Journal of Mathematical Psychology 115 (2023) 102766

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

A queueing model of visual search ☆

Yiqi Li [a,*], Martin Schlather [b], Edgar Erdfelder [c]

[a] University of Mannheim, Graduate School of Economic and Social Sciences, B6, 30-32, 68159 Mannheim, Germany
[b] University of Mannheim, Chair of Applied Stochastics, School of Business Informatics and Mathematics, B6, 26, 68159 Mannheim, Germany
[c] University of Mannheim, Cognition and Individual Differences Lab, School of Social Sciences, A5, 6, 68159 Mannheim, Germany

ARTICLE INFO

Article history:
Received 8 August 2022
Received in revised form 25 February 2023
Accepted 1 March 2023
Available online xxxx

Keywords:
Visual search
Queueing model
Serial/parallel processing
RT distribution modeling
Error modeling
Search termination

ABSTRACT

Understanding how attentional resources are deployed in visual processing is a fundamental and highly debated topic. As an alternative to theoretical models of visual search that propose sequences of separate serial or parallel stages of processing, we suggest a queueing processing structure that entails a serial transition between parallel processing stages. We develop a continuous-time queueing model for standard visual search tasks to formalize and implement this notion. Specified as a finite-time, single-line, multiserver queueing system, the model accounts for both accuracy and response time (RT) data in visual search on a distributional level. It assumes two stages of processing. Visual stimuli first go through a massively parallel preattentive stage of feature encoding. They wait if necessary and then enter a limited-capacity attentive stage serially where multiple processing channels ("servers") integrate features of several stimuli in parallel. A core feature of our model is the serial transition from the unlimited-capacity preattentive processing stage to the limited-capacity attentive processing stage. It enables asynchronous attentive processing of multiple stimuli in parallel and is more efficient than a simple chain of two successive, strictly parallel processing stages. The model accounts for response errors by means of two underlying mechanisms, namely, imperfect processing of the servers and, in addition, incomplete search adopted by the observer to maximize search efficiency under an accuracy constraint. For statistical inference, we develop a Monte-Carlo-based parameter estimation procedure, using maximum likelihood (ML) estimation for accuracy-related parameters and minimum distance (MD) estimation for RT-related parameters. We fit the model to two large empirical data sets from two types of visual search tasks. The model captures the accuracy rates almost perfectly and the observed RT distributions quite well, indicating a high explanatory power. The number of independent parallel processing channels that explain both data sets best was five. We also perform a Monte-Carlo model uncertainty analysis and show that the model with the correct number of parallel channels is selected for more than 90% of the simulated samples.

1. Introduction

A key question in attention research involves the allocation of cognitive processing resources to stimuli in the environment. A widely-used experimental paradigm to study visual attention is the standard visual search task (e.g., Geng & Behrmann, 2003; Wolfe, 2018), sometimes also referred to as the single-frame search task in the literature (e.g., Fisher, 1982). In a standard visual search experiment, participants are required to look for a single, predefined target among several isolated visual objects in the display. They are instructed to indicate as quickly and as accurately as possible whether the target is present or not. The display remains visible and static until a response is made. To infer how cognitive processing resources are allocated, the relation between stimulus inputs (i.e., task demands) and responses (i.e., processing performance) is investigated. The number of items in the display, called set size, is typically manipulated. The response time (RT) is recorded as a core performance measure. In the literature, the label "RT-method" (e.g., Wolfe, 2018) is also used for this paradigm. Accuracy rates are also recorded but usually treated as control variables to check for trade-offs between speed and accuracy. In past research, trials with incorrect responses were typically excluded, and the analysis has focused on the slope of the regression line of the mean RTs of correct responses on the set

☆ This work is based on the doctoral dissertation of the first author (Li, 2019). All data and R scripts for data fitting are available in the supplementary material at https://osf.io/k9573/.

