Competitive guided search: Meeting the challenge of benchmark RT distributions

Rani Moran*
School of Psychological Sciences, Tel Aviv University, Ramat Aviv, Israel

Michael H. Zehetleitner*
General and Experimental Psychology, Ludwig-Maximilians-Universität München, Munich, Germany

Hermann Müller
General and Experimental Psychology, Ludwig-Maximilians-Universität München, Munich, Germany
School of Psychology, Birkbeck College, University of London, UK

Marius Usher
School of Psychological Sciences, Tel Aviv University, Ramat Aviv, Israel

Historically, visual search models were mainly evaluated based on their account of mean RTs and accuracy data. More recently, Wolfe, Palmer, and Horowitz (2010) have demonstrated that the shape of the entire RT distributions imposes important constraints on visual search theories and can falsify even successful models such as guided search, raising a challenge to computational theories of search. *Competitive guided search* is a novel model that meets this important challenge. The model is an adaptation of guided search, featuring a series of item selection and identification iterations with guidance towards targets. The main novelty of the model is its termination rule: A quit unit, which aborts the search upon selection, competes with items for selection and is inhibited by the saliency map of the visual display. As the trial proceeds, the quit unit both increases in strength and suffers less saliency-based inhibition and hence the conditional probability of quitting the trial accelerates. The model is fitted to data the data from three classical search task that have been traditionally considered to be governed by qualitatively different mechanisms, including a spatial configuration, a conjunction, and a feature search (Wolfe et al., 2010). The model is mathematically tractable and it accounts for the properties of RT distributions and for error rates in all three search tasks, providing a unifying theoretical framework for visual search.

Introduction

Visual search is an experimental paradigm that has attracted intensive research over the last 50 years in the vision-research literature for three important reasons. First, it is commonly used in daily activity and is essential to our survival, as when we search for a target (say, our child) among similar distractors (children of other parents). Second, it has been at the forefront of the research on attentional processing, with a number of alternative theories suggesting that, depending on the search material, observers either spread their attention over the whole display (Shaw, 1982; Palmer, Verghese, & Pavel, 2000) or sequentially deploy it from one item to another (Treisman & Gelade, 1980; Wolfe, Cave, & Franzel, 1989). Third, it is a complex task, involving both perceptual and decision processes (Palmer & McLean, 1995; Zehetleitner & Müller, 2010; Zehetleitner, Rangelov, & Müller, 2012), which has generated a number of alternative ways to account for the data by shifting the explanatory locus from attentional to decision processes.

The typical finding that has driven research in visual search is the *set-size* effect: the finding that, depending on the type of search stimuli, the mean reaction time (RT) for target detection increases (or does not

Citation: Moran, R., Zehetleitner, M. H., Müller, H., & Usher, M. (2013). Competitive guided search: Meeting the challenge of benchmark RT distributions. *Journal of Vision*, *XX*(XX):XX, XX–XX, http://www.journalofvision.org/content/XX/XX/XX, doi:10.1167/XX.XX.XX.

doi: 10.1167/XX.XX.XX Received January 31, 2013; published Month 0, 2013 ISSN 1534-7362 © 2013 ARVO

increase) with the number of items in the display (set size). Typical data shows that, in easy search tasks (feature search, e.g., finding a red among green circles; e.g., Egeth, Jonides, & Wall, 1972), the mean RT is flat with set size, in contrast to difficult search tasks in which the slope of the RT/set-size function (i.e., the mean RT increment when the set size is increased by one item) is considerably higher (e.g., finding a 2 among 5s or a red vertical line among red horizontal and green vertical lines; e.g., Atkinson, Holmgren, & Juola, 1969). In such difficult tasks, the time to find a target in the displays increases approximately linearly with the number of distractors. This effect of set size, together with the finding that the search slope is about twice as high when the target is absent as compared to present, has motivated a number of theories about the deployment of attention in search. For example, it was concluded that, while in simple feature search one can use bottom-up feature contrast signals to directly allocate attention to the target position, in more complex search tasks one typically has to shift attention serially between the items in order to locate the target (e.g., Treisman & Gelade, 1980; Wolfe et al., 1989; Itti & Koch, 2000, 2001). For several years, search tasks with flat set-size functions have been referred to as *parallel* and those with steeper functions as *serial* searches.

Early on, however, it has been noted that mean RT data provide only weak constraints for search models. Empirical mean-RT patterns are not a signature of a parallel versus serial processing architecture. Rather, parallel and serial models could mimic each other; in particular, both serial and parallel models can generate either positive or flat slopes (Townsend & Ashby, 1984; Townsend & Nozawa, 1995). As a consequence, Wolfe (1998) has proposed to use the continuum between efficient and inefficient to characterize search tasks, rather than terms that postulate an underlying cognitive architecture, such as parallel/serial search.

A potential remedy for the problem of model mimicry is to consider entire RT distributions: Whereas mean RT is merely a central tendency summary of many observations, the shape of the whole RT distribution contains much more information about the underlying decision processes (Ratcliff, 1978). Thus, challenging models with the necessity to account for entire distributions may permit alternative models to be distinguished which are indistinguishable on the basis of mean RTs alone (Cousineau & Shiffrin, 2004; Wolfe, Palmer, & Horowitz, 2010; Balota & Yap, 2011).

Recently, Wolfe et al. (2010) have published an extensive data set which includes data from three of the standard tasks in the visual-search literature. For these three basic search tasks, the authors have collected a solid basis of more than 100,000 trials. The tasks are a color feature, a conjunction, and a spatial configuration search. In the color feature task, the nontargets were green vertically oriented bars and the target, when present, was a red vertical bar. In the conjunction task, the nontargets were green vertical and red horizontal bars, the target a red vertical bar. In the spatial configuration task, the target had the shape of the digit 2 presented amongst digit 5 nontargets. These tasks are considered to be paradigmatic: search slopes are either flat (1 ms/item), intermediate (9 ms/item), or steep (43 ms/item) for feature, conjunction, and spatial-configuration search, respectively. For each participant and condition (target present/absent, set size, task), some 500 observations had been collected. This large number of observations permits for a rather precise estimate of the RT distributions, as the variance of the estimate of a population's quantile RT obtained from the quantile of a sample of RT observations is inversely proportional to the sample size (Stuart & Kendall, 1977, cf. Ratcliff & Tuerlinckx, 2002). By virtue of the large ?5 number of observations and the selection of prototypical search tasks, these RT distributions provide an important benchmark for models of visual search. Currently, there is as yet no model of visual search which meets this benchmark. One of the most popular models of visual search—guided search—appears to fail this benchmark in its published versions (Wolfe et al., 2010, p. 1309; see discussion below).

The aim of our paper is to develop a computational model of visual search that can account not only for mean RTs but also for RT distributions and error rates. As our model is based on the guided-search framework, we start with reviewing guided search and the problem it encounters in accounting for distributional RT data. We will then elaborate on one key mechanism to any visual search process—*search termination*—which we will argue to be critical for yielding a good account of distributional RT data. In what follows, we present our model and the data fits that it achieves to the benchmark data of Wolfe et al. (2010).We end with a discussion of the main properties that enable the model to achieve a good account of the data and with a discussion of the wider implications for other types of visual search models and to search optimality.

Guided search

Guided search (Wolfe et al., 1989; Wolfe, 1994, 2007) is one of the most popular/prominent models of visual search. It assumes a two-stage processing architecture, where in a first stage the display is processed in parallel (see also Hoffman, 1979). This parallel processing stage results in a salience value for each location in the display. Salience here is understood as a measure of local conspicuity: how distinct a visual item is, compared to all other items, with respect to color,

orientation, luminance, and motion (Koch & Ullman, 1985; Itti & Koch, 2000; Bruce & Tsotsos, 2009). The second processing stage consists of serial selection of items in descending order according to their salience. In each attentional selection, one item is scrutinized for its target identity; that is, each selection leads to a target versus nontarget decision for the selected item. Search is assumed to terminate with a positive response when the first item is classified as a target (i.e., self-terminating search). This model was highly successful in explaining a number of search RT patterns, such as the continuous distribution of search slopes (Wolfe, 1998), the interplay between top-down and bottom-up factors in the guidance of attention, and the high efficiency of some conjunction searches (Wolfe et al., 1989). Guided search can account for high efficiency conjunction search by allowing the first parallel stage to depend not only on bottom-up stimulus features, but also on top-down modulations of the salience map, which influence the order in which items are searched. For instance, in a conjunction search, when the target is a red vertical item, guided search assumes the feature contrasts of "red" and "vertical" to have a greater impact on the salience map than feature contrast generated by other features.

Search termination

A critical component of any search model, including guided search, is the search *termination* mechanism. When a target is identified, this mechanism is assumed to terminate the search generating a "target-present" response. However, the literature is less consistent with respect to search termination when no target is identified. The first possibility is that of an exhaustive search: search terminates, generating a "target-absent" response, only when all items in the display have been rejected (i.e., classified as a nontarget). Note that this search termination rule can be applied not only to two-stage models (such as guided search), but also to parallel models (Ward & McClelland, 1989; Palmer & McLean, 1995) in which all items are processed simultaneously though with varying classification durations. A number of more complex termination rules have also been suggested.

Wolfe et al. (2010) discuss a schematic (simplified) implementation of guided search which is self-terminating when a target is found and exhaustive when no target is present. Specifically, search is terminated when the target has been identified, or when all items in the display have been checked and rejected. The model was aimed at simulating the hardest task (2 amongst 5s) for which it is assumed that the target provides no bottom-up guidance, that is, items are selected for identification in an equiprobable random order (e.g., Treisman & Gelade, 1980; Wolfe, 1994). In their implementation, attention was deployed to an item on average every 98 ms (the deployment times being drawn from a gamma distribution). For target-absent trials (exhaustive search), the number of items inspected prior to response execution was the set size. For target-present trials, the number was drawn uniformly from the integers between one and the set-size (serial self-terminating search, with no saliency-based guidance for the target). The simulated RT was the sum of the times for each deployment plus a motor/nondecision component that was also gamma distributed (200 ms on average). Although the simulation parameters were chosen in such a way that the model produced mean RTs that were similar to the empirical data, Wolfe et al. (2010) pointed out several important discrepancies between the RT distributions of the simulated and of the empirical data.

As can be seen in Figure 1, the simulated target-absent and target-present RT distributions for Set Size 18 exhibit only a small overlap, as opposed to the substantial overlap of the respective empirical distributions (note, in particular, the leading edge); a similar discrepancy appeared for Set Sizes 12 and 6. Furthermore, the simulated target-present RT distributions change their shape from peaked to rectangular as set size increases, whereas the target-absent RT distribution has a much more narrow distribution, unlike that observed in the data. Given this failure of a plausible implementation of guided search to fit the benchmark data, Wolfe et al. (2010, p. 1310) concluded that this model faces a serious challenge in accounting for RT distributions.

Wolfe et al. (2010) point out that the distributional data is also problematic for some other search termination rules. One such rule is to terminate search after a temporal interval has elapsed that would allow successful target detection on P% (e.g., 95%) of the trials. The idea, instantiated in some version of the guided search model, is that target-present responses are determined by a stochastic timer (a stochastic quit accumulator racing to threshold) that is independent of the identification process. However, as Wolfe et al. (2010) argued, such a mechanism mandates a high miss rate if the RT distributions for target-present and target-absent trials overlap substantially. To see this, consider a target-present trial with a "lingering" target-present response, that is, an RT that is longer than the median RT for target-absent responses. Arguably, this response ended up as a hit owing to the target-absent timer ticking slower than its median speed. However, since the quit timer operates independently of the identification process, there was an equal chance that the timer would have expired sooner than its median response time, in which case the target would have been missed. Thus, on average, any lingering hit trial should

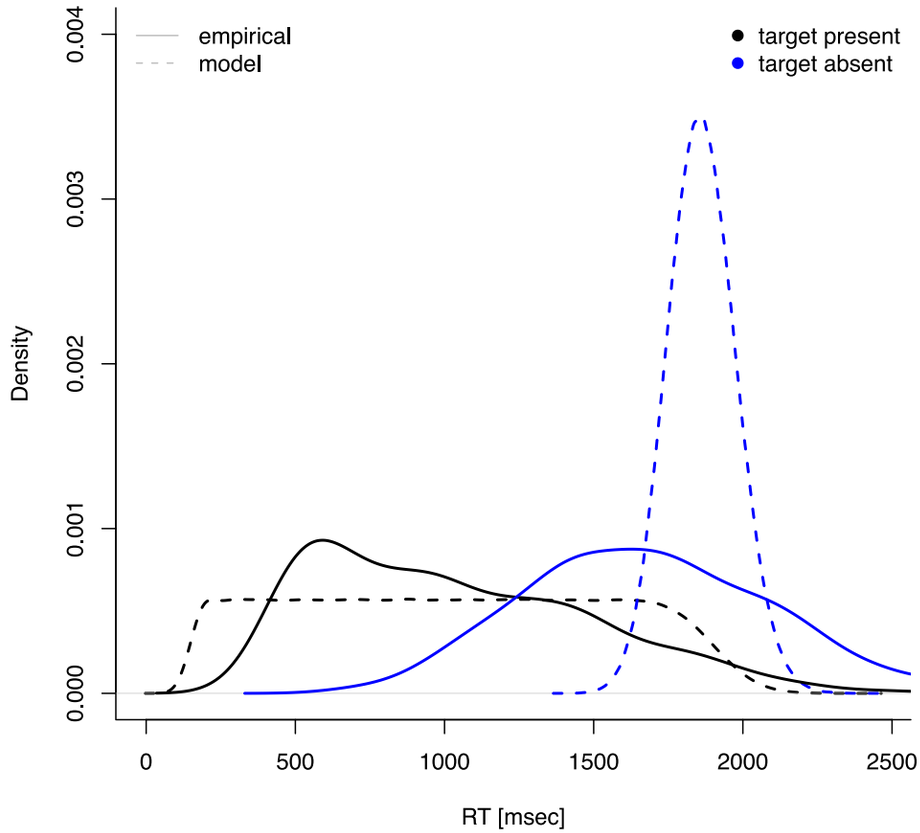

Figure 1. Data reproduced from Wolfe et al. (2010). Target-present and absent RTs of the spatial configuration task for Set Size 18. Black lines correspond to target-present trials and blue lines to target-absent trials. Solid lines refer to empirical data of one prototypical subject. Dashed lines refer to the simplified serial search model with exhaustive termination on nontarget trials (reproduced with the parameter values given in Wolfe et al., 2010).

be matched with a miss trial. This results in a lower bound on miss rates based on the overlap between the target-present and target-absent RT distributions (see Wolfe et al., 2010, Figures 4 and 5). Wolfe et al. (2010) thus concluded that the empirical combination of low miss rates with large distributional overlap renders such a termination mechanism implausible as an account of their data.

Alternative termination rules

An alternative to the timer-based termination scheme is to quit after a certain proportion of items has been examined. This has been implemented in guided search-like two-stage architectures in several ways. First, terminating after checking only a proportion of the items has been implemented via a cutoff in salience (Chun & Wolfe, 1996). Second, the conditional probabilities of terminating search after checking k out of n ($\geq k$) items have been proposed to be free model parameters (e.g., Cosineau & Shiffrin, 2004). And third, termination could be implemented as an item-based (rather than time-based) termination criterion (Wolfe, 2007; Wolfe & Van Wert, 2010). We briefly describe these termination rules below.

In the model of Chun and Wolfe (1996), in rejecting some items as nontargets based on their below-cutoff salience levels without identifying them, a single free parameter controls the average proportion p (<1) of items that are eligible for examination and this proportion is set-size-invariant: The number of eligible items is also the number of items that will be checked unless a target is identified first, whether correctly or mistakenly. For example, for a given cut-off parameter which results in rejecting 25% of the items, on average six items are eligible for Set Size 8, and 12 items for Set Size 16. In the Consineau and Shiffrin (2004) model, the termination probabilities after each nontarget identification are free parameters (estimated from the model), and each additional empirical set size n requires $n-1$ new free parameters. A more parsimonious termination rule was suggested by Donkin and Shiffrin (2011). According to this rule, the termination probability is a logistic function of the proportion of already inspected items, which is characterized by only two free

?6

parameters (location and scale). In the Wolfe and Van Wert (2010) model, item-based termination occurs when a stochastic accumulator that is incremented following each nontarget rejection reaches a boundary (Wolfe & Van Wert, 2010). This strategy also requires a specific strategic criterion for each set size (Wolfe, 2007). None of these three termination rules have as yet been tested on the Wolfe et al. (2010) RT distribution benchmark data.

