grid locations in our displays. To help visualize the results, the diameter of each circle was made proportional to the mean RT of each target location minus 300 ms. This visualization demonstrates how RTs increase as eccentricity of the target increases.

(Appendices continue)

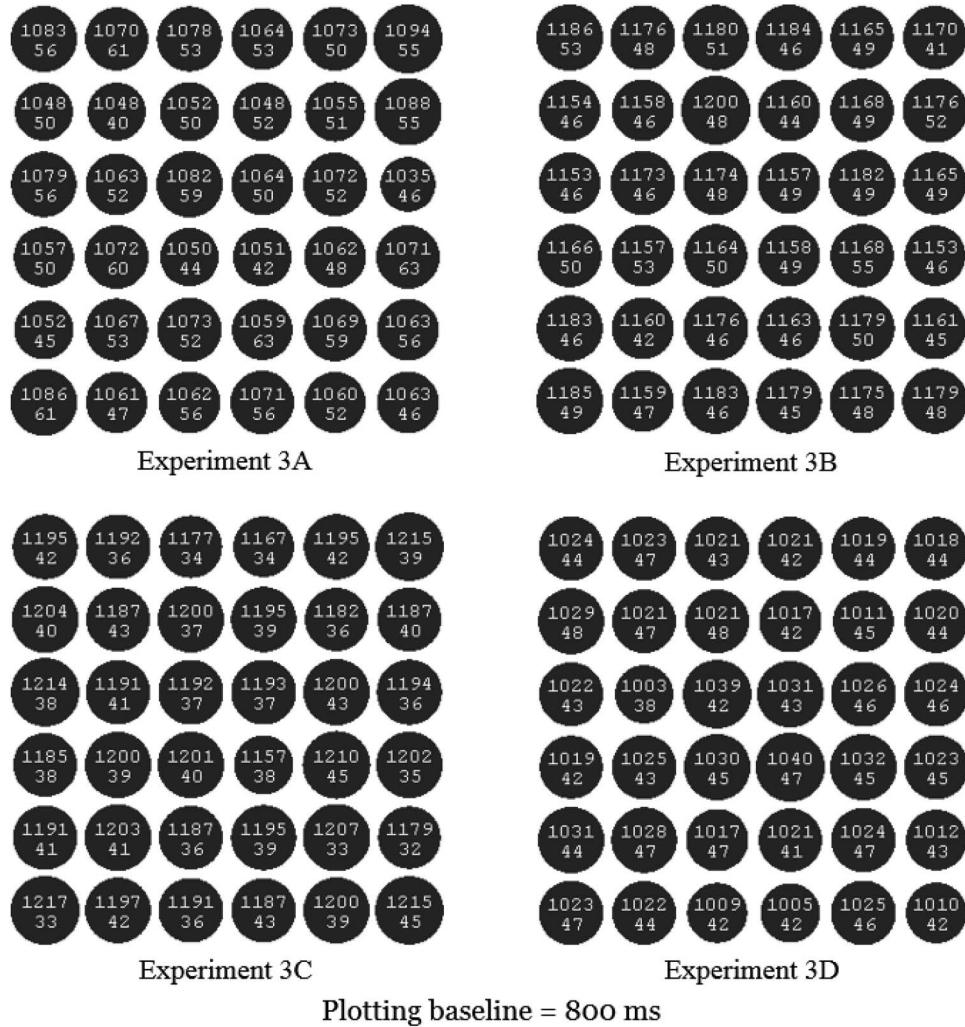


Figure A3. Mean response time in Experiments 3A through 3D as a function of target position in our displays. Top number in each circle indicates the mean RT value and bottom number is the standard error of mean for RTs with each target position. Diameters of circles are proportional to mean RT value minus 800 ms.

factors still, like signal-to-noise ratio. As such, these types of factors may be fundamentally different from the lure-target similarity effects we demonstrated in Figures 5 and 6 because we proposed lure-target similarity affects decision thresholds, rather than rates of evidence accumulation. That said, we do not have enough data at the moment to meaningfully test this interpretation and more research will be needed to investigate whether the eccentricity effect is additive or interactive with the lure-target similarity effect.

### Experiments 3A Through 3D

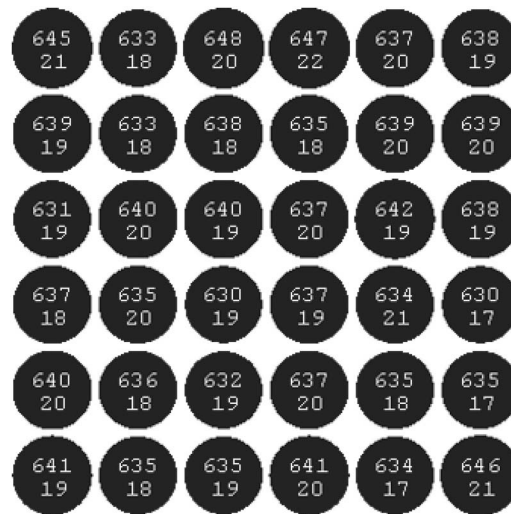
As can be seen in Figure A3, there were no systematic effects of position on RT in any of the experiments.<sup>4</sup> Because we were expect-

ing that multiple eye movements would be required on each trial to find the target, we anticipated that the initial eccentricity of the target would matter little to the overall RT. This contrasts nicely with the eccentricity effects found in efficient search, as in Experiments 1A and 1B. In spite of the search task being substantially more difficult than in those experiments (as indicated by longer RTs), there was also no indication that participants systematically used display scanning strategies such as left-to-right or top-to-bottom inspections.

<sup>4</sup> Experiment 3A:  $F(2, 192) = 0.2144, p = .8072$ ; Experiment 3B:  $F(2, 288) = 0.1385, p = .8707$ ; Experiment 3C:  $F(2, 264) = 0.4094, p = .6645$ ; Experiment 3D:  $F(2, 276) = 0.1486, p = .8620$ .

(Appendices continue)





Plotting baseline = 300 ms

*Figure A4.* Mean response time in Experiments 4 as a function of target position in our displays. Top number indicates the mean RT value and bottom number is the standard error of mean for RT with each target position. Diameters of circles are proportional to mean RT value.

#### Experiment 4

Although the statistics reveal a significant effect of eccentricity,  $F(2, 38) = 4.573$ ,  $p = .017$ ,  $f = 0.491$ , it is readily visible in [Figure A4](#) that the mean effect of 7 ms (computed as the difference between the outermost and the innermost eccentricities) is much less pronounced than in Experiments 1A (82ms) and 1B (136ms). We take this result as further evidence of the different way in which displays were processed in Experiment 4 compared with Experiments 1A and 1B, when the target was fixed.

#### Experiment 2

Because of an oversight in our part, we failed to record the target position on every trial, making a position analysis for this experiment impossible.

Received March 4, 2015

Revision received February 2, 2016

Accepted February 18, 2016 ■

#### E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://notify.apa.org/> and you will be notified by e-mail when issues of interest to you become available!