* Corresponding author at: JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Room 409, 4/F, School of Public Health, Prince of Wales Hospital, Shatin, N.T., Hong Kong Special Administrative Region of China.
E-mail addresses: yiqi.li.mathpsych@gmail.com (Y. Li), schlather@math.uni-mannheim.de (M. Schlather), erdfelder@uni-mannheim.de (E. Erdfelder).

https://doi.org/10.1016/j.jmp.2023.102766

size, denoted as "the mean RT × set size slope" (cf. Wolfe, 2018). This slope characterizes the input–output relation observed under the target-present or target-absent condition. Once the processing structure is specified, explicit predictions on the slope can be derived. For instance, in serial exhaustive searches, the slope provides a rough estimate of the incremental time required to process each item added to the search set. Because different processing structures propose different explanations for the mean RT × set size slope, this slope has been considered one of the key *explananda* in the serial/parallel debate.

On the theoretical level, the serial/parallel debate is an enduring controversy in the visual search literature, focusing on the existence of an attentional "bottleneck", in particular among theoretical models that conceive visual processing as a sequence of serial and/or parallel stages. According to Moore and Wolfe (2001), serial and parallel accounts of visual search diverge on the question of whether accomplishing a visual search task involves any mandatory serial processing stage. Serial accounts assume at least one mandatory stage in visual search that requires an object-by-object processing by a single processing unit, whereas parallel accounts maintain that more than one item can be processed at any time and any level of processing, using parallel lines and parallel processing units. Accordingly, many prominent visual attention theories and models (e.g., Bundesen, 1990; Duncan & Humphreys, 1989; Treisman & Gelade, 1980; Wolfe, 1994), among others, those that are composed of a series of processing stages, are linked to either the serial or the parallel account. Although there are several sets of empirical findings in line with either account, the body of existing empirical research did not yield a conclusive result on this debate. Neither account can fully explain all empirical findings (Moore & Wolfe, 2001; Wolfe, 2021). For example, although the well-replicated finding of positive and steep mean RT × set size slopes (in the range of 20–60 ms/item, e.g., Treisman & Gelade, 1980; Wolfe, 1998) is consistent with the serial account, a limited-capacity parallel model with sequential sampling can also account for the same pattern (e.g., Townsend, 1971, 1990; Townsend & Nozawa, 1995). Resembling typical visual search findings, attentional dwell time, that is, the minimal time period during which attention is occupied once it is committed to a certain location or object (Duncan, Ward, & Shapiro, 1994), also does not discriminate between serial and parallel processing conclusively and complicates matters even further. Attentional dwell time estimates typically fall between 200 to 500 ms (e.g., Duncan et al., 1994; Moore, Egeth, Berglan, & Luck, 1996; Theeuwes, Godijn, & Pratt, 2004). Obviously, this finding is difficult to reconcile with the rapid rate of processing suggested by the mean RT × set size slope, assuming exhaustive, strictly serial processing.

Against this background, Moore and Wolfe (2001) explained how a hybrid structure that integrates serial and parallel features could resolve this paradox. As illustrated in the lower part of Fig. 1, a hybrid structure in the sense of Moore and Wolfe (2001) receives stimuli in a critical stage in series, yet several stimuli can be processed in parallel after entry. Clearly, this notion "is not a proposal for distinct serial and parallel stages of processing" (Moore & Wolfe, 2001, p. 191). However, such a hybrid structure resembles serial models because the deployment of attention can happen to only one stimulus at a time. This notwithstanding, it also resembles parallel models because the identification of several stimuli can take place simultaneously. These similarities might be seen as reasons for classifying models with such a structure as serial or parallel. However, we agree with Moore and Wolfe (2001) that they belong neither to the classically serial nor to the classically parallel camp. Importantly, such hybrid structures differ from serial structures by allowing subsequent stimuli to receive attention before the processing of precedent stimuli has finished. They also differ fundamentally from standard parallel structures because stimuli attain attentional resources one at a time, resulting in asynchronous beginnings of attentive processing. These properties also discriminate between hybrid models in the sense of Moore and Wolfe (2001) and alternative structures that have been called "hybrid" in a broader sense (cf. Townsend, 1971) to indicate chains of successive processing stages that include both purely serial and parallel ones (e.g., Bundesen, 1990; Hoffman, 1978; Treisman & Gelade, 1980). Moore and Wolfe (2001) argued that the hybrid structure they advocated (i.e., serial entry and subsequent parallel processing of stimuli) is compatible with existing empirical findings. For example, if stimuli enter an attentive processing stage one-by-one every 50 ms and stay there for about 300 ms, it is possible to observe a "processing rate" of 50 ms/item along with a dwell time around 300 ms.