The most important alternative to guided search is the family of single-stage parallel models. Such models account for set-size effects either on the basis of capacity limitations on the processing rates (Bundesen, 1990) or by adjusting the response criteria (Palmer ?7 1995). Several single-stage parallel models have been devised to specifically account for accuracy (though not RT) data (e.g., Eckstein, 1998; Palmer et al., 2000), while others have accounted for mean-RT data (Ward & McClelland, 1989; Palmer & McLean, 1995; Thornton & Gilden, 2007). The probabilistic parallel search model (PPSM; Dosher, Han, & Lu, 2004, 2010) has been successful in accounting for the temporal course of visual search in speed-accuracy trade-off paradigms (in which accuracy is measured as a function of the manipulated processing time) under conditions of brief display presentation. To the best of our knowledge, none of these models (like the serial ones discussed above) have as yet attempted to fit the benchmark of RT distributions for the most basic search tasks (Wolfe et al., 2010).

Typically, parallel models for RT (Ward & McClelland, 1989; Palmer & McLean, 1995; Thornton & Gilden, 2007) assume that for each item in the display, there is a parallel diffuser whose boundary crossing determines whether the corresponding item is a target or a nontarget. That is, one boundary of each diffusor corresponds to target identification and the other boundary to nontarget identification of that item. A "target-present" response is triggered as soon as any of the diffusers hits the target boundary, and a "target-absent" response as soon as all diffusers hit the nontarget-absorbing boundary. In that sense, search is self-terminating when a target is found and exhaustive until all items are verified as being nontargets. The diffusion process is assumed to be noisy and consequently a nontarget diffuser could reach the target boundary, thus generating a false alarm (on a target-absent trial). A miss occurs when a target is present but all diffusers (including the target diffuser) reach the nontarget boundary.

Consider the behavioral consequences of maintaining the decision boundaries at a fixed level across different set sizes in such a parallel model: As set size increases, so does the probability that one of the nontarget diffusers will accidentally reach the target boundary. This will increase both false-alarm and hit rates. Most typically, however, empirical miss rates increase, rather than decrease, with increasing set size, whereas false-alarm rates stay rather constant (e.g., Chun & Wolfe, 1996). Thus, response boundaries must be adjusted for different set sizes if a parallel model is to capture these qualitative error data patterns. At the same time, such adjustments change item identification durations, accounting for RT set-size effects. The upshot of this discussion is that parallel models require strategic set-size-dependent parameters in order to account for search RT data.

The present model: Competitive guided search

We describe our model in three levels of technical detail. First, we provide a general description which is aimed at any reader regardless of their interest in technical details. The second description is more elaborate, presenting the model implementation in full detail. Finally, we provide, in Appendix A, a comprehensive (mathematically) analytic derivation of the behavioral predictions (accuracy rates and RT distributions) of the model. The model's amenability to mathematical explication equips it with substantial theoretical and computational advantages, which will be discussed.

General description

We now present an implementation of a two-stage search model with a novel termination rule, which can account for the RT distribution benchmarks of Wolfe et al. (2010). The basic model is an adaptation of guided search in its iterative implementation of two stages: selection and identification. The main novelty of our approach lies in the termination rule, which combines two main principles: first, we assume that the search termination mechanism increases the conditional probability of quitting after the inspection of each additional item (e.g., Cosineau & Shiffrin, 2004; Wolfe & Van Wert, 2010); second, we assume the operation of inhibitory connections from the salience map to a termination unit (Zehetleitner, Müller, & Wolfe, 2009).

Let us first consider the flow of processing and decisions in our model. Each iteration begins with a decision whether to terminate search: *quit selection*. If search is terminated, a "target-absent" response is issued. If search is not terminated, a selection decision is made, resulting in attentional selection of another item from the visual display. Next, the selected item is identified either as the target—resulting in a "target-

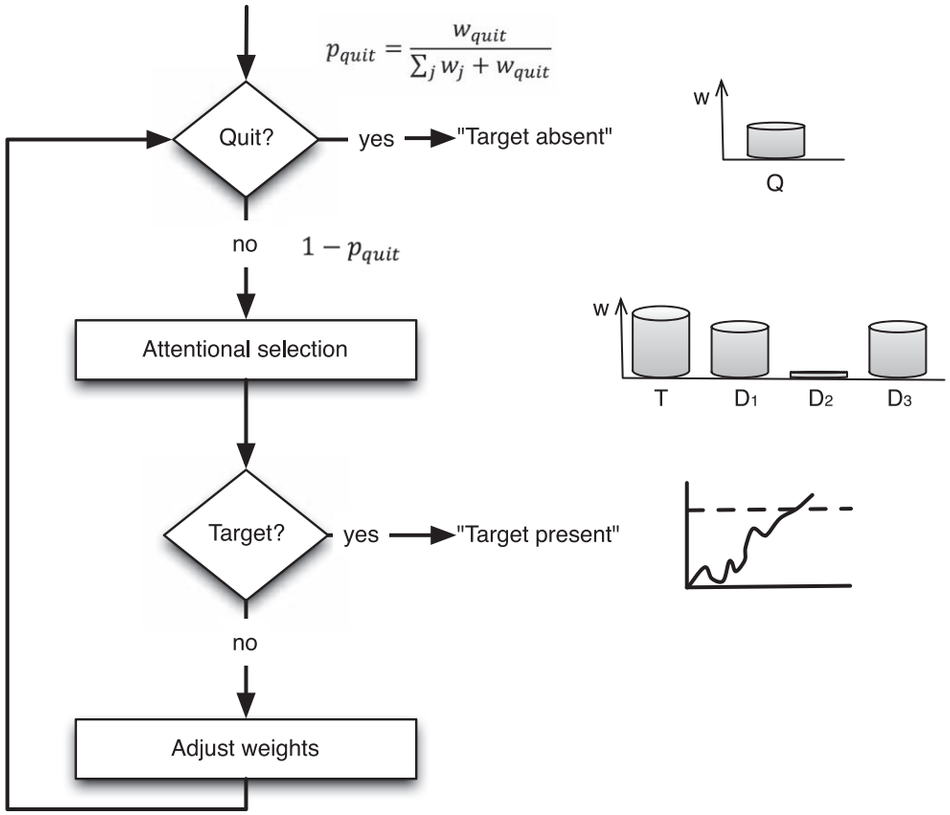

Figure 2. Flow chart depicting the sequence of decisions. When a trial is started, first a quit decision is made. The probability of quitting is described by the equation for p_{quit} :ilt is equal to the weight associated with the quit unit, relative to the overall weight associated with the quit unit and with the display items, w_i. Because w_{quit} is initialized with zero, this never results in a "target-absent" response at the beginning of a trial. Also, search is definitely terminated if all items in the display have been checked, because, in this case, the denominator in the equation is equal to w_{quit}, resulting in a quitting probability p_{quit} of one. If search is not terminated, an item is selected for serial inspection. The probability of selecting the target (if present) is described by Equation 1 (see text below). If the target is selected, a "target-present" response is issued. If a nontarget has been selected, the weights are adjusted, that is w_{quit} is increased by Δw_{quit} and the weight of the just inspected item is set to zero, after which the next quitting decision is made. Responses are subject to a small proportion of motor errors. The icons to the right of the quitting decision and the attentional selection unit denote the weights for the quitting unit as well as the weights of one target, T, and three distractors, D1 through D3. D2 has already been identified as a nontarget and its weight reset to zero. Also, the quitting weight has already been increased. The example illustrates some guidance in the system, as the target weight is slightly larger than the distractor weights. The duration of reaching a target/no-target decision is drawn from a Wald distribution, corresponding to a noisy accumulation process to a single boundary (Luce, 1986). Note that motor errors are not depicted in this schematic.

present" response—or as a nontarget. Following the identification of a nontarget, the nontarget is inhibited and an adjustment is made to the likelihood of terminating the search (see detailed description below), and the next iteration begins (see Figure 2). This search cycle continues with either, eventually, an identification decision determining that a selected item is a target, or a decision to terminate the cycle by quit-selection.

Before discussing each of these decisions in detail, we describe their general properties. First, the selection decision is guided by salience: the more salient an item is, the greater the probability that it is selected first. If the salience difference between targets and nontargets is large enough, the probability that the target is the first item to be selected is very close to one for all empirical set-sizes, resulting in *pop out*: flat set size RT functions. If, on the other hand, there is no salience difference between target and nontargets, search becomes unguided and the order of item selection is random. As such, the model is indifferent as to how item salience is generated, that is, it allows for both bottom-up and top-down influences influences on salience. The more physically distinct the target is from the distractors, but also the more attentional weight is assigned to the target feature (Wolfe et al., 1989) or the target dimension (Found & Müller, 1996), the higher target salience becomes.

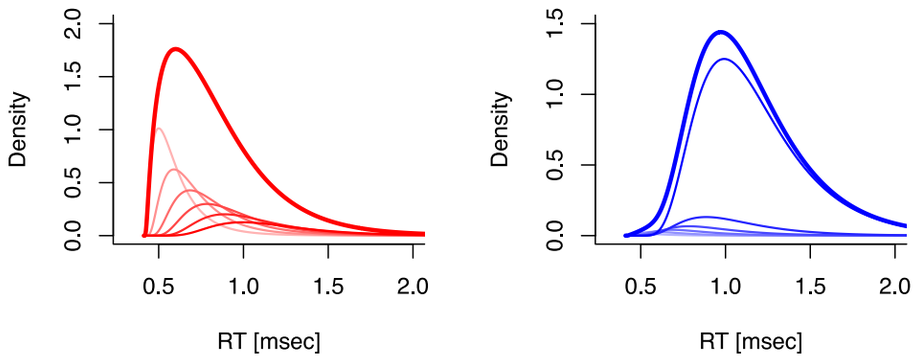

Figure 3. RT probability density functions for hits (left panel) and correct rejections (CR) (right panel) as probability mixtures across the number of item identifications prior to trial termination. The figure shows the constituent defective shifted ex-Wald distributions (thin lines) that correspond to the RTs in the event that *k* items are identified on a hit trial or a CR trial. The mixture (thick) distributions are just the functional (vertical) sum of these defective distributions. The plot is based on the best-fitting parameters of the average observer on the spatial configuration task for Set Size 6.

The termination rule has several general properties. First, search terminates when all items in the display have been identified as nontargets, assuming perfect memory for visited locations (Horowitz, 2006).[1] Second, our termination decision allows for search to be terminated even before all items have been identified. This takes place especially in efficient pop-out searches: After the first item has been identified as a nontarget, it would make little sense to verify that all other items are also nontargets before issuing a "target-absent" response. The reason is that in such tasks, the probability that the first item selected is the target is close to one when a target is present. Consequently, if the first selected item turns out to be a nontarget, it is highly unlikely that a target is present at other locations. Thus, it is relatively safe to quit the search while maintaining a low miss rate.

This nonexhaustive search termination is important for generating a nearly complete overlap of RT distributions for "target-present" and "target-absent" responses as well as a flat search slope for mean "absent" RTs, as is typical for efficient searches (see, e.g., Wolfe et al., 2010, Figure 4, or our Figure 1). Thus, our termination decision allows for both exhaustive and nonexhaustive types of search. Furthermore, in the current termination mechanism, the conditional probability of search termination (prior to target identification) increases following the identification of each nontarget. Regulating the size of this item-to-item increase in termination probability between tasks determines how early the search will terminate: A very large increase results in search termination directly after the first item has been identified as a nontarget (suitable for pop-out searches, where the target—if present—is the first item to be inspected). By contrast, a zero increase results in an exhaustive search, that is, checking all items for their identity. Intermediate levels of the probability increment allow for a smooth interpolation between these two extremes.

Finally, the core novelty of our termination mechanism is its link to the salience map. The overall amount of activity on the salience map regulates (i.e., down modulates) the probability of quitting a trial in the termination decision. This inhibitory link between

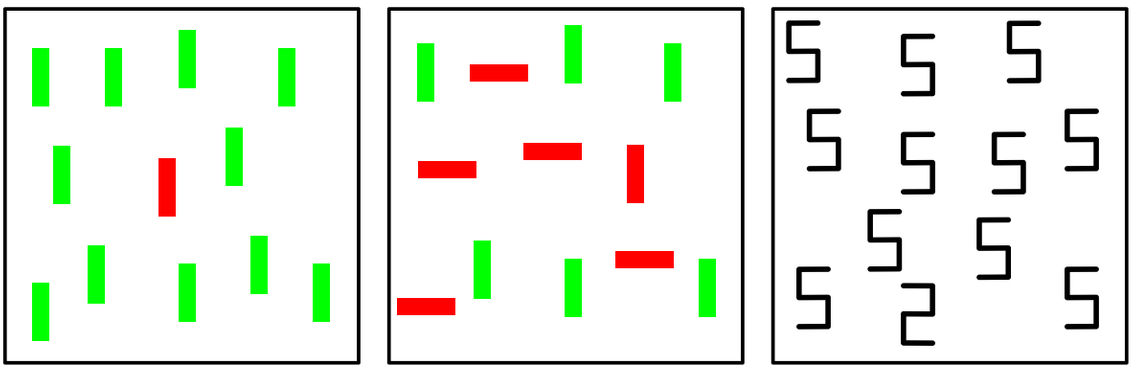

Figure 4. Stimuli for the feature, conjunction, and spatial configuration tasks. The target is the red vertical bar (in the left and middle panels) or, respectively, the digit 2 (in the right panel). Redrawn after Wolfe et al. (2010) with permission.

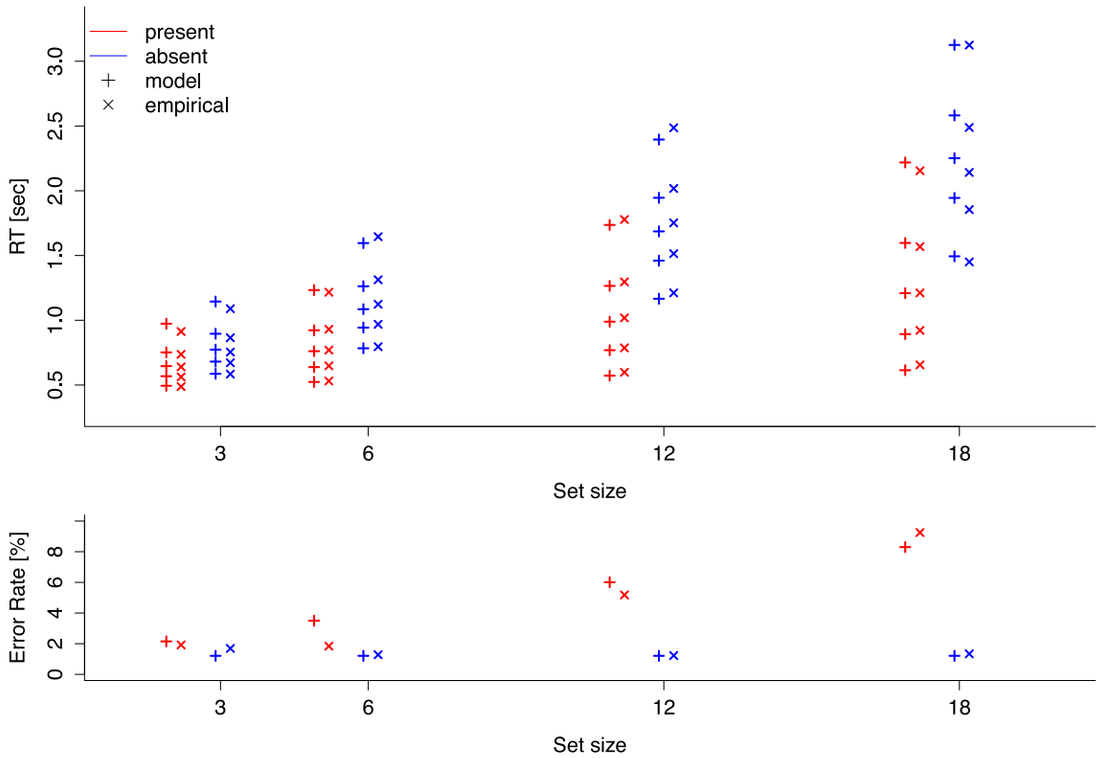

Figure 5. RT distribution and error data for the spatial configuration search task. The top panel depicts quantile RTs and the lower panel error rates for Set Sizes 3, 6, 12, and 18 items, for target-present (black color) and target-absent (blue color) trials. Empirical data from Wolfe et al. (2010) are denoted by symbols x and data predicted by the best-fitting model by symbols +. The empirical quantiles and error rates are obtained by averaging empirical quantiles and error rates, respectively, across individual participants. The model predictions are based on the parameters from the fit of the RT distribution averaged across all observers (the mean distribution).

the salience map and the termination mechanism has several desirable properties. For instance, the amount of activity on the salience map is proportional to set size: the larger the set size, the more nontargets are represented on the salience map. Thus, due to the inhibitory link between the salience map units and the termination unit, an increase in set size causes later search terminations. In other words, the larger the display size, the more items are checked before search is terminated. Consequently, the empirical finding that target-absent RTs increase with set size can be generated by the current model even for nonexhaustive searches (i.e., not checking all items before quitting).