This notion seems promising in view of the results reported by Wolfe (2021) and Fisher (1982); however, it has yet to be formalized and elaborated into a model that can make quantitative and testable predictions and is equipped with suitable parameter estimation procedures for rigorous empirical tests. In the latest version (6.0) of the Guided Search model, Wolfe (2021) incorporated an asynchronous diffusion component to represent a hybrid selection and recognition process. He demonstrated through simulation that the model is capable of mimicking basic patterns found in RT and accuracy data. There are other precedent theories and models that conceptualize cognitive processing as being both serial and parallel, for instance, the models proposed by Harris, Shaw, and Bates (1979); Fisher (1982, 1984); Miller (1993); Liu (1996, 2013); Liu, Feyen, and Tsimhoni (2006); and Wu and Liu (2008). Some of them have been stated as quantitative models, but the time-dependent limited-channel model by Fisher (1982) is the only one developed specifically for visual search. Fisher (1982) derived expected values of RTs assuming error-free processing. By minimizing the sum of squared differences between expected values and data, he estimated the model parameters from empirical mean RTs of correct responses and demonstrated a good model fit.

In this paper, we propose a mathematical model that implements the notion of a hybrid structure with both serial and parallel features — a continuous-time queueing model of visual search that accounts for RTs of both correct and incorrect responses on a distributional level. So far, the statistical analysis of RT in visual search has focused on the mean RT of correct responses. This does not provide a rigorous test of visual selective attention theories because many models can be shown to produce patterns of mean RTs as a function of set size that are similar to the empirically observed ones (see e.g., Palmer, Horowitz, Torralba, & Wolfe, 2011; Wolfe, Palmer, & Horowitz, 2010). Our model, in contrast, makes full use of the information contained in the RT data by modeling entire RT distributions. It thus enables a more comprehensive comparison of the theoretical RTs obtained from the model with empirical data, and provides a more rigorous test of the model's validity.

In addition to RT distributions, our model also accounts for accuracy data. Previous studies in visual search usually excluded incorrect trials from the analysis of RTs. The analysis of accuracy data in visual search has rarely been addressed, leaving sources of response errors and the consequences of excluding incorrect trials undiscovered. However, response errors in standard visual search appear systematic because particular patterns depending on set sizes have consistently been found (Wolfe et al., 2010), see Section 2.2 for details. To our knowledge, only a few papers addressed error rates of responses in standard visual search tasks specifically (e.g., Moran, Zehetleitner, Liesefeld, Müller, & Usher, 2016; Moran, Zehetleitner, Müller, & Usher, 2013; Zenger & Fahle,

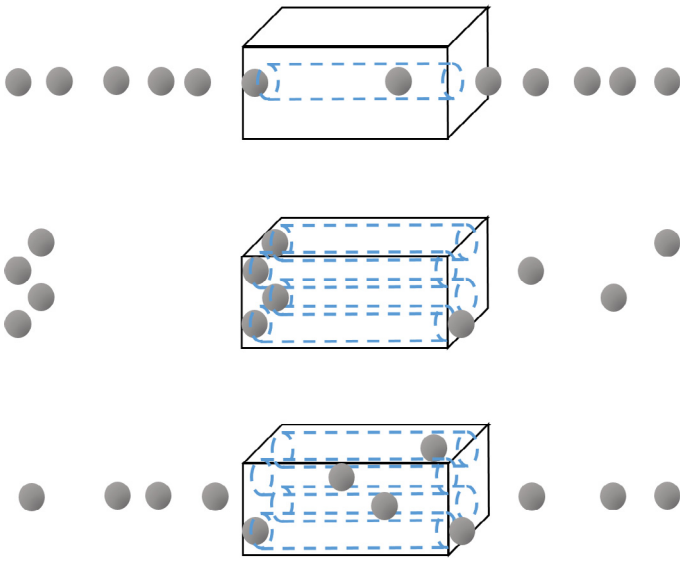


Fig. 1. Illustration of a strictly serial system (top), a limited-capacity parallel system with sequential sampling (middle), and a hybrid structure with both serial and parallel properties (bottom). Dotted tubes within cuboids illustrate parallel processing channels.

1997). Further authors modeled the RTs of incorrect responses in visual search but not the error rates themselves (e.g., Cousineau & Shiffrin, 2004; Donkin & Shiffrin, 2011). Still others adapted quantitative models of speed-accuracy trade-off from other fields, for instance, Ratcliff's (1978) diffusion model, to visual search data with a single, fixed set size (e.g., Corbett & Smith, 2020; Thornton & Gilden, 2007). Despite these achievements, fully explaining observed error patterns in a psychologically meaningful way remains challenging. We tackle this problem by modeling error probabilities using two underlying principles, namely, premature search termination and misidentification. Embedded in the framework of a queueing process, these two principles also constitute the final pieces of the puzzle required for the derivation of the joint distribution of responses and RTs.

On a more general level, our goal is to develop an alternative formal model framework for standard visual search tasks that goes beyond the idea of a simple chain of serial or parallel processing stages. This new model framework is intended to account for the joint distributions of responses and RTs for different set sizes, thereby extending the empirical scope of existing models. In substantive terms, the proposed framework formalizes a generalized version of the feature integration theory of visual attention (Treisman & Gelade, 1980) by connecting parallel processing within the feature-encoding and the subsequent feature-integration stage via a single-line queueing mechanism, thereby enriching the explanatory power of a classical theory in the field.

This paper is structured as follows. In the subsequent section, we formalize the proposed theory of visual search as a stochastic process model using standard terms of queueing theory. From the mathematical model, we derive formulas for the joint distributions of responses and RTs. To enable data fitting, we then develop appropriate parameter estimation procedures in Section 3. In Section 4, we evaluate the model empirically by fitting it to two large prototypical data sets collected and published by Wolfe et al. (2010). Finally, we perform a model uncertainty analysis using bootstrap methods in Section 5 to test whether model-based analyses can identify the number of servers in a single-line queueing model reliably. Appendix A delivers detailed derivations and comments on the subsequent equations.

2. Mathematical framework

In a queueing system, customers arrive as discrete units for service, wait in line if necessary and leave after being served. Queueing systems are widely used and are fully described by six characteristics (see e.g., Bhat, 2015; Gross, Shortle, Thompson, & Harris, 2008; Stewart, 2009): the interarrival times A between successive arrivals, the service times S of servers, the number of parallel servers c, the waiting room capacity, the total number of customers k, $k \in \mathbb{N} \cup \{\infty\}$, and the queue discipline, for example, the first-come-first-served (FCFS) discipline.

In our model, a continuous-time queueing process represents the course of visual search on a single trial. The k visual stimuli in the display correspond to the customers who compete for limited attentional resources. The queueing system determines when and which stimulus gets the service of being identified as target or non-target. Every stimulus is allowed to enter the queueing system once. This assumption of a non-recurrent customer source corresponds to perfect inhibition of return in visual search (e.g., Klein, 1988, but see also Horowitz & Wolfe, 1998; Wolfe, 2012 for different viewpoints). Furthermore, our model deliberately ignores processes that might be involved in visual search experiments, but lack a visual processing component, such as pure guessing or motor execution errors. The reasons for this choice will be elaborated in Section 2.2.

2.1. Model specification

The processing of a single stimulus is described assuming two stages, namely, preattentive processing and attentive processing (Treisman, 1986; Treisman & Gelade, 1980). In the preattentive stage, features of visual objects are extracted and encoded. Feature encoding is hypothesized to occur prior to attentive processing, automatically without attentional control, and simultaneously across the entire visual field, that is, in a massively parallel manner. In our model, the arrival process in the queue corresponds to feature encoding. The beginning of the arrival process, t_0, is defined as the time of stimulus onset. At this point, there is no customer in the queueing system, as illustrated in Fig. 2.