Selection decision

Although the flow chart describing the decision process of our model starts with the termination decision, we commence the detailed description with the selection decision, because this permits the basic computational concepts also used in the termination mechanism to be readily introduced.

In general, the selection stage is conceived as a competition between the display items. In the model, each item in the display is represented by a saliency value. Similar to the guided search model, these saliency values are assumed to be computed in parallel for all items in the display with no temporal cost. Similar to guided search, the saliency values are not maintained at a fixed level during the trial, but rather they change as a consequence of distractor identification and inhibition (see below) in order to prevent reselection of a previously rejected item—under the assumption of perfect memory for visited locations (see also Itti & Koch, 2001). We denote the saliency (or the weight) of item i by $w_i \geq 0$.

The weight of each display item is proportional to its salience level. Salience here is understood as physical (bottom-up) feature contrast of each item to its surround (as quantified for instance by Itti & Koch, 2000, 2001; Gao & Vasconcelos, 2007; Bruce & Tsotsos, 2009; and others). It is well established that bottom-up stimulus factors are not solely responsible for attentional selection. Rather, intentional top-down signals are able to modulate bottom-up salience, for instance, when an observer is looking for a red item, or

?8

for a color pop-out target the specific color of which is not known in advance. In such cases, salience signals from red or from color channels, respectively, are amplified (e.g., Wolfe et al., 1989; Found & Müller, 1996; Navalpakkam & Itti, 2007). The exact computation of salience values is beyond the scope of the present model. It is however based on the assumption that item weights w_i are proportional to the items' salience.

The selection process itself is probabilistic and based on Luce's choice axiom (Luce, 1959). According to Luce's choice axiom, each competitor is selected with a probability that is equal to its relative weight, which is its weight in relation to the sum of weights for each competitor. Thus, each item i is selected with probability

$$p_i = \frac{w_i}{\sum_j w_j}, \quad (1)$$

where the numerator w_i is the weight (salience) of item i and the denominator consists of the sum of saliencies for all items in the display. That display item which is selected is the next one to be (see below).

Note that similar to guided search, target relative to nontarget salience determines the order in which items are selected in a probabilistic stochastic way. In guided search, however, the actual salience value of an item on a single trial is drawn from a normal distribution centered on the theoretical average salience of the item. Zehetleitner, Koch, Goschy, and Müller (2013) have proposed to pay terminological credit to this distinction by referring to the numerical value that an item has on the salience map on a given trial as *selection* salience and the expected value of the distribution from which selection salience values are drawn as *stimulus* salience. Stimulus salience thus denotes the physical property of the display item. All item (selection) saliences are then rank ordered and items selected and identified in descending order of their associated salience. In other words, the actual selection salience value of an item varies from trial to trial stochastically. In contrast, in competitive guided search, the salience value in the selection Equation 1 is assumed to be constant from trial to trial, thus corresponding to the constancy of the stimulus salience value in guided search. In the present model, stochasticity in selection is introduced by Luce's choice axiom, rather than a signal-detection type of mechanism as in guided search. Importantly, in both the present conception and that of guided search, it is possible that a nontarget is selected prior to the target, even though the target has a higher (in guided search: stimulus) salience value.

In competitive guided search, in each iteration of the search process, a new competition takes place. Thus, items are not ranked (according to selection saliency) once and for all in a manner that determines their selection order during the entire trial. Rather, each competition results in a unique winner, and in subsequent selections, new and independent competitions occur. One way to conceive of the above competition is by assuming that each item reflects a neuronal population that fires at a rate that follows a Poisson process with rate w_i. If the selection process consists of eavesdropping on the first incoming neuronal pulse, then it follows that each item is selected according to the probabilities given above (Townsend & Ashby, 1984). Since only relative weights are operative in the current model (i.e., selection will operate in the same manner if all the weights are multiplied by a constant), we adopt the convention that for a display with homogeneous nontargets, $w_{\text{nontarget}} = 1$. Thus, the weight of the distractors serves as a weight-scaling factor.

Guidance favors selection of targets when the target has higher saliency than any of the distractors, that is, when $w_{\text{target}} > 1$. If so, then the target has a higher probability of being selected than any of the distractors. Note that, in principle, the model is flexible enough to allow for targets with salience less than nontargets: If $w_{\text{target}} < 1$, than the target is less salient than the distractors and has a lower chance of being selected. Another important assumption of the model is that the initial weights for the nontargets ($= 1$) and the target (if present) are independent of set size. Note that despite this assumption of set-size-independent target weights, the probability of target selection decreases with set size, because its relative weight decreases (because set size affects the denominator in calculating the selection probability in Equation 1). In the current model, we neglect the duration of the selection process. In other words, item selection is assumed to occur instantly and does not contribute to the trial RT.

Search termination

Having prepared the ground, it becomes apparent how the termination decision is implemented. First, the model assumes the existence of a quit unit (as for instance in Wolfe & Van Wert, 2010). This unit is also characterized by a nonnegative activation weight w_{quit}, which reflects the dynamic tendency to terminate the search, rather than to select a new item from the visual display: The larger w_{quit}, the more likely the search will be terminated. The core novelty of the present model is that the quitting unit is linked to activity on the salience map. That is, the quit unit issues a termination probability proportional to its own activity, w_{quit}, relative to the overall activation on the salience map:

$$p_{quit} = \frac{w_{quit}}{\sum_j w_j + w_{quit}}. \quad (2)$$

If the quit unit is selected, search on this trial terminates with a "target-absent" response. What are the consequences of the inhibitory connection from the salience map to the quit unit? First, for a given quitting weight w_{quit}, the probability of terminating search following the kth distracter identification is reduced as set size increases, because set size increases the summed activity on the salience map. Second, when all items have been checked and rejected, activity on the salience map is zero (see distracter inhibition below) and consequently $p_{\text{quit}} = 1$ (Equation 2). Third, on target-present trials, that is, with targets whose salience may even be just a little larger than that of nontargets, search termination would tend to occur later than on target-absent trials. This is because the total weight of the saliency map would be higher for target-present, compared to target-absent, displays, as it would include the (if even slightly increased) weight of a target. Hence, the quitting probability (Equation 2) would smaller for target-present than for target-absent trials. Consequently, information about the likelihood of target presence, which is latently present within the salience map, is used to modulate the termination probability.

On trial initiation, the model assumes that $w_{\text{quit}} = 0$, which yields a zero probability of terminating the trial before any display item is selected. Later on, as distractors are identified as nontargets, the weight of the quit unit increases, and so does the likelihood of search termination (see below for details). If and when the quit unit is finally selected, a "target-absent" decision is made and the observer proceeds to the response execution stage. Similarly to the duration of item selection, we also neglect the duration of the quit unit selection (though see the General discussion).

Target identification

Once a display item has been selected, it goes through a process of item identification, that is: Is it a target or a nontarget? When a selected item is identified as a target, search is terminated and a target-present response is produced. The identification time is modeled by a stochastic accumulation to a boundary (here, we neglect the probability that a target is misidentified; but see the General discussion). The temporal density of item identifications follows a Wald distribution with three parameters: drift rate, threshold, and noise variance (Luce, 1986):

$$f_{id}(t|\theta, v, \sigma) = \sqrt{\frac{\left(\frac{\theta}{\sigma}\right)^2}{2\pi t^3}} * e^{\frac{-\left(\frac{\theta}{\sigma}\right)^2 \left(t - \frac{\theta}{v}\right)^2}{2\left(\frac{\theta}{v}\right)^2 t}}, \quad (3)$$

where θ denotes the threshold and v the drift rate of the identification decision; the noise level σ is a scaling factor and was fixed at 0.1. For the sake of simplicity, the drift rate, threshold, and noise level parameters for the Wald distribution were identical for targets and distractors and were maintained at a fixed level across set size. In addition, the times required for the identification of different display items are assumed to be independent. Note that the mean item identification time is given by the ratio of the threshold and the drift rate: θ/v.

Despite our assumption of perfect identification, the model can account for errors, that is, misses and false alarms. Misses occur when the quit unit is selected prior to selection of the target. In addition, motor errors (see below) generate both misses and false alarms, even when the decision with respect to target presence is correct. Our adoption of a perfect identification approach was motivated by the very low rate (usually below 2%) of false alarms in the empirical data sets we fit. Provided with data sets with higher false-alarm rates, the model should be elaborated by adopting an error-prone binary choice approach to identification (see also the General discussion).

In summary, when a target is identified, a "target-present" decision is made and the observer proceeds to the response execution stage. Alternatively, if a distractor is identified, the item is inhibited, and the weight of the quit unit is incremented—a process we describe next.

Nontarget inhibition and quit unit activation

When a nontarget is identified, it is immediately and fully inhibited. This is implemented in the model by setting its weight (for the rest of the trial) to zero. In effect, this eliminates the possibility of item reselection (similar to Itti & Koch, 2001). Thus, the model implements *perfect memory*: Observers never return to reexamine a rejected item. Note that this assumption could be relaxed if necessary by replacing full inhibition with a partial inhibition mechanism. For example, the weight of an identified distractor could be inhibited by $p\%$ where p is a free parameter. This would allow the model to implement imperfect memory: Rejected items could be reselected, albeit with a reduced probability.

A second event that follows item inhibition (after having rejected an item as a nontarget) is that the weight of the quit unit is increased according to the following rule:

$$w_{quit,new} = w_{quit,old} + \Delta w_{quit}. \quad (4)$$

The increase of the quitting weight reflects an increased tendency to terminate the search following each additional distractor rejection. Hence, after n rejec-

tions, the weight of the quit unit will be $n\Delta w_{\text{quit}}$ (recall that the initial weight is zero). The size of Δw_{quit} is a free parameter and considered to be under strategic control of the observer. For a more difficult task, Δw_{quit} would probably be smaller than for an easier task, to allow for inspection of a higher proportion of the display items prior to search termination. For a pop-out task, Δw_{quit} would be so large that after the first selection-identification cycle, the probability of terminating the search is very close to one. Note that as more distractors are identified and rejected, two mechanisms are responsible for increasing the probability of terminating search. First, the quit unit increases in absolute (and thus in relative) strength as it is further activated. Second, as distractors are inhibited and fewer items are active on the salience map, the weight of the quit unit increases relative to the total weight of the saliency map, as a consequence of the reduction in effective set size. Note also, that the weight of the quit unit does not depend directly on the elapsed duration of the trial, but rather on the number of distractors that have already been rejected.

An appealing consequence of the (full) distractor inhibition and the quit-activation processes is that following the rejection of all items in a target-absent display, the search automatically terminates. In fact, once all distractors have been rejected, they all have weight zero so that the saliency map is effectively weightless and the relative weight of the quit unit is one, which means that it is necessarily selected. Of course, the quit unit could also be selected already prior to full display identification.

As discussed above, we conceive of the weight of the quit unit as a strategic variable. As such, Δw_{quit} could depend on set size, and it could even vary within a given trial. Nevertheless, guided by parsimony, we chose to maintain Δw_{quit} fixed within trials and across set size. Future implementation might provide extended flexibility to the quit unit activation process.

Finally, we assume that both distractor inhibition and quit-unit activation are immediate or, alternatively, that their durations are absorbed in the identification stage (the Wald distribution). Of course, this assumption, too, could be relaxed in future applications.

Motor errors

We assume that the response execution stage can be distorted by motor errors with a probability m, which is a free parameter of the model. In case of a motor error, the alternative decision-incongruent response is executed. We assume that the parameter m is maintained at a fixed level across set size and that motor errors occur independently across trials and of the preceding processing stages.

Since item identification is perfect in the model, false alarms can only ensue as a consequence of motor errors on target-absent trials. Similarly, on target-present trials, if the target has been identified, then a motor error would yield a miss. Nevertheless, on target-present trials, motor errors do not necessarily cause response errors. For instance, if the quit unit is selected prior to target identification, then an erroneous target-absent decision is made. Consequently, a motor error would yield a target-present response, which is classified as a hit, saving the observer from a miss.

Residual time

The residual time T_{er} accounts for all reaction time variance which is not explicitly accounted for by the search processes described above. It, thus, incorporates encoding times, the time necessary for the first target to be selected, and post-decisional processes such as response planning and execution. We assume that the residual time is independent of set size. As soon as the decision with respect to target presence has been made, a compatible response (subject to motor errors) is executed.

The residual time is modeled as a shifted exponential distribution with shift T_{min} and with rate γ (Schwarz, 2001, 2002). Frequently, residual times are modeled with a uniform distribution with mean T_{er} and range $s_{er.}{}^2$ Schwarz used an exponential distribution for residual time. We added a shift parameter since in the data analysis, a cutoff for fast RTs was applied (see Data analysis methods). Wolfe et al. (2010) discussed the idea that part of the difference in the RT distributions between target-present and target-absent trials might be attributable to differences in the residual time for "target-present" and "target-absent" responses. We decided to put this idea to a test by allowing the shift parameter to vary freely across responses. Thus, the model includes two shift parameters: T_{min}^{yes} and T_{min}^{no}. This model was compared with one that included a single shift parameter to test which of the two alternatives would account better for the data (see statistical methods). The residual time parameters are maintained at a fixed level across set size. In addition, the residual time is independent across trials and independent of the decision time components (item identification):

$$f_{T_{er}}(t|T_{\text{min}},\gamma) = \begin{cases} \gamma e^{-\gamma(t-T_{\text{min}})} & t \geq T_{\text{min}} \\ 0 & t < T_{\text{min}} \end{cases}. \quad (5)$$

In the present model, the target-identification process and the residual time "carry the RT burden." As item (or quit-unit) selection, nontarget inhibition, and quit-unit activation make no contributions to RT, trial RTs are the sum of a residual time and a multiple

of independent identification times. In summary, if k display items were identified prior to response execution, then the trial RT would be distributed as the sum of $k + 1$ independent samples: a single shifted exponential residual time and k identification Wald samples. The overall RT distributions are then probability mixtures across different values of k weighted by the probabilities that the trial terminated following the kth item selection (see Appendix A for mathematical details; see also Figure 3 for an illustration).

Behavioral predictions of competitive guided search

A highly attractive feature of the current model is that it is mathematically tractable: Predictions for behavioral dependent variables such as accuracy rates and RT distributions could be derived in analytic formulas, rather than by simulation. This is advantageous from a theoretical but also from a practical computational perspective.

First, when a model is fully described analytically, its behavioral predictions for accuracy and RT distributions can be calculated precisely. By contrast, if the model is solvable only by methods of computational simulation, its predictions can only be *estimated* and are thus subject to simulation error (running the simulation twice would produce somewhat different results). In the latter case, when the model predictions are examined vis-à-vis empirical data, one is forced to consider sampling errors both in the data and in model predictions, complicating data analysis procedures. Deriving predictions based on simulations of huge sets of trials could somewhat mitigate this concern, though nevertheless potentially resulting in enormous computational complexity. Indeed, when the model is fitted to empirical data, the parameters space is searched for the "best-fitting" parameters, that is, those that yield the closest match between model predictions and data. In the process, model predictions should be derived for (usually, a very large set of) candidate parameters. If each of these derivations is based on an enormous set of simulated model trials, then the fitting procedure might run excruciatingly slowly, exhausting the computational resources. Consequently, one would be forced to derive predictions based on smaller simulated samples. But then, prediction errors might become substantial and compromise the fitting procedure in its quest for the best-fitting parameters. In contrast, by utilizing analytical formulas, one can derive model predictions much faster, sometimes by orders of magnitude. Fitting procedures are hence both more efficient and more robust. In summary, mathematical tractability is a highly desirable property of a model, which should render the model preferable to work with, ceteris paribus. We refer readers who are interested in the analytical derivation of the error rates and RT distributions predicted by the present model to the Appendix A.

Methods

Brief description of the experimental methods of Wolfe et al. (2010)

Wolfe et al. (2010) collected data from a total of 28 participants for three classic search tasks: 9 participants in a feature search (with target defined by color), 10 in a conjunction search (with target defined by a combination of color and orientation), and 9 in a spatial configuration search (with a target 2 among distractor 5s). In each task, four set sizes (3, 6, 12, and 18 items) were crossed with two trial types (target present vs. absent) to create a factorial design with a total of eight conditions. For each participant, about 500 trials were run for each of the eight factorial cells. Both factors were intermixed within experimental blocks, that is, they varied randomly from trial to trial.[3]

Data analysis methods

As a first step in the data analysis, "contaminated" trials were excluded. For the sake of maintaining consistency and comparability with Wolfe et al. (2010), we used their exclusion criteria: Assuming that unreasonably fast and slow trials represent anticipations and attentional lapses, respectively, all trials with RTs < 200 ms and RTs > 4000 ms were removed for the feature and conjunction search tasks, and all trials with RTs < 200 ms and RTs > 8000 ms were removed for the spatial configuration task. Thus, a total of 80 trials, or 0.07% of the entire data set (for all participants and across the three search tasks), was eliminated from analysis (see Wolfe et al., 2010, for further details).

Model fitting

Each of the three search tasks were fitted separately, that is, without parameters being constrained across the tasks. For each data set, we fit the model to the RT distributions of correct responses and to the accuracy rates. Due to the scarcity of errors, we did not attempt to fit error RT distributions. Fits were conducted at the level of individual participants as well as for the average observer, obtained by averaging data across participants (see below).

We now describe the process of fitting the model to the empiric data for an individual participant (in a single search task). We adopted Heathcote, Brown, and Mewhort's (2002) Quantile Maximal Probability (QMP) method for our purposes: For each participant, set size s, and trial type t (target present or absent), we computed the 0.1, 0.3, 0.5, 0.7, and 0.9 quantiles of the correct RT distributions. These quantiles serve as bin or category separators, hence six bins were generated. A seventh category was defined by all the error trials. We then calculated the frequency of trials in each category. For a set size s and a trial type t, we denote these frequencies by $O(t, s)_1$, $O(t, s)_2$, . . . , $O(t, s)_7$. The goal in fitting the model is to find a parameter set which predicts a distribution of trials (across the seven categories) that will best match the empirical distributions.

For a given parameter set and for each set size and trial type, we computed the proportion of simulated trials that occupy each of the seven categories: $x(t, s)_1$, ?10 $x(t, s)_2$, . . . , $x(t, s)_7$. For example, $x(t, s)_1$ is the proportion of simulated responses (for set size s and trial type t) that were correct and whose RT was below the 0.1 quantile of the correct empirical distribution. Similarly, $x(t, s)_5$ is the proportion of simulated responses that were correct and whose RT was between the 0.7 and 0.9 quantiles of the correct empirical distribution. $x(t, s)_7$ is just the proportion of simulated error response. Note that the sum of these category proportions is 1: $\sum_{i=1}^{7} x_i^{t,s} = 1$. Calculation of the $x_i^{t,s}$ was based on the analytical equations we derived (see Appendix A). Thus, $x_7^{t,s}$ are simply P_{miss}^s or P_{FA}^s, that is, the predicted probabilities of miss and false alarms associated with set size s (see Equations 31, 35). The calculations for $1 \leq i \leq 6$ rely on the RT quantiles. For a given RT quantile q, the predicted proportion of hits and correct rejections with RT $\leq q$ are given by $P_{hit}^s \cdot F_{hit}^s$ (Equations 30, 32) and by $P_{CR}^s \cdot F_{CR}^s$ (Equations 34, 46), respectively. P_{hit}^s and P_{CR}^s are the predicted probabilities of hit and correct rejections (for set size s), respectively, and F_{hit}^s and F_{CR}^s are the predicted cumulative density functions (for set size s) for the RTs for hits and correct rejections, respectively. The $x_i^{t,s}$s were readily obtained from those predictions. For example, $x_2^{t,s}$ was the difference of the predicted proportions for the 0.3 and the 0.1 quantiles for trial type t and set size s.

In the QMP method, the data is interpreted as being drawn from a multinomial distribution: The model predicts categorical probabilities, and each empirical trial is classified as belonging to one of the (seven) categories. In this perspective, the likelihood of the empirical data given the model predictions for the current parameter set is: $\prod_{i=1}^{7} x(t, s)_i^{O(t,s)_i}$; and across all set-sizes and data types the likelihood is:

$$L = \prod_{s=3,6,12,18} \prod_{t=target\ present,\ absent} \prod_{i=1}^{7} x(t, s)_i^{O(t,s)_i}. \quad (6)$$

Finding the best-fitting parameters involves identifying the parameters that will maximize the likelihood of the data—or, equivalently, that will minimize the negative value of twice the log likelihood of the data:

$$2\ln(L) = -2 \sum_{s=3,6,12,18} \sum_{t=target\ present\ or\ absent} \sum_{i=1}^{7} O(t, s)_i \log\left(x(t, s)_i\right). \quad (7)$$

The search for the minimizing parameters was conducted with an iterative Nedler-Mead (Nedler & Mead, 1965) method, implemented as the Nedler-Mead simplex routine ("fminsearch," available in Mathwork's MATLAB). To minimize the risk of getting ?11 caught in a local minimum, we iterated the simplex routine several times (typically between 4 to 10 times), with each new iteration starting with the parameters obtained from the previous iteration.[4]

Initially, we fit the models with no constraint on the mean identification time per display item (see Appendix B). This resulted in very low identification times for all participants in the conjunction task (range: 9–46 ms, mean 25 ms). Because these identification times seemed too low to be psychologically plausible, we decided to constrain the fits by imposing a minimal mean identification time of 50 ms. For both the spatial configuration and feature search tasks and for all participants, the mean identification time exceeded 50 ms; thus, there was no need to explicitly impose this constraint in the fits for these tasks. Importantly, imposing this constraint trades goodness of fit for psychological plausibility and interpretability.

As there are eight (2 Trial Type × 4 Set Sizes) combinations and seven categories for each combination, there is a total of $8 \times (7 - 1) = 48$ free data points to which the model was fit. The full model contains right free parameters, summarized in Table 1. Thus, the model is highly constrained.

Note that the model contains two additional parameters that were maintained at a fixed level and serve as scaling variables: The distractor identification weight $w_{\text{distractor}} = 1$ and the standard deviation of the identification noise $\sigma = 0.1$.

In addition to fitting the model to data of individual participants, we constructed an average observer and fit the model to her. The empirical quantiles for the average observer were obtained by averaging the quantiles of the individual participants. The empirical category frequencies $O(t, s)_i$ were obtained by summing the category frequencies across participants. We could thus proceed in fitting the model to the average observer using the methodology described above.

Parameter	Meaning
w_{target}	Target search weight
Δw_{quit}	Quit unit weight increment
v	Identification drift
θ	Identification threshold
m	% of motor errors
T_{er}^{yes}, T_{er}^{no}	Shift of residual RT for yes and no responses
γ	Rate of residual RT

Table 1. Free parameters of the model.

X-score transformations and distributions

Wolfe et al. (2010) introduced a nonparametric normalization procedure dubbed the "x-score transform," which linearly scales distributions via quantile alignment. Thus, scaling differences are removed, whereas nonlinear properties such as kurtosis and skew of distributions are preserved. They then used this procedure to study properties of RT distributions and to compare empirical versus model-predicted distributions. For the sake of comparability with their work, we utilized the x-scoring procedure as well.

The x-score distributions for the empirical data were calculated for each set size and trial type separately: For each individual, we calculated the 0.25 and 0.75 percentiles of the appropriate (trial type, set size) correct RT distribution. The correct RTs were then linearly transformed such that these percentiles were transformed to the (arbitrary) values −1 and 1, respectively, to obtain the x scores. Next, the x scores for the different individuals where pooled together and the distribution of the pooled x scores was estimated via kernel-estimation procedures (implemented by the "ksdensity" function in MATLAB).

To calculate x score for model predictions, we generated simulated data. For each set size, trial type, and individual, we simulated 100,000 model trials from the best-fitting parameters for that individual. We then proceeded in the calculations as above with the simulated data instead of the empirical data.

Model comparison

In addition to the full eight-parameter model, we also fitted certain submodels which were obtained by constraining parameters of the full model. First, to determine whether incorporating two mean residual RTs (one for each of the detection responses) improves fits, relative to having a single mean residual RT parameter, we compared the fit of the eight-parameter full model to the fit of a seven-parameter submodel that was obtained by constraining $T_{er}(yes) = T_{er}(no) \triangleq T_{er}$. In addition, in order to test whether having guidance towards targets improves the fits in the spatial configuration search task, we compared the full model to a seven-parameter submodel that was constrained by $w_{target} = 1$.

We conducted a further model comparison analysis, which was motivated by suggestions that in a conjunction search, participants segment the search display into two subsets (e.g., the red vs. the green items) and first examine items whose color matches the target's color on the previous trial (e.g., Kaptein, Theeuwes, & van der Heijden, 1995; Geyer, Müller, & Krummenacher, 2006). In CGS, such a strategy would be implemented by top-down assignment of higher salience weights to items of one color and lower salience weight to items of the other color. Specifically, we constructed an alternative conjunction model in which we set the weights of all green items to zero, reflecting the assumption that from display onset, green items are fully inhibited (recall that the target was always red vertical)—which is tantamount to excluding the green items from the search set. We dub this model variant the *half-set-size model*, to distinguish it from the *full-set-size model* in which all distractors have weight one. For the half-set-size variant, model fitting procedures were identical, save a single change: Set sizes were adjusted to half values of 2, 3, 6, and 9 instead of the actual set-sizes 3, 6, 12, and 18.

Model comparisons were performed with the Akaike information criterion (AIC; Akaike, 1973) and the Bayesian information criterion (BIC; Schwarz, 1978). In principle, both criteria implement a tradeoff between goodness of fit, gauged by $-2\ln(L)$ for the best fitting parameters, and model parsimony, measured by the number of free parameters. They thus penalize models for their complexity, so as to test whether the improvement in fits provided by the more complex models justifies their reduced parsimony. When selecting one from several alternative models, the model with the minimal criterion value is preferred. Both criteria implement different penalties on the number of parameters. AIC taxed each of the k parameter by two, whereas BIC by $\log(N)$, where N is the total number of observations per participant. Consequently, AIC is more liberal than BIC with respect to incorporating additional model complexity (as long as $\log(N) > 2$). Note that $k = 8$ or $k = 7$ for the full or the submodels, respectively, and $N \sim 4000$ per participant. For the half-set-size variant of the conjunction task, $k = 8$.

$$BIC = -2\ln(L) + \log(N) \cdot k \tag{8}$$

$$AIC = -2\ln(L) + 2k. \tag{9}$$

We can also utilize the AIC and BIC values obtained for the individual participants to calculate the AIC and BIC values for the entire group. In this approach, the individual fits are adjoined and perceived as a single fit for the entire data set across participants. Thus, the

above AIC and BIC formulas are used but with group negative twice log likelihood, group-k, and group-N values that are obtained by summing the values (of the negative twice log likelihood, k, and N, respectively) across individuals. An alternative approach for performing model comparisons at the group level is to calculate AIC and BIC for the average observer. In this approach, we use the negative twice log likelihood from the fit for the average observer, k is still eight (full model) or seven (submodels), and N is obtained by summing the Ns for the individual participants.

Results

First, we present the best-fitting model and several qualitative properties of the empirical data which are captured by the model performance. Second, we will quantitatively test the implicit assumption that a spatial configuration (2 vs. 5) task is prototypical for unguided search, in which the target salience is equal to that of nontargets. Third, we explore the difference in residual times between target and nontarget trials.

Qualitative properties of the model distributions

Wolfe et al. (2010) displayed the benchmark RT distribution data mainly in the form of density functions. Here, we display the same data along with the data produced by the best-fitting model in form of five quantiles, specifically, the 0.1, 0.3, 0.5, 0.7, and 0.9 quantiles. In Figures 5–7, the four set sizes (3, 6, 12, and 18 items) are depicted on the x axes. The quantile data for one condition are displayed as vertical "stacks." The data point in the middle is the median and thus comparable to displaying the central tendency, as is frequently done in studies reporting mean or median RT only. The distance between the lowest and highest data point (0.1 and 0.9 quantiles) indicates the dispersion of the data. For instance, it becomes apparent that for the spatial configuration task (Figure 5), both the central tendency and the spread increase as set size increases. Further, the shape of the distribution can be read from the distance of the 0.1 and the 0.9 quantiles from the median. In all conditions, it can be seen that the RT distributions have a positive (right) skew, that is, the slower tail is longer than the fast tail, with both the empirical and the model data.

Visual inspection reveals a very close similarity between the empirical and the model data, for both the RT distributions and the error rates (see Figures 5–7) The mean differences between the empirical and model RT quantiles across both target-present and target-absent trials and across all set sizes are 35 ms, 22 ms, and 5 ms for the spatial configuration, conjunction, and feature search tasks, respectively. Error rates across target presence and set size differ between the empirical data and model fits by, on average, 0.5%, 1.4%, and 0.3% for the spatial configuration, conjunction, and feature search tasks, respectively.

Second, the effects of set size are very minor in the feature search task. For the conjunction and, even more so, the spatial configuration search, the shapes of the RT distributions change substantially with set size. Specifically, the whole distributions are shifted to slower RTs with larger set sizes. That is, not only the mean RT but specifically also the head of the distribution, the 0.1 quantiles, become slower as set size increases. Further, the distributions increase in spread (apparent in the increasing distance between the 0.1 and 0.9 quantiles in Figures 5–7) as set size increases. With set size, "target-absent" RT distributions, too, become slower and wider and miss rates increase for the conjunction and spatial configuration tasks. Both in the model and the empirical data, miss rates for the inefficient tasks show a dependency on set size, whereas false-alarm rates stay relatively constant (the model constancy of false-alarm rates is an intrinsic property of our model implementation, as the sole cause of false alarms are motor errors, which are independent of set size; see above).

Third, the RT distributions of target present and absent responses strongly overlap. This becomes apparent in Figures 5–7 by comparing the "stacks" of target present and absent quantiles for one set size: the closer the 0.1 quantile of absent RTs is to that of present RTs, the larger the overlap.

Finally, we applied the x-score transform of RT distributions proposed by Wolfe et al. (2010). They presented this method of normalization as a means of comparing the shapes of different distributions with each other. The purpose of the x-score method is to remove scaling differences in distributions while preserving nonlinear properties such as kurtosis and skew (see statistical methods).[5] The x-transformed RT distributions for target-absent and target-present trials for the empirical and model data are presented in Figures 8, 9, and 10 for the spatial configuration, conjunction, and feature search tasks, respectively.

Note that the simulated data in the exhaustive model of Wolfe et al. (2010) qualitatively failed to catch the shape of empirical RT distributions (see their Figure 8), especially with respect to the skew of the distributions: For the spatial configuration task, the empirical data were strongly skewed, but the simulated data were symmetrical. In our model, the normalized x-score distributions are remarkably similar to the empirical data.

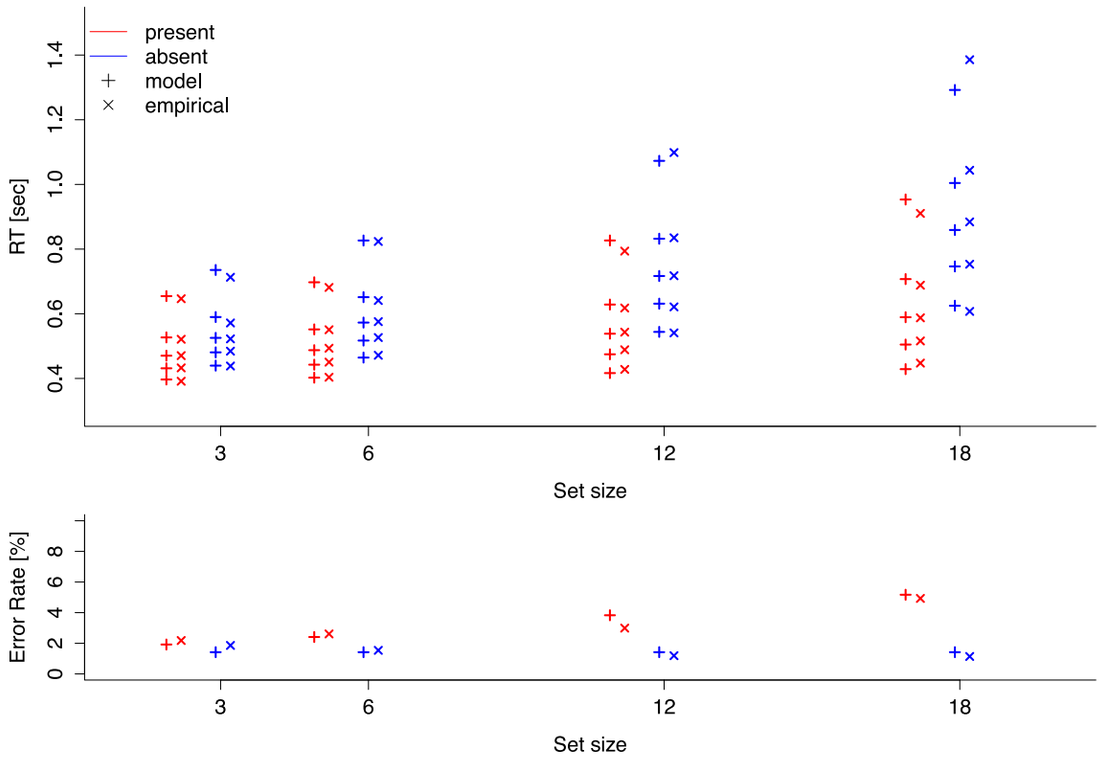

Figure 6. RT distribution and error data for the conjunction search task for the half-set-size variant. The RT distributions for the full set size variant are presented in Figure C1, Appendix C. The structure of the figure is equivalent to Figure 5.