An item "arrives" in the queue at the precise instant that all features necessary to construct a representation of it have

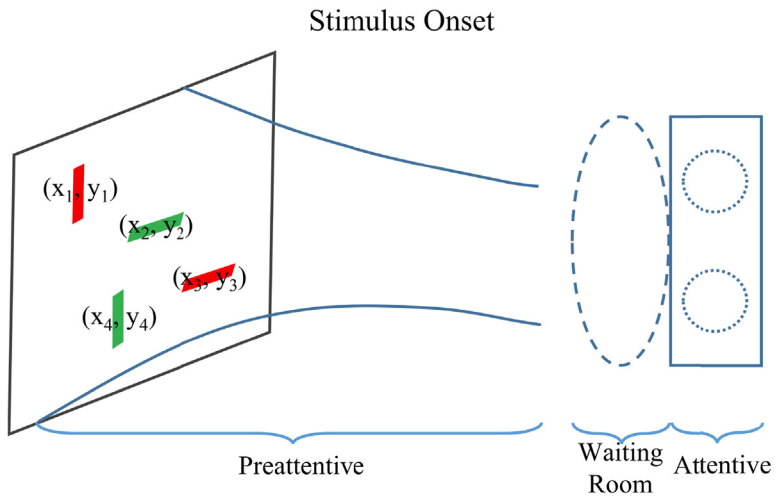


Fig. 2. Visual information in a paradigmatic queueing system at stimulus onset t_0. A total of $k = 4$ visual stimuli are denoted unambiguously by the coordinates (x_i, y_i) of their locations in the display, respectively, for $i = 1, \ldots, k$. The stimuli are characterized by two features each, orientation (horizontal vs. vertical) and color (red vs. green), but have not been processed so far. The dashed ellipse illustrates the waiting room for the queue, the rectangle the server system, and the dashed circles the $c = 2$ parallel processing channels (i.e., servers) in the server system.

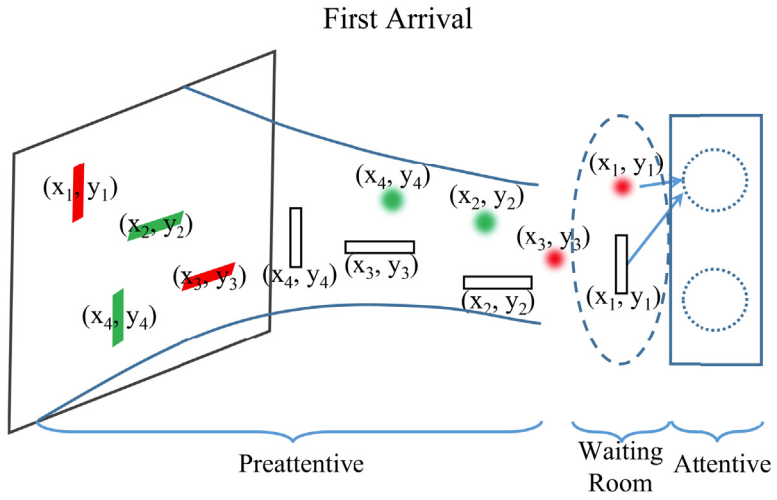


Fig. 3. Visual information in the queueing system when the first item arrives in the queue. The first stimulus with the coordinate (x_1, y_1) is ready for being served since both features (i.e., vertical, red) have been encoded. In contrast, none of the other stimuli has arrived because only one feature (i.e., red) has been encoded for stimulus (x_3, y_3), which stays in front of the waiting room, and no feature at all for the remaining two stimuli.

been encoded (cf. Treisman, 1986, 1988). This is illustrated in Fig. 3. Let $n_A(t)$ denote the cumulative number of arrivals up to time t. Then, the stochastic process n_A takes values in $\{0, \ldots, k\}$, has a jump at each arrival epoch, and remains constant until the next arrival occurs. It can be assumed to be càdlàg (right-continuous with left limits; see e.g., Protter, 2005). Regarding t_0 as the moment of "arrival" for $i = 0$, the interarrival time A_i between the $(i-1)$th and ith arrival for $i = 1, \ldots, k$ is assumed to be exponentially distributed with mean $1/(\lambda(k-(i-1)))$, where λ is a positive constant. Furthermore, A_i is assumed to be conditionally independent of A_{i-1} given $n_A(t) = i-1$. This implies an effective arrival rate $\lambda(k - n_A(t))$ at time t. Note that such a state-dependent arrival pattern mirrors the massively parallel processing in the preattentive stage because, at any instant, the number of arrivals per unit time is proportional to the number of stimuli that are still undergoing feature encoding, that is, stimuli that have not yet arrived in the queue. We further assume that any arrival order is equiprobable.