Table 2 presents the best-fitting parameters for the average participant in the feature, conjunction, and spatial configuration tasks, respectively (please refer to Tables C1–C5 in Appendix C for the recovered parameters for the individual participants).

The effect of guidance

The investigation of search behavior is usually based on observable measures such as RTs or accuracy, from which conclusions are drawn with respect to the nature of the constituent mental processes. These variables are hidden and hence must be inferred from observable variables (e.g., slopes of RT/set size functions). A computational model provides a convenient framework for the investigation of hidden variables, such as guidance or search order. In fact, an important advantage of computational model lies in their formal explication of hidden cognitive processes. The recovered parameters of the model allow quantification of these processes, thus providing an inferential route to backstage cognitive operations.

For example, utilizing the recovered parameters permits tracking of how guidance by salience actually works in the model for the three different tasks. A very high salience (w_{target}) should result in pop out, that is, the target should almost always be the first item to be selected. For the conjunction task, it is usually assumed that there is some guidance, that is, the target has a salience somewhat greater than the distractors, making it more probable that the target is found in (relatively) early, rather than late, selections. The spatial configuration task is usually assumed to be prototypical of a nonguided search task in which the saliency map conveys no information with respect to which item is more likely than others to be the target. Thus, it is usually assumed that the search order is completely random and that the target is equally likely to be selected at each position in the search order. Supporting these expectations, Wilcoxon signed-rank tests confirmed that guidance was larger in the feature (pop out) relative to the conjunction task ($z = 3.633$, $p < 0.001$); in addition, it was larger in the conjunction than in the spatial configuration tasks ($z = 3.06$, $p = 0.002$). Note that in both comparisons, we used the full-set-size variant of the conjunction task.

Consider next the left panel of Figure 11, which displays the probability that the target is selected as the first, second, etc. item for the best-fitting parameters of the three tasks (to the average observer). The right panel displays the probability that a target-absent trial terminated after selecting (and rejecting) k display items. As can be seen, the assumptions with regard to

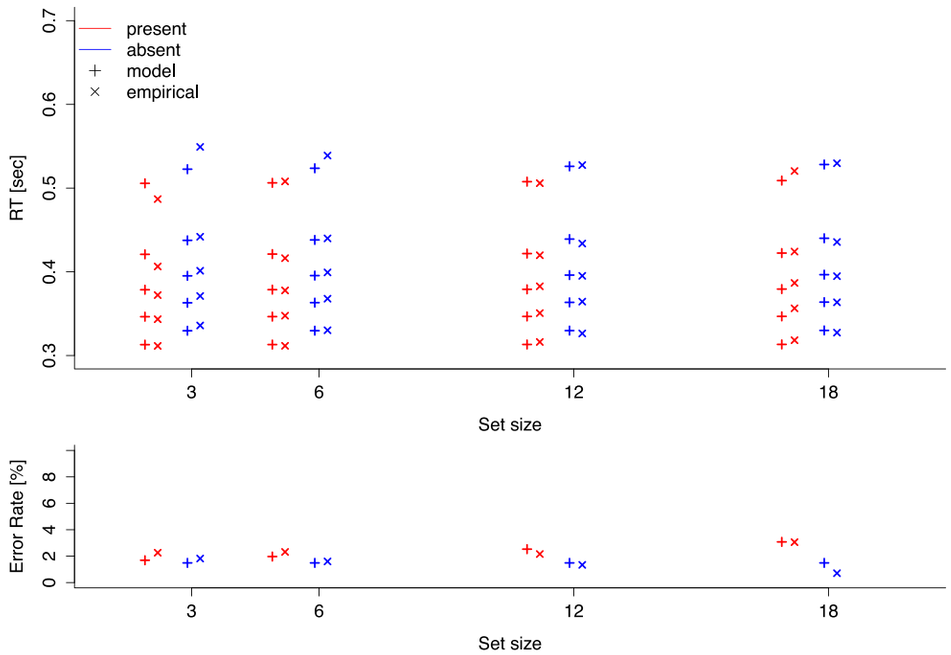

Figure 7. RT distribution and error data for the feature search task. The structure of the figure is equivalent to Figure 5.

substantial guidance in both the feature and the conjunction search task were supported by the model; surprisingly, though, there also seems to be some guidance for the spatial configuration search. Specifically, it is more likely to find the digit 2 in early selections than predicted by random search order; restated, the target is less likely to be selected at later positions than would be predicted by a random search order (the solid horizontal line represents the model prediction without guidance, i.e., with random search order).

This observation is also corroborated by the best-fitting parameters: the guidance parameters w_{target} for the spatial configuration task in Table 2 were greater than one for seven out of nine of the participants. A Wilcoxon signed-rank test for the guidance parameters revealed them to be significantly larger than one, $z = -2.073$, $p = 0.038$.[6]

We also conducted a model comparison analysis. For that purpose, we tested whether a model in which the salience of targets is enforced to be equal to that of the distractors ($w_{\text{target}} = 1$) does indeed produce a worse fit than the model reported thus far, in which w_{target} is a free parameter. We used the BIC and AIC criteria to quantify whether the improved fit (i.e., reduced negative log likelihood) due to the additional guidance parameter is justified by the reduced complexity. Table C1 presents the model comparison analysis. For eight and, respectively, seven out of nine of the individuals, the full model is preferable to the no-guidance submodel version, according to AI and BI criteria, respectively. The same conclusion holds for the average observer as well as for the group AIC and BIC values. On this basis, we conclude that a small guidance component towards the target is operative even in the spatial configuration search task, under the multiple session protocol that allows intensive training with the stimuli; it is possible that this component would disappear in single-session data.

Searching thorough half of the set in the conjunction task

Recall that we obtained two sets of fits for the conjunction tasks: The full-set-size variant in which all nontargets had equal weight of one, and the half-set-size variant in which the weight of all green items was set to zero and the weight of the remaining nontargets was set to one. Note that because the number of free parameters in both variants is identical (i.e., eight), comparing the AIC or the BIC values (or simply the negative twice log likelihood) is equivalent. In these analyses, the half-set-size variant was superior to the full-set-size variant for 7 out of 10 of the participants as well as for the average participant and the group (see Table C5). These results indicate that most participants are able to inhibit at least some of the distracters (e.g., the green distracters) using top-down modulation of attention towards the target and away from the nontarget color, thus effectively reducing the search set size.

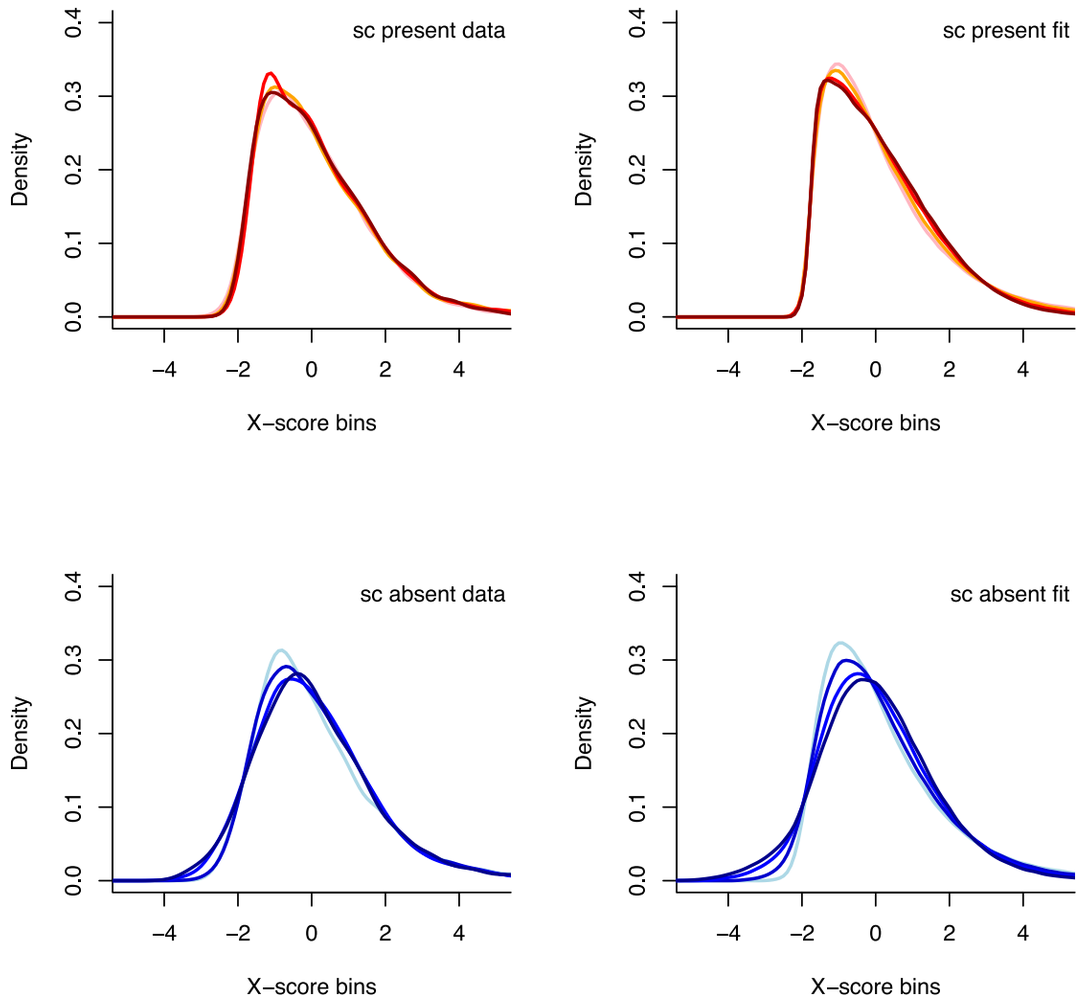

Figure 8. X-score distributions for the spatial configuration task correct RTs. The panels in the left column depict the empirical data, those in the right column the model predictions. The top row corresponds to target-present trials and the bottom row to target-absent trials. In each panel, the different lines represent different set sizes, with smaller set sizes corresponding to more desaturated colors. The x-score data is pooled across individuals.

The role of residual times for target-present versus target-absent trials

Wolfe et al. (2010) considered the possibility that part of the difference in RT distributions between target-present and target-absent trials might be attributable to differences in the residual time for "target-present" and "target-absent" responses. We decided to test this idea by comparing the full model with two shift parameters—$T^{yes}_{\min}$, $T^{no}_{\min}$—with a submodel obtained by constraining these shifts to be equal $T^{yes}_{\min} = T^{no}_{\min} \equiv T_{\min}$.

Analyzing the estimated residual RTs, a Wilcoxon signed-rank test showed that the shifts in residual times for "target-present" and "target-absent" responses do not differ across participants for the spatial search task, $z = -0.415$, $p = 0.678$. By contrast, for both the conjunction task ($z = -1.784$, one sided $p = 0.037$) and the feature task ($z = -2.173$, $p = 0.03$), the shift in residual RT is longer for the "target-absent" response, $T^{no}_{\min} > T^{yes}_{\min}$.

Additionally, a model comparison analysis revealed that for six (five) out of the nine individual participants in the spatial configuration task, AIC (BIC), preferred the full model over the submodel obtained by constraining the residual times to be equal. This was also the case for the entire group, but not for the average observer, whose fit preferred a single shift parameter. For the conjunction task, the picture was more consistent:T model with two shift parameters was preferred for 7 out of the 10 individual participants, as well as for the average observer and the group, according to both AI and BI criteria. Note that in this analysis we used the half-set-size variant (which yielded overall better fits than the full-set-size variant). Finally, for the feature task, the full model was preferred for eight (six) out of the nine individual participants as well

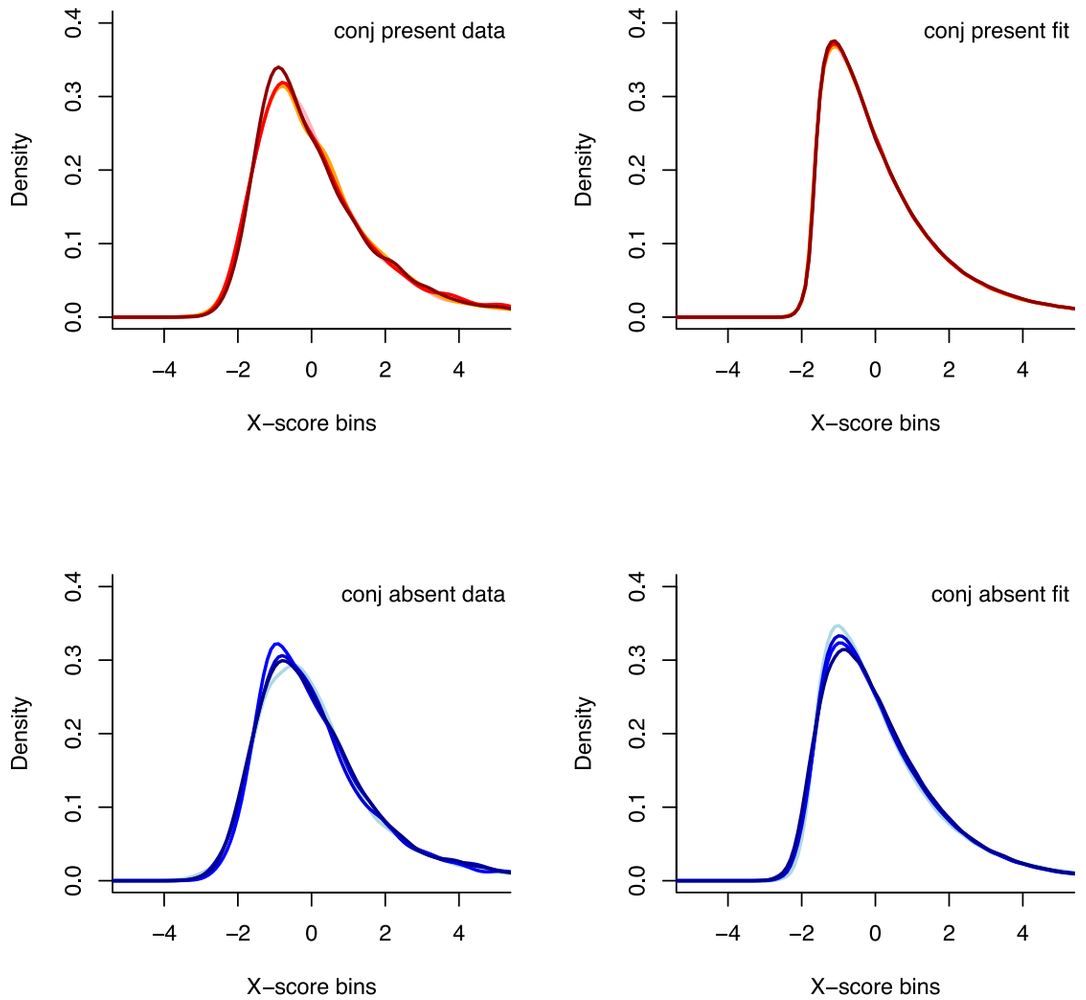

Figure 9. X-score distributions for the conjunction task correct RTs (full-set-size variant—the resulting x-score distributions for the half-set-size variant are indistinguishable from the present figure by naked eye). The organization of the graph is the same as that of Figure 8.

as by the average observer and the entire group, in terms of both AI and BI criteria.