If there are no attentional resources (i.e., no free servers) available upon arrival, an item has to wait in the waiting room, as illustrated in Fig. 4. The waiting room is assumed to have unlimited capacity and consists of a single queueing line, which corresponds to an attentional bottleneck. Once processing capacity becomes available again (i.e., a server is free for service), the first item in the line is assigned to this server (i.e., FCFS). At this moment, the item advances to the attentive stage of processing.

According to Treisman (1986) and Treisman and Gelade (1980), the attentive stage selects features present at particular locations and integrates them to form coherent object representations. It is hypothesized to be more sophisticated, to require focal attention and to have limited processing capacity. In our model, the service process corresponds to the identification of an item by a server based on the integration of its features. Our model assumes c parallel, independent servers, with $c \in \mathbb{N}$ fixed but unknown. In visual search, the number of parallel processing units that require attention is considered limited by the capacity of the visual short term memory (Bundesen, 1990; Duncan & Humphreys, 1989). Thus, we restrict the range of c to $1, \ldots, 10$. At any server, the service time S_i of the ith identification, $i = 1, \ldots, k$, is assumed to be i.i.d. exponentially distributed with mean μ^{-1}.

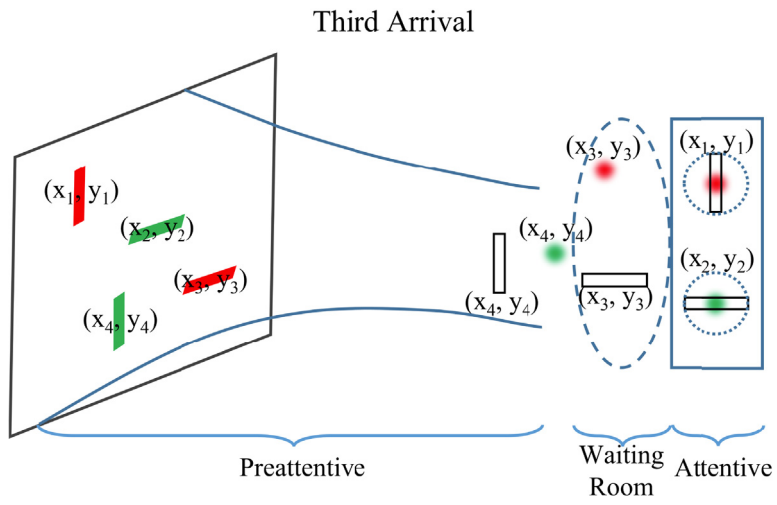


Fig. 4. Visual information in the queueing system at the moment when the third item (x_3, y_3) arrives in the waiting room of the queue with both features encoded. Because the two available servers are occupied by the first two stimuli (x_1, y_1) and (x_2, y_2), the third stimulus has to wait. Only the color but not the orientation of the fourth stimulus (x_4, y_4) has been encoded at this point, thus it is not ready for entering the queue.