In summary, in both the conjunction and the feature task, model comparison as well as estimated parameter comparison approaches converge on the conclusion that the shift in the residual time for "target-absent" responses is larger than that for "target-present" responses. For the spatial configuration task, the results are mixed: While for most of the participants and the entire group, the fit with two shift parameters is advantageous, a comparison of the fitted shift parameters reveals no difference between "target-present" and "target-absent" responses at the group level.

The larger shift in residual time for "target-absent" as compared to "target-present" responses is important for the model's ability to account for the RT data in the feature task. As discussed above, the large weights of both w_{target} and w_{quit} result in examination of a single item prior to response, on both target-present and target-absent trials, in a substantial majority of the trials (more than one display item is examined in only a small minority of trials). Thus, the major part of the difference in response RTs between the "target-present" and the somewhat slower "target-absent" responses is attributable to the larger residual shift for "target-absent" responses.

Consider now a subtle difference between target-present and target-absent trials in which a single item was checked prior to response. On such a target-present trial, the model terminates after a single iteration: The target is selected and identified, and a response ensues. On target-absent trials, by contrast, on the first iteration a distracter is selected, identified, and inhibited, and the weight of the quit unit is incremented. Only at the beginning of the second iteration, the quit unit is selected and the trial terminates. The upshot of this comparison is that target-absent trials require additional processes, such as inhibition, quit-unit

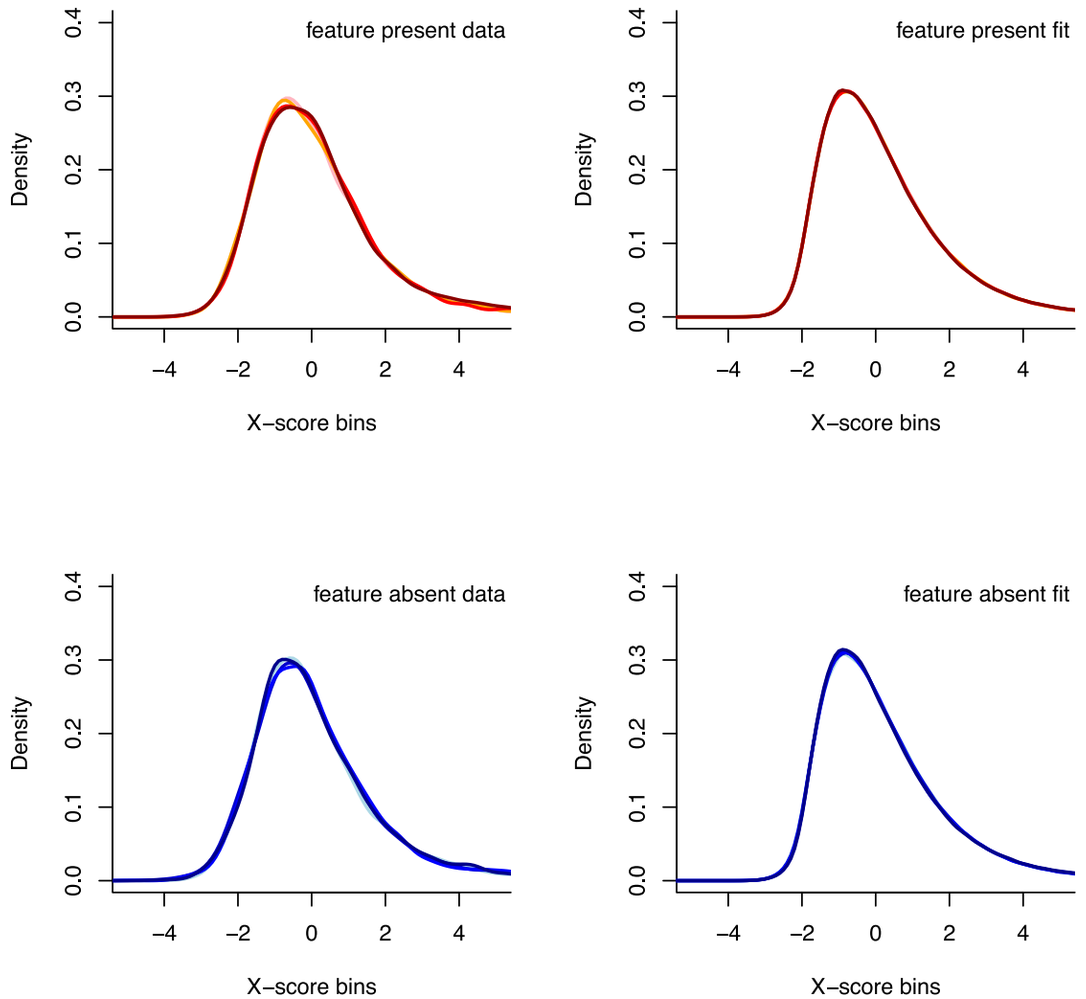

Figure 10. X-score distributions for the feature task correct RTs. The organization of the graph is the same as that of Figure 8.

activation, and quit-unit selection, which are not operative on target-present trials (with a single selection). More generally, for target-present and target-absent trials that both terminate after selection of k display items, the first involves k model iterations, whereas the latter involves $k + 1$ iterations. Recall that in the current model, we assumed that inhibition and quit-unit inhibition and selection occur instantaneously. Modeling durations for these processes might thus do away with the necessity of having two shift parameters with $T^{no}_{\min} > T^{yes}_{\min}$.

Discussion

For both humans and animals, searching the environment for a target is essential to their survival. Laboratory visual search tasks offer a controlled proxy for the more ecological situations, in which one forages for food, prey, or mates, and in which mistakes are costly, while speed is also of the essence. Over the last 30 years, intensive research has been conducted within this experimental paradigm, resulting in the development of a number of visual search models that account reasonably well for mean RT and error rate data. In a recent, seminal paper, Wolfe et al. (2010) reported benchmark RT distribution data for three classic visual search tasks (feature, conjunction, and spatial configuration searches), which challenge all visual search models, including their previous guided search model. Here, we presented a computational model of visual search—competitive guided search (CGS)—a variant of the guided search model that is able to account for the RT distributional data in all three tasks. We have shown that this model accounts simultaneously for both error rates and RT distributional properties (quantiles and the form of x-transformed density distribution; see Figures 8–10), for all set sizes and for both target-present and target-absent displays. In particular, the model is able to meet the challenge posed by Wolfe et al. (2010): low miss rates along with

Task	Variant	w_{target}	δ	θ	Δw_{quit}	T_{min}^{yes}	T_{min}^{no}	γ	m	θ/δ^a
2 vs. 5	Guidance	1.511	0.252	0.029	0.019	0.413	0.41	11.793	0.012	0.115
2 vs. 5	No guidance	1	0.23	0.02	0.01	0.41	0.44	12.48	0.01	0.103
Conjunction	Full set	4.958	0.286	0.014	0.162	0.368	0.367	14.810	0.011	0.05
Conjunction	Half set	1.57	0.25	0.01	0.02	0.36	0.39	15.32	0.01	0.056
Feature		599.8	0.457	0.063	869.7	0.239	0.256	48.104	0.015	0.138

Table 2. Best fitting parameters for the average participant in the three different tasks. For the spatial configuration (2 vs. 5) and the conjunction task, both model variants are presented. The recovered parameters for individual participants are presented in Appendix C. *Notes*: The right-most column is the estimation of the mean identification time per display item. It is obtained by dividing the identification threshold by the identification drift and is not another free model parameter.

highly overlapping target-present and target-absent RT distributions. Moreover, the model achieves this with a relatively small number of free parameters (eight free parameters and 48 free data points per task per participant). In addition, by fitting data to individual participants, we were able to estimate individual underlying search parameters (e.g., search guidance, identification drift rate, etc.).

We begin by discussing the novel components of the model that allow it to face the challenge posed by distributional RT data and we consider its implications for search processes. Next, we discuss the model's limitations and possible extensions and its relation to other search models.

Competitive guided search (CGS): What's new?

CGS is a two-stage model like guided search, consisting of a selection-identification cycle, where salience guides the order in which items are selected for identification. A critical component of the model is the termination mechanism, which tends to terminate the search, following additional nontarget identifications, with an increasing conditional probability (see also Cosineau & Shiffrin, 2004; Wolfe & Van Wert, 2010). Several aspects of our implementation are novel. First, GCS has a mechanistic model for the decision termination process which is integrated into the search architecture (see Zehetleitner et al., 2009). Second, CGS implements a novel implementation of the selection decision, which is based on Luce's choice axiom. Third, the model incorporates a shifted exponential distribution for the residual times (T_{er}), rather than the usual rectangular distribution (e.g., Ratcliff, 1978; but see Schwarz, 2001, and discussion in Wolfe et al., 2010). We will discuss each of these issues in turn, before briefly considering the model's implications for guidance in visual search.

Termination rule

CGS implements a competitive connection from the units of the salience map to the termination unit. That is, the probability of terminating search is reduced in proportion to the total (summed) activity on the salience map. In addition, we assumed that following nontarget identification, the salience of an item is inhibited (for the remainder of that trial). This has three important consequences. First, as more and more

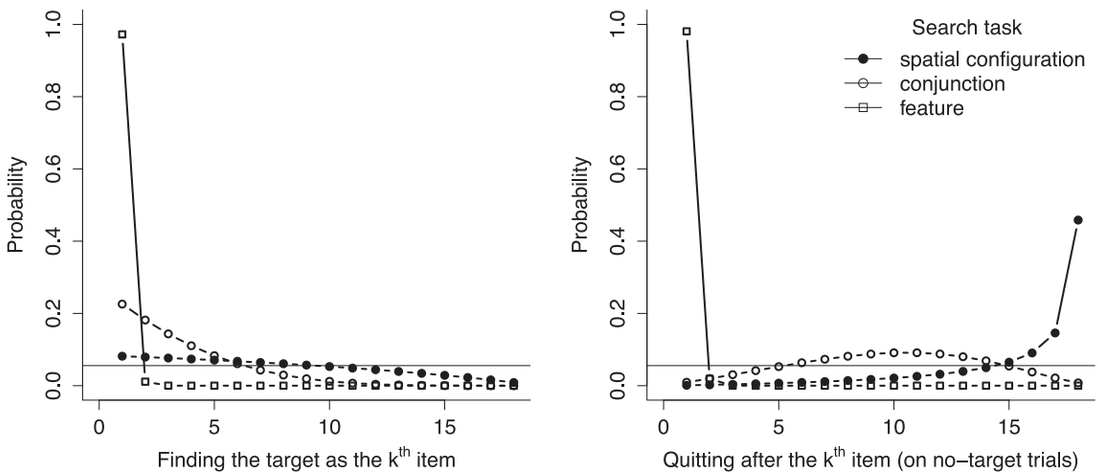

Figure 11. The left panel presents the probability of selecting the target (on target-present trials) at the *k*th position. The right panel presents the probability of terminating search after having checked *k* nontarget items.

items are selected and identified as nontargets, and consequentially inhibited, search termination probability increases. Eventually, when no items are left on the salience map for further inspection, the probability of terminating search becomes one. Second, as set size increases, the total activity on the salience map increases; as a consequence, search termination is delayed, resulting in slower target-absent RTs. In this way, CGS provides a mechanism that makes the termination probability dependent on the number of unexamined items in the display, based on interactions alone. Finally, when a salient target is present and $w_{target} > 1$, the quit unit faces a stronger competition for selection compared to when a target is absent. Consequently, the propensity to terminate the search after k nontargets are identified is stronger for target-absent than for target-present trials. This tendency benefits the searcher as it acts to reduce the prospect of missing the target (there is a weaker tendency for early termination when a target is present), while at the same time speeding up correct rejections (a stronger tendency for early termination when a target is absent).

A second component of the search termination mechanism is that the "weight" of the termination unit increases with each selection. This weight increment is governed in the model by the parameter Δw_{quit}, which determines the extent to which the propensity to terminate the trial increases. In the extreme, if Δw_{quit} is small enough, all items have to be examined before search is terminated automatically; by contrast, if Δw_{quit} is large, search is terminated after the first item has been identified. In our data fits, Δw_{quit} is estimated for each task separately, reflecting three different strategic settings for the three different tasks. In the feature search, a "target-absent" response is triggered if the first item turns out to be a nontarget. Only in a very small proportion of all trials, a second item is identified. For both the conjunction and the spatial configuration search, in about 50% of all trials, search is terminated after all items have been scrutinized. In the other 50%, fewer items are identified before quitting (see Figure 4, right panel). Here, the difference with exhaustive search, which would result in 100% of search terminations after checking all items, becomes apparent. This premature termination can be regarded as a type of informed guess (Cosineau & Shiffrin, 2004). As is apparent in Figure 4 (left panel), the higher the salience of the target, the more likely it is that the target is selected early rather than late. For example, for pop-out targets, the probability that a target is present if the first selected item was identified as a nontarget, is very close to zero: $P(\text{no target}|\text{the first item is a nontarget}) \sim 0$. Therefore, continuing to search is far from optimal, because this effort would hardly increase the evidence for whether a target is present or absent: The probability that the display does not contain a target is already close to one.

As noted above, Δw_{quit} is a strategic parameter under the observer's control and can be adjusted for each search task. Wolfe and Van Wert (2010) proposed that observers can strategically adjust their tendency to terminate the search even *within* a given task in response to target prevalence, that is, the proportion of target-present trials: They modeled the quitting signal as a stochastic accumulator that is incremented following each item that has been rejected as a nontarget. When the accumulator reaches a threshold, the search is terminated with a target-absent response. According to their account, when target prevalence is high (e.g., 98%), observers strategically increase this target-absent termination threshold. Consequently, more items are inspected, resulting in both slower termination times for target-absent trials and in decreased miss rates (see also the discussion of false-alarm rates, which increase with high target prevalence, below). In CGS, a reduction of Δw_{quit} leads to the inspection of more items and can thus produce slower target-absent RTs and lower miss rates (as target prevalence increases). Future research should test how well CGS accounts for the quantitative patterns and RT distributions in the Wolfe and Van Wert (2010) data set and compare it with their model.

The selection process

Selection in our model is based on salience: the more salient an item is, the higher its selection probability. Note that salience is potentially affected by both bottom-up and top-down factors. The selection process itself is stochastic and relies on Luce's Choice Axiom to translate the items' salience into selection probabilities (one possible neuronal mechanism of such translation is described below in the section Temporal dynamics of selection). This is somewhat different from the signal detection-based approach proposed by Wolfe et al. (1989) and Wolfe (1994). In signal detection-based approaches, saliency values for display items are sampled from Gaussian distributions, one distribution for targets and another for nontargets. The means of the target and nontarget distributions may differ (guidance toward targets is modeled by a larger mean for targets) and items are rank ordered for selection according to their saliency. Thus, while the signal detection approach assumes a noisy salience estimation, followed by a deterministic selection mechanism, CGS takes an opposite approach with respect to the sources of randomness in the salience estimation and selection processes: The salience values are deterministic, but the selection process itself is noisy. We are not aware of important functional differences between these mechanisms. We opted for utilizing the Luce

Choice Rule approach as it provided a coherent and common mechanism/framework for addressing both the item selection process and the *termination selection*, that is, the decision of whether or not to terminate the search process on a trial.

Residual times

The residual RTs in our models were assumed to be distributed with shifted exponential distributions. We also carried out simulations using the more standard rectangular distribution of residual times, which yielded reasonably good quantile-RT and error rate data fits. However, in some instances, we found the shape of the x-transformed RT density to show some deviations from the data—in particular, for small set sizes in the conjunction search task: For such displays, identification decisions are rather fast, and the variance in RT is dominated by the variance of the residual component, making the model distributions more rectangular than the empirical data. The use of exponentially shifted residual RTs solved this problem. The idea that the residual RTs are exponentially distributed was discussed by Palmer, Horowitz, Torralba, and Wolfe (2011b), following work by Schwarz (Schwarz 2001, 2002).

The role of shifted ex-Wald distribution in CGS

Fitting the Wolfe et al. (2010) data with several different RT distribution functions (Gamma, ex-Gaussian, ex-Wald, and Weibull), Palmer, Horowitz, Torralba, and Wolfe (2011) found that the ex-Wald provided some of the best fits. In order to avoid confusion, it is important to point out the different roles of the ex-Wald distribution there and in the present model. First, in Palmer et al. (2011), for each participant and condition of task, target presence, and set size, three parameters of the ex-Wald (and other) distributions were estimated. These fits resulted in a parsimonious description of the data: the 500 observed RTs per cell were reduced to three parameters of the ex-Wald distribution.

?12

In contrast, observed RTs in CGS are modeled as a mixture (across k) of sums of $k+1$ random variables: k item identifications plus a residual time. The mixture weights are determined by the likelihood that the target is found as the first, second, etc. and kth item (see Equations 32, 33, 36, and 37 in Appendix A for details). The Wald distribution was utilized to model one cycle of identification time. For each k, the time to identify the target is the sum of k independent Wald-distributed random variables—which, in turn, is Wald distributed. However, the overall distribution of the time spent on item identifications is a probability mixture, which is no longer Wald distributed.

Apart from these technical details, the main difference between the CGS and the Palmer et al. (2011) fitting approach is that CGS consists of a decision mechanism modeling chains of decisions (see Figure 2) which produces an RT distribution, whereas the distributions used in Palmer et al. are not the result of a decision mechanism. Moreover, for each observer and task, two parameters for the Wald distribution were estimated, whereas in Palmer et al. (2011), 24 ex-Wald parameters were fitted per task. Note that, overall, CGS required eight free parameters per task.

Guidance in spatial configuration?

Spatial configuration searches (for example, searching for the digit 2 amongst digits 5) are usually assumed to be free of influences from guidance implying that the order of item selection in such searches is virtually random (e.g., Braun & Sagi, 1990; Egeth & Dagenbach, 1991; Moore, Egeth, Berglan, & Luck, 1996; c.f. Wolfe, 1998). Somewhat surprisingly, however, the CGS fits for the spatial configuration task revealed the target to be more salient than the distracters. First, a model comparison analysis demonstrated that the model with a guidance parameter (a free target weight parameter) was superior to the nonguidance version (in which the weight of the target was constrained to one). Furthermore, the estimated guidance parameters of spatial configuration search were significantly above one (i.e., 1.5), which was smaller than the estimated guidance for the conjunction search, (i.e., 16.1). These finding support the notion that there is some, albeit small, degree of guidance toward targets in the spatial configuration task. One possibility is that guidance was acquired due to the substantial opportunities for practice in the multiple-sessions protocol. Future research should revisit this issue and in particular investigate whether the 2 among 5s task used in Wolfe et al. (2010) is indeed stereotypical for nonguided search. One possibility is that guidance will be distinguished by adopting a single-session protocol or by rotating each letter randomly to some degree. Alternatively, more complicated spatial configurations, such as circles from which radial spokes extrude at 4 of 12 possible positions (Cosineau & Shiffrin, 2004), may be better suited for the investigation of unguided search.

Limitations and possible extensions of the present model

A number of simplifications were made in the present model. (a) The target/nontarget identification was assumed to be perfect, that is, without allowing for

the possibility of misidentifications; (b) attentional selection is modeled as being instantaneous; (c) memory for rejected items in the model as perfect (items are never reselected); and (d) all of the distractors were modeled to have the same saliency weight. Importantly, these simplifications did not hamper the model's ability to account for RT distributions and error rates. Nonetheless, these assumptions could be relaxed in future studies. In fact, in our opinion, one virtue of GCS is that it provides a highly flexible framework that permits additional parameters to be incorporated naturally in order to implement further, or alternative, search mechanisms. Presumably, equipping the model with further mechanisms might take a toll with respect to complexity (increasing the number of model parameters) but may prove to be unavoidable in accounting for other data sets (depending on the task variations) and rewarding with respect to augmenting explanatory power. Below, we discuss the adopted simplifications and numerous potential relaxations/alternatives.

Misidentification

Error rates are very low in the benchmark data of Wolfe et al. (2010), as they are in a large body of behavioral visual search studies (usually below 5%). Our model can produce misses in two separate paths: First, it is possible that search is terminated before the target item has been selected; second, response keys may be confused (a motor error) even though the observer correctly identified a target. Motor errors are also the sole cause of false alarms in the present implementation.

A further source of errors stems from the identification decision: In the present model, when an item is selected, its identity is faultlessly determined. However, in reality, it is reasonable to assume that a target could be mistaken for a nontarget, and vice versa. These types of identification mistakes can be incorporated in future versions of CGS by modeling an imperfect identification process as a two-boundary diffusion (Ratcliff & McKoon, 2008), as a race model (Pike, 1973; Van Zandt, Colonius, & Proctor, 2000; Usher, Olami, & McClelland, 2002; Brown & Heathcote, 2008), or as a leaky-competing accumulator (Usher & McClelland, 2001). When we add the possibility of misidentification, CGS can easily account not only for the target-absent RT and miss data pattern of the prevalence effect, but also for false alarms. Specifically, it has been empirically shown that as target prevalence increases, false-alarm rates increase (while, concomitantly, miss rates decrease and target-absent RTs increase; Wolfe, Horowitz, & Kenner, 2005; Wolfe & Van Wert, 2010). Wolfe and Van Wert (2010) proposed that the prolonged target-absent RTs can be accounted for by a strategic shift in the termination criterion, leading to more items to be selected and checked for target identity. Checking more items increases mean target-absent RTs and reduces miss rates. As argued above, in the present model, this change in the termination criterion could be implemented as a decreased Δw_{quit}. Adding the possibility of misidentification to the present model can also, possibly, explain the empirical false-alarm rate pattern: When nontargets may be mistakenly identified as targets (generating false alarms), the overall false-alarm rate increases as more items are identified, because the probability that at least one of the identification errs builds up according to $1 - (1 - p)^k$, where p is the probability of a false alarm in one identification cycle and k is the number of cycles. However, apart from these qualitative considerations, it remains to be investigated how well the present model extended for misidentification can account for RT distribution and error data with target prevalence manipulations. Note that the Wolfe and Van Wert model was designed to account for RT means rather than distributions, as each selection and identification was counted simply as one cycle of a fixed duration. It remains to be seen whether their model could be elaborated to account for RT distributions as well and how it would compare to an extended CGS.

Temporal dynamics of selection

We have assumed that the selection process is instantaneous (obviously a simplification). Future versions of the model may include a separate RT distribution for the selection process, whose characteristic time may depend on item salience. In this case, the absolute target weights will affect RT distributions. In fact, GCS offers a natural way to model item selection times. Recall that the selection process in CGS can be conceived as a competition between items, where it is assumed that each item reflects a neuronal population whose firing is characterized by a Poisson process with rate w_i. If the selection process consists of eavesdropping on the first incoming neuronal pulse, then each item is selected with a probability that is proportional to its relative weight. More relevant here is that, these selection times will be distributed exponentially with a rate that is the sum of the weights of the (noninhibited) display items.

For feature (pop-out) search, there have been two recent proposals devised to understand the time until the first item is selected as a decision that can be modeled with an accumulator racing for a selection criterion (Purcell et al., 2010; Zehetleitner et al., 2013). According to these accounts, each item on the salience map is represented by an accumulator which races toward a selection criterion. There is also electrophysiological evidence from humans that the time until the

first item is selected depends on its salience: Töllner, Müller, and Zehetleitner (2012) have demonstrated that the N2pc (an electrophysiological marker of spatial-attention shifts; e.g., Eimer, 1996) has a latency that is the shorter the more salient a feature search target is. These findings refer to the timing of the first selection of an item. When more items are inspected, each additional selection may impose a different RT cost, as well.

Imperfect memory for visited locations

We have assumed that the search process has perfect memory. This was implemented by fully inhibiting an identified nontarget for the remainder of the trial, thus in effect isolating it from future search stages. In a number of difficult search experiments, it has been found that this is not necessarily the case (Horowitz & Wolfe, 1998; but see Peterson, Kramer, & Wang, 2001; von Mühlenen, Müller, & Müller, 2003). Specifically, observers sometimes forget which items they have already checked and revisit locations that they have already scrutinized. Future versions of CGS can implement a time-confined "inhibition of return" by setting the weight of a nontarget to zero (or higher value) and let it recover to its original value according to a certain time course. Alternatively, the model could adopt a *partial-inhibition* strategy, according to which the weight of an identified nontarget is inhibited with a certain percentage (an additional free parameter; e.g., 70% of the item weight is inhibited) for the rest of the trial. The gain might be reaped in the form of enabling a smooth interpolation (of memory levels) between the extremes of no memory whatsoever and full memory (as in the current model version) by letting the percentage of inhibited weight is zero and one, respectively. It remains for further investigations to test whether or not an extended partial-memory CGS model would improve the fits for the benchmark data and, if so, what the recovered estimates of memory efficiency would be.

Variability in salience weights

For the feature and spatial configuration search tasks, all nontargets had identical salience values. For the conjunction search task, we demonstrated the capacity of CGS to relax this restriction. In general, the assumption of fixed weights for all nontargets can easily be relaxed by endowing different distractors with different weights. For example, central distractors may be more salient that peripheral nontargets (Carrasco & Yeshurun, 1998), or distracters may be nonhomogeneous and, consequently, some nontargets may be more salient than others (e.g., Avraham et al., 2009). In fact, considering again the conjunction task, we note that the display consists of two different types of nontarget: red horizontal and green vertical bars (the target is a red vertical bar). In the full-set-size variant, we assumed an egalitarianism of nontarget saliency. This is tantamount to assuming that the red versus green color and the horizontal versus vertical orientation make identical contributions to item saliency. But this need not be the case and, in principle, each distractor type should receive its own weight (i.e., salience) parameter. A first promising step in this direction has been the fitting of the half-set-size variant, in which we assumed that participants can fully inhibit all the green items, effectively cutting the set size by half. For most participants, the fits of this variant were superior to the full-set-size fits. This demonstrates the benefit of studying variability in variance. We note that the assumption of perfect green inhibition may be over-simplistic or drastic, because it is reasonable that participants can inhibit the green items to some, yet not full, extent. A more general approach for studying the conjunction task could endow the two different distractors with two different (free) salience parameters and test whether this can further improve the fits. Furthermore, if the fits did improve, the recovered weight parameters might be instructive for determining which of the two distractors is more salient and to what extent. The upshot is that the half-set-size variant is only a preliminary step in studying salience variability and there is potential for further improvement.

Removing the constraint that all distractors are of the same salience weight could extend the ability of CGS to account for several additional finding. For example, consider a conjunction search task in which some of the distractors share no features with the target (e.g., green horizontal distracters when the target is red vertical). Such distracters are likely to receive lower weights than distracters that share target features, enhancing the relative saliency of the target and thus search efficiency (see, e.g., von Mühlenen & Müller, 2000).

Further, conjunction search is easier when the target is a red vertical among 20% red horizontal and 80% green vertical nontargets. In that case, according to guided search 2.0, the feature red would benefit from an intentional top-down boost, and green would be suppressed. In CGS, this would presumably imply a lower weight for all green items, and a higher weight for all red ones, turning the total balance of saliency in favor of the target.

When the target can be a singleton defined randomly in one of two possible dimensions (e.g., color or orientation), it has been demonstrated that correctly cueing or simply repeating the target dimension on the subsequent trial speeds up mean RTs (dimension weighting account, e.g., Found & Müller, 1996; Müller et al., 2003). In CGS this could be implemented by a

trial-by-trial adjustment of target weights depending on dimension cue or dimension repetition. Of course, these are speculations that do not replace the necessity of formal model fitting, evaluating how well CGS fares under such conditions and what the recovered saliency patterns are.

With respect to the discussion above, it is important to note that, qua model of search decisions, CGS estimates the salience (weight) parameters from empirical accuracy rates and RT distributions—it does not provide a mechanism to directly derive salience values from the search stimuli. However, instead of fitting saliency weights, these weights could be computed, and serve as the salience weights in CGS, by any of a number of salience computation algorithms (e.g., Itti & Koch, 2001; Gao & Vasconcelos, 2007; Bruce & Tsotsos, 2009; or others). In this way, a salience "module" could be plugged into CGS exempting it from the burden of fitting saliency parameters. Such an extension would benefit the model by virtue of extending its scope while constraining the salience (weight) values. Importantly, this would also serve to reduce the number of free model parameters, thus benefitting model fitting procedures with respect to robustness and efficiency.

To illustrate, consider search asymmetries. Search asymmetries refer to the fact that for some stimulus features, changing the role of target and nontarget affects search efficiency. Consider, for instance, a red disk among green disks, which is found efficiently; when swapping the target/nontarget features, presenting a green among red disks, search is still efficient. By contrast, when the target is a curved line among straight line segments, search is efficient—whereas the reverse is not true: a straight line among curved lines is found less efficiently. In general, if a salience module can account for saliency asymmetries in such situations, it is an appropriate candidate for being plugged into CGS.

Relation to parallel models

A number of parallel search models have also been proposed in the literature to account for visual search phenomena. Some of these models were shown to provide a good account for accuracy-based search variants with briefly presented stimuli (Palmer et al., 2000; Verghese, 2001), by relying on signal detection mechanisms and without assuming a sequence of attentional allocations.[7] A number of models were also proposed to account for search times in RT paradigms with unlimited viewing times (Ward & McClelland, 1989; Palmer & McLean, 1995; Thornton & Gilden, 2007) as well as for speed-accuracy trade-off paradigms with time-limited viewing (Dosher et al., 2004, 2010). These models typically require strategic parameters to account for set-size effects and have not yet been tested on RT distribution data. We have already embarked on another project in which we fit parallel models to the data. Some preliminary experimentation with an RT variant of the Verghese (2001) model—where at each time step an accumulator is incremented by a sample that is the maximum of n Gaussian samples for set size n, where targets have a greater mean than nontargets—suggests that the quality of the data fits in the 2 versus 5s search is worse than in CGS, even though more free parameters are used. We see this as a very preliminary investigation, as it is beyond the scope of this paper to assess the power of parallel models, which will require future dedicated research.

Here it should be noted that two different paradigms have been utilized in investigations of the temporal dynamics of visual search. The paradigm discussed so far and used by Wolfe et al. (2010) is a free-response paradigm in which the search display is presented until the observer responds. The observer is instructed to respond as fast and accurately as possible. An interrogation paradigm, where the search display is presented for only a brief period of time, say 50 ms, and observers have to submit their responses after an experimentally manipulated temporal interval, such as 100 ms, 150 ms, or 400 ms after display onset, has also been used (a speed-accuracy trade-off paradigm). For the latter paradigm, the resulting accuracies showed a set-size effect that was successfully fit by PPSM (Dosher et al., 2004, 2010). Their reason to prefer the interrogation paradigm with limited stimulus exposure was to exclude the possibility that eye movements dilute processing. It remains to be studied whether or not CGS can also account for data from a brief-presentation interrogation paradigm, thus providing a unifying model with respect to display presentation conditions.

The PPSM and parallel diffusion models feature a similar decision rule, but differ in the nature of the item-classification units. Instead of a diffusion implementation for item identification in which the boundary controls both accuracy and duration, in PPSM, the abstract decision units have one set of two parameters controlling the identification time and an independent set of two additional parameters controlling identification accuracy. Accounting for data of an inefficient difficult heterogeneous search task (Dosher, Han, & Lu, 2010) mandated set-size adjustment to the item classification accuracy parameters. PPSM has not been tested on RT distributions from a free-response paradigm and unlimited presentation time such as the data of Wolfe et al. (2010) and, in fact, it has not originally been developed for such purposes. However, PPSM—in its present form—might not be able to account for both substantial RT search slopes and low error rates in conjunction and spatial configuration

tasks, because set-size-dependent changes in the decision criterion captures error data but leaves decision times per unit unaffected. The core difficulty might be that the low false-alarm rate indicates that distracters are hardly ever classified as targets and, thus, that hits are almost exclusively based on correct identification of targets. However, as described above, in PPSM, target identification time is modeled independent of set size.

Apart from these considerations, parallel models meet some additional challenges. First, parallel models are unable to account for variations in the search task, where one has to report an attribute of the target, instead of its presence (Bravo & Nakayama, 1992). Second, eye movements and movements of covert attention are assumed to be intimately linked (Rizzolatti, Riggio, Dascola, & Umiltá, 1987; Hoffman & Subramaniam, 1995; Deubel & Schneider, 1996). That is, ideally, there should be one single model that accounts both for search with and without eye movements. Two-stage models are more prone to incorporate eye movements as an overt (instead of covert) shift of attention, whereas single-stage models would probably require a substantial change in architecture.

Conclusion

Competitive guided search is a novel model of visual search that meets the challenge of accounting for RT distribution in three benchmark search tasks. The model thus provides a unifying theoretical framework for prototypical search tasks that have been traditionally considered to be governed by qualitatively different mechanisms ranging from nonguided serial (spatial configuration task) to guided serial (conjunction task) and to parallel (feature) search. The model assets are its parsimony—it is based on a small set of parameters, its flexibility—it provides a theoretical framework that can readily be extended and elaborated to incorporate further or alternative search mechanism and its mathematical tractability.