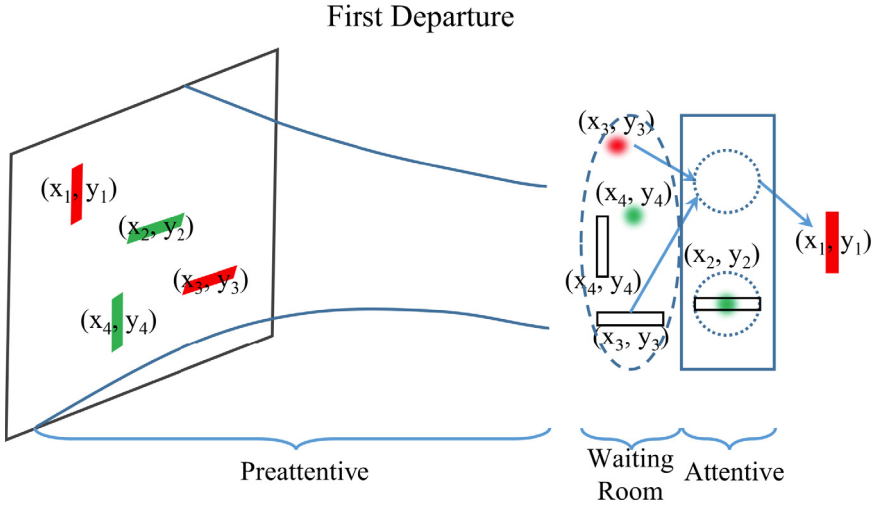


Fig. 5. Visual information in the queueing system at the first departure epoch, that is, when the first item (x_1, y_1) has been identified by a server and leaves the system. Since this server is free now, the same server can now process the next item in the queue, (x_3, y_3). Stimulus (x_4, y_4) has to stay in the waiting room with both features encoded because the second server is still occupied by stimulus (x_2, y_2).

Once the identification is completed, the item is released from the queueing system, as illustrated in Fig. 5. This instant is then a departure epoch. Although services start in the same order as arrivals because of the FCFS queue discipline, items may depart in a different order due to service time variations at parallel servers (cf. Wolfe, 2021). The assumptions on the service pattern imply that the effective departure rate of the whole system at time t is given by $\mu \min\{Q(t), c\}$, where $Q(t)$ denotes the number of customers in the system, that is, at a server or in the waiting room, at time t. The stochastic process Q is also càdlàg.

The above assumptions fully describe the probabilities of arrivals and departures at any moment t in continuous time, that is, the transitions between different system states $(n_A(t), Q(t))$. Consequently, the time course of any event defined in terms of the queueing model, in particular the departure epochs, can be derived based on the parameter vector (λ, μ, c).

In sum, the queueing model of visual search proposes that each item that has been identified in visual search experiments typically passes through two stages, namely, (1) arrival in the queue, which is achieved by the completion of feature encoding, and (2) processing by one of the c servers of the system to form an integral object representation. In between, there might or might not be an intermediate state, where the processing is suspended in the waiting room. The first stage corresponds to the preattentive stage and the second stage to the attentive stage of processing. Cognitive processing occurs only in these two stages and is characterized by unlimited-capacity parallel processing (i.e., automatic feature encoding) in the first stage and limited-capacity parallel processing (i.e., attention-based stimulus identification) in the second stage. The waiting room, in contrast, does not involve cognitive processing, although staying there takes some time in case all servers are occupied. It is best understood as a serial transition from the preattentive to the attentive stage of stimulus identification. Due to the FCFS discipline, the time any item has spent in the waiting room will be reflected in the departure times of those items that arrive after it. Thus, the waiting time is considered part of the visual processing time.

The theoretical construct of the waiting room is a core feature of our model. The serial transition from the preattentive to the attentive stage is not a designed property of this waiting room,

but rather a natural consequence of arrival and subsequent service processes in continuous time. To see this, suppose there are two or more stimuli in the waiting room. This implies that all servers must be currently occupied. The next transition of an item from the waiting room to the attentive stage of processing will happen at the same moment an item departs from the attentive stage so that the respective server becomes available again. Since the probability of two or more departures occurring at exactly the same instant is zero, no more than one server will become available at a time.

Note that this one-by-one transition is more efficient than a chain of two successive strictly parallel processing stages. This is because standard synchronous parallel processing stages (cf. Townsend, 1971) must necessarily postpone the start of attentive processing for those stimuli that arrived earlier in the queue until the required number of servers becomes available. This refers, for example, to visual search models that postulate attentional selection as the result of parallel competition among stimulus representations (e.g., Bundesen, 1990; Duncan & Humphreys, 1989). Such models necessitate a simultaneous start of parallel processing for the stimulus representations that take part in the competition.