Keywords: guided search, computational modeling, parallel versus serial search, sequential sampling, RT distributions, salience, search termination

Acknowledgments

We thank Wolfe et al. (2010) for providing access to their data set and to R. Shiffrin and C. Donkin for sharing a draft of their visual search manuscript. The research was funded by a grant 158/2011 from the German-Israel Foundation (M. Z., M. U., & H. J. M), by the I-CORE Program of the Planning and Budgeting Committee and The Israel Science Foundation (grant no. 51/11, M. U.), and DFG grant ZE 887/3-1 (M. Z. & H. J. M.).

* RM and MHZ contributed equally to this article.

Commercial relationships: none.
Corresponding authors: Michael Zehetleitner; Rani Moran.
Emails: mzehetleitner@psy.lmu.de; ranimora@post.tau.ac.il.
Addresses: Department of Psychology, Ludwig-Maximilians-Universität München, München, Germany; School of Psychological Sciences, Tel-Aviv University, Ramat Aviv, Israel.

Footnotes

[1]Note that it is also possible to implement imperfect memory for visited locations (Horowitz & Wolfe, 2003), though this has not been made use of in the present study.

[2]Fitting the model with uniform residual time, we found that while the quantile-RT data fits were not worse than those presented here, the more precise shape of the RT densities exhibited some discrepancies from the data (and relative to the shifted exponential residuals). This was especially evident for the smaller set sizes, for which the residual constitute a larger component of the RT. We thus opted for the shifted exponential residual times.

[3]We are grateful to Wolfe et al. (2010) for making their data set publicly available.

[4]Repeating the simplex routine several times increases the chances of escaping a local minimum because the step size with which the parameter space is sampled (the length of the edges of the simplex polygon) is generally decreasing as the simplex routine progresses. Rerunning the algorithm with the best-fitting parameters from the previous run thus starts with a larger step size.

[5]In Wolfe et al. (2010), the x-score transform is referenced as "Palmer, Horowitz, & Wolfe, submitted," because it was very similar to a prior method. The code and details are available from E. Palmer (personal communication: E. Palmer, Feb. 22, 2013). We apply this method here in order to allow for a direct comparison with Wolfe et al. (2010).

[6]The statistical test actually compared the logarithms of the guidance parameters to zero. Logarithms were taken so that the assumption of a symmetrical distribution [for log(guidance)] would be plausible.

[7]However, recently, visual search data with briefly presented inefficient search stimuli have been reported,

which are more consistent with models assuming serial shifts of attention than with a large class of single-stage parallel search models (Palmer, Fencsik, Flusberg, Horowitz, & Wolfe, 2011a).

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *International Symposium on Information Theory* ?16 (2nd edition, pp. 267–281). Tsahkadsor, Armenian.

Atkinson, R. C., Holmgren, J. E., & Juola, J. F. (1969). Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics, 6*(6), 321–326, doi:10.3758/BF03212784.

Avraham, T., Yeshurun, Y., & Lindenbaum, M. (2008). Predicting visual search performance by quantifying stimuli similarities. *Journal of Vision, 8*(4):9, 1–22, http://www.journalofvision.org/ ?17 content/8/4/9, doi:10.1167/8.4.9.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20*(3), 160–166, doi:10.1177/0963721411408885.

Braun, J., & Sagi, D. (1990). Vision outside the focus of attention. *Perception & Psychophysics*, *48*(1), 45–58.

Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Attention, Perception & Psychophysics*, *51*(5), 465–472.

Brown, S., & Heathcote, A. (2008). The simplest complete model of choice reaction time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153–178.

Bruce, N. D. B., & Tsotsos, J. K. (2009). Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision, 9*(3):5, 1–24, http://www.journalofvision.org/content/9/3/5, doi: 10.1167/9.3.5.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*(4), 523–547, doi:10.1037/0033-295X.97.4.523.

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*(2), 673–692, doi:10.1037/0096-1523.24.2.673.

Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology, 30*(1), 39–78.

Cousineau, D., & Shiffrin, R. M. (2004). Termination of a visual search with large display size effects. *Spatial Vision*, *17*(4-5), 4–5.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827–1837.

Donkin, C., & Shiffrin, R. M. (2011). Visual search as a combination of automatic and attentive processes. In L. Carlson, C. Hoelscher, & T. Shipley (Eds.), (pp. 2830–2835). Presented at the Proceedings of the 33rd Annual Conference of the Cognitive Science Society. Austin, TX. ?18

Dosher, B. A., Han, S., & Lu, Z.-L. (2004). Parallel processing in visual search asymmetry. *Journal of Experimental Psychology: Human Perception and Performance, 30*(1), 3–27, doi:10.1037/0096-1523.30.1.3.

Dosher, B. A., Han, S., & Lu, Z.-L. (2010). Information-limited parallel processing in difficult heterogeneous covert visual search. *Journal of Experimental Psychology. Human Perception and Performance, 36,* 1128–1144, doi:10.1037/a0020366.

Eckstein, M. P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science*, *9*(2), 111–118.

Egeth, H., & Dagenbach, D. (1991). Parallel versus serial processing in visual search: Further evidence from subadditive effects of visual quality. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 551–560.

Egeth, H., Jonides, J., & Wall, S. (1972). Parallel processing of multielement displays. *Cognitive Psychology, 3*(4), 674–698, doi:10.1016/0010-0285(72)90026-6.

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*(3), 225–234.

Found, A., & Müller, H. J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a "dimension-weighting" account. *Attention, Perception & Psychophysics*, *58*(1), 88–101.

Gao, D., Mahadevan, V., & Vasconcelos, N. (2008). On the plausibility of the discriminant center-surround hypothesis for visual saliency. *Journal of Vision, 8*(7):13, 1–18, http://www.journalofvision.org/content/8/7/13, doi:10.1167/8.7.13. ?19

Geyer, T., Müller, H. J., & Krummenacher, J. (2006). Cross-trial priming in visual search for singleton conjunction targets: Role of repeated target and

distractor features. *Perception & Psychophysics, 68*(5), 736–749, doi:10.3758/BF03193697.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9,* 394–401.

Hoffman, J. E. (1979). A two-stage model of visual search. *Perception & Psychophysics*, *25*(4), 319–327.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*(6), 787–795.

Horowitz, T. S. (2006). Revisiting the variable memory model of visual search. *Visual Cognition, 14*(4-8), 668–684, doi:10.1080/13506280500193958.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*(6693), 575–577, doi:10.1038/29068.

Horowitz, T. S., & Wolfe, J. M. (2003). Memory for rejected distractors in visual search? *Visual Cognition, 10*(3), 257–298, doi:10.1080/13506280143000005.

Kaptein, N. A., Theeuwes, J., & van der Heijden, A. H. C. (1995). Search for a conjunctively defined target can be selectively limited to a color-defined subset of elements. *Journal of Experimental Psychology Human Perception & Performance*, *21*(5), 1053–1069.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10-12), 1489–1506, doi:10.1016/S0042-6989(99)00163-7.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203.

Kendall, M. G., & Stuart, A. (1977). *The advanced theory of statistics* (Vol. 1). New York, NY:
?20 MacMillan.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4*(4), 219–227.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York, NY: Wiley.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. Oxford, UK: Oxford University Press.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition*,
?21 *22*(6), 657–672.

Moore, C. M., Egeth, H., Berglan, L. R., & Luck, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin & Review*, *3*(3), 360–365.

Navalpakkam, V., & Itti, L. (2006). Top-down attention selection is fine grained. *Journal of Vision, 6*(11):4, 1180–1193, http://www.journalofvision.org/content/6/11/4, doi:10.1167/6.11.4. ?22

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7,* 308–313. ?23

Palmer, E. M., Fencsik, D. E., Flusberg, S. J., Horowitz, T. S., & Wolfe, J. M. (2011a). Signal detection evidence for limited capacity in visual search. *Attention, Perception & Psychophysics*, *73*(8), 2413–2424.

Palmer, E. M., Horowitz, T. S., Torralba, A., & Wolfe, J. M. (2011b). What are the shapes of response time distributions in visual search? *Journal of Experimental Psychology: Human Perception and Performance, 37*(1), 58–71, doi:10.1037/a0020747.

Palmer, J., & McLean, J. (1995). Palmer: Imperfect, unlimited-capacity, parallel search. Presented at the Society for Mathematical Psychology. ?24

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*(10), 1227–1268.

Peterson, M. S., Kramer, A. F., & Wang, R. F. (2001). Visual search has memory. *Psychological Science*, *12*(4), 287–289.

Pike, R. (1973). Response latency models for signal detection. *Psychological Review, 80,* 53–68, doi:10.1037/h0033871.

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*(4), 1113–1143, doi:10.1037/a0020311.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia, 25*(1A), 31–40.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464.

Schwarz, W. (2001). The ex-Wald distribution as a

descriptive model of response times. *Behavior Research Methods, Instruments, & Computers, 33,* 457–469.

Schwarz, W. (2002). On the convolution of inverse Gaussian and exponential random variables. *Communications in Statistics: Theory & Methods, 31,* 2113–2121.

Shaw, M. L. (1982). Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology, 14*(3), 353–409, doi:10.1016/0010-0285(82)90014-7.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606.

Thornton, T. L., & Gilden, D. L. (2007). Parallel and serial processes in visual search. *Psychological Review, 114*(1), 71–103, doi:10.1037/0033-295X.114.1.71.

Töllner, T., Müller, H. J., & Zehetleitner, M. (2012). Top-down dimensional weight set determines the capture of visual attention: Evidence from the PCN component. *Cerebral Cortex, 22*(7), 1554–1563, doi:10.1093/cercor/bhr231.

Townsend, J. T., & Ashby, F. G. (1984). Measurement scales and statistics: The misconception misconceived. *Psychological Bulletin*, *96*(2), 394–401.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and co- active theories. *Journal of Mathematical Psychology, 39,* 321–359, doi:10.1006/jmps.1995.1033.

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

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*(3), 550.

Usher, M., Olami, Z., & McClelland, J. L. (2002). Hick's law in a stochastic race model with speed–accuracy tradeoff. *Journal of Mathematical Psychology*, *46*(6), 704–715.

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, *7*(2), 208–256.

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, *31*(4), 523–535.

von Mühlenen, A., & Müller, H. J. (2000). Perceptual integration of motion and form information: Evidence of parallel-continuous processing. *Perception & Psychophysics*, *62*(3), 517–531.

von Mühlenen, A., Müller, H. J., & Müller, D. (2003). Sit-and-wait strategies in dynamic visual search. *Psychological Science, 14*(4), 309–314, doi:10.1111/1467-9280.14441.

Ward, R., & McClelland, J. L. (1989). Conjunctive search for one and two identical targets. *Journal of Experimental Psychology: Human Perception and Performance, 15*(4), 664–672, doi:10.1037/0096-1523.15.4.664.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202–238.

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

Wolfe, J. M. (8AD). Visual search. In H. Pashler (Ed.), *Attention* (Vol. 1, pp. 13–73). London, UK: University College London Press.

Wolfe, J. M. (2007). Guided search 4.0. In W. D. Gray (Ed.), *Integrated models of cognitive systems (cognitive models and architectures)* (pp. 99–120). Oxford, UK: Integrated Models of Cognitive Systems.

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*(3), 419–433.

Wolfe, J. M., Palmer, E. M., & Horowitz, T. S. (2010). Reaction time distributions constrain models of visual search. *Vision Research, 50*(14), 1304–1311, doi:10.1016/j.visres.2009.11.002.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature*, *435*(7041), 439–440.

Wolfe, J. M., & Van Wert, M. J. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. *Current Biology, 20*(2), 121–124, doi:10.1016/j.cub.2009.11.066.

Zehetleitner, M., Goschy, H., & Müller, H. J. (2012). Top-down control of attention: It's gradual, practice-dependent, and hierarchically organized. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 941–957, doi:10.1037/a0027629.

Zehetleitner, M., Koch, A. I., Goschy, H., & Müller, H. J. (2013). Salience-based selection: Attentional capture by distractors less salient than the target. J. J. Geng, (Ed.), *PLoS ONE, 8*(1), e52595, doi:10.1371/journal.pone.0052595.s002.

Zehetleitner, M., & Müller, H. J. (2010). Salience from the decision perspective: You know where it is before you know it is there. *Journal of Vision*,

10(14):35, 1–16, http://www.journalofvision.org/content/10/14/35, doi:10.1167/10.14.35.

Zehetleitner, M., Müller, H. J., & Wolfe, J. M. (2009). Accumulation of salience: Modeling the effects of target distractor similarity in visual search. Vol. *38*, p. 19. Presented at the European Conference of Visual Perception.

Zehetleitner, M., Rangelov, D., & Müller, H. J. (2012). Partial repetition costs persist in nonsearch compound tasks: Evidence for multiple-weighting-systems hypothesis. *Attention, Perception & Psychophysics, 74*(5), 879–890, doi:10.3758/s13414-012-0287-y.

Queries for jovi-13-08-16

1. **Author: This article has been lightly edited for grammar, style, and usage. Please compare against your original document and make changes on these pages. Please limit your corrections to substantive changes that affect meaning. If no change is required in response to a question, please write "OK as set" in the margin. Copy editor**
2. **Author: Please provide email address and websites for all authors if available. Copy editor**
3. **Author: Make sure that all funding/financial support is listed in the Acknowledgments section. Copy editor**
4. **Author: Please define RT upon first use in the Abstract. Copy editor**
5. **Author: The Stuart & Kendall (1977) in-text callout does not have a corresponding reference. Copy editor**
6. **Author: The Cosineau and Shiffrin (2004) in-text callouts do not have a corresponding reference. Copy editor**
7. **Author: The Palmer (1995) in-text callout does not have a corresponding reference. Copy editor**
8. **Author: The Gao & Vasconcelos (2007) in-text callouts do not have a corresponding reference. Please check and correct. Copy editor**
9. **Author: The Navalpakkam & Itti (2007) in-text callout does not have a corresponding reference. Copy editor**
10. **Author: Please clarify the series at the end of the sentence "For a given parameter set and for each set size. . ." Are items missing? Copy editor**
11. **Author: The Nedler & Mead (1965) in-text callout does not have a corresponding reference. Copy editor**
12. **Author: For all Palmer et al. (2011) in-text callouts, please clarify whether you are referring to (2011a), (2011b), or both. Copy editor**
13. **Author: Please clarify the phrase "letting the percentage of inhibited weight is zero and one" in the sentence beginning "The gain might be reaped. . ." Copy editor**
14. **Author: There is no reference corresponding to the Avraham et al. 2009 callout. Please check and correct. Copy editor**
15. **Author: There is no corresponding reference for the Müller et al. (2003) in-text callout. Copy editor**
16. **Author: Please clarify "Tsahkadsor, Armenian" in the Akaike (1973) reference and add a location of publication and publisher to the reference. Copy editor**
17. **Author: The Avraham, Yeshurun, & Lindenbaum (2008) reference is not called out in the text. Please check and correct. Copy editor**
18. **Author: Please clarify the Donkin & Shiffrin (2011) reference. Was the paper presented at a conference or published in a book with editors? If both, which version of the paper did you use in your research? Please modify the reference accordingly. Copy editor**
19. **Author: The Gao, Mahadevan, & Vasconcelos (2008) reference is not called out in the text. Please check and correct. Copy editor**
20. **Author: The Kendall & Stuart (1977) reference does not have a corresponding in-text callout. Copy editor**
21. **Author: The Maljkovic & Nakayama (1994) reference does not have a corresponding in-text callout. Copy editor**
22. **Author: The Navalpakkam & Itti (2006) reference does not have a corresponding in-text callout. Copy editor**
23. **Author: The Nelder & Mead (1965) reference does not have a corresponding in-text callout. Copy editor**
24. **Author: Please double check the title of the Palmer & McLean (1995) reference. Additionally, please provide the month and location of the presentation. Copy editor**
25. **Author: The Theeuwes (1992) reference does not have a corresponding in-text callout. Copy editor**
26. **Author: Please provide the volume number of the journal for the Triesman & Gelade reference. Copy editor**
27. **Author: Please correct the year in the Wolfe (8AD) reference and ensure the reference has a corresponding in-text callout. Copy editor**
28. **Author: There is no corresponding in-text callout for the Zehetleitner, Goschy, & Müller (2012) reference. Please check and correct. Copy editor**
29. **Author: For the Zehetleitner, Müller, & Wolfe (2009) reference, please clarify if it was published in a journal or presented at a conference and modify the reference accordingly. Copy editor**
30. **Author: In Figure 2 caption, please verify the characters following p_{quit} in the second line. AP/typesetting**