2.2. System responses and theoretical RT

The courses of visual search on the trials of a standard visual search experiment are modeled as independent realizations of the same queueing process with the transition structure characterized by (λ, μ, c), as described above. For each experimental trial, we denote the set size by k and the objective status of the target by q, with $q = 1$ for presence and $q = 0$ for absence of the target in the stimulus display. The RTs, $X_{k,q}$, and responses, $Y_{k,q}$, observed on different trials under the same experimental condition (k, q) are considered i.i.d. The experimental manipulation of the factors k and q is assumed to impact the outcomes of identification at individual servers, which then jointly determine the stopping time of the queueing process and the final system response. Thus, to model RT and response jointly as a random vector $(X_{k,q}, Y_{k,q})$, we need to specify how a queue terminates under each level combination (k, q).

Assumption A1 (*Termination Rules*)**.** The entire queueing process is assumed to terminate in finite time under two mutually exclusive conditions:

C1. The system finishes serving an item classified as a target. Since target detection is conceived as a self-terminating process, this instantly triggers a target-present (TP) response.
C2. The system finishes serving l items $(0 \leq l \leq k)$ and classifies them all as distractors. This triggers a target-absent (TA) response.

The termination criterion $l = l(k)$ is a function of set size k and is determined based on the maximization of search efficiency under an accuracy constraint (see Section 2.3 for details). We denote the occurrence of a TP and a TA response under condition (k, q) by $Y_{k,q} = 1$ and $Y_{k,q} = 0$, respectively. Note that Assumption A1 implies an asymmetry between a TA and a TP response. Given insufficient evidence for the presence of the target, TA is always the "default" response. This differs conceptually from models that attribute bias between positive and negative responses completely to asymmetric boundaries for gradual, noisy evidence-accumulation processes (e.g., Ratcliff, 1985).

Assumption A1 implies that an incorrect response can result from (a) omission due to early search termination (i.e., $l < k$) or (b) an erroneous service outcome, that is, the misidentification of an item. An omission occurs on a target-present trial when the target is not included in the first l items searched. If, in addition, these l non-target items are processed in an error-free manner, an incorrect TA response will be triggered automatically, without any intrinsic errors in the attentive stage. The subsequent assumption (Assumption A2) specifies how a misidentification at a server (i.e., a false positive or a false negative classification) is invoked. For the service outcome of any item i on a trial, we use the term "$\mathcal{T}_i = 1$" | $\mathcal{T}_i = 0$ to indicate a false positive classification of item i, i.e., mistaking a distractor for the target, and "$\mathcal{T}_i = 0$" | $\mathcal{T}_i = 1$ to indicate a false negative classification, i.e., mistaking the target for a distractor. Note that "$\mathcal{T}_i = 1$" | $\mathcal{T}_i = 0$ for any item i in a target-absent display necessarily results in a false alarm $Y_{k,0} = 1$ because the search process is assumed to be self-terminating. Conversely, "$\mathcal{T}_i = 0$" | $\mathcal{T}_i = 1$ does not imply a miss $Y_{k,1} = 0$ in target-present displays because another item in the queueing system might still be classified as a target by any of the c servers.

Assumption A2 (*Misidentification Probabilities*)**.** For any fixed set size k, "$\mathcal{T}_i = 1$" | $\mathcal{T}_i = 0$ and "$\mathcal{T}_i = 0$" | $\mathcal{T}_i = 1$ emerge at each server independently with probabilities p_1 and p_2, respectively, for each item i in the system.

The probabilities p_1 and p_2 are functions of set size k and are determined by a threshold adaptation. Adjourning further specifications of the dependencies of l, p_1, and p_2 on set size k to Sections 2.3 and 2.4, we can now derive the error probabilities of the responses from Assumptions A1 and A2 (cf. Appendix A.1 for comments on the derivation):

$$\pi_1(k; l, p_1, p_2) := \mathbb{P}(Y_{k,0} = 1) = 1 - (1 - p_1)^l; \tag{1}$$

$$\pi_2(k; l, p_1, p_2) := \mathbb{P}(Y_{k,1} = 0) = (1 - p_1)^l \frac{k - l}{k} + (1 - p_1)^{l-1} p_2 \frac{l}{k}. \tag{2}$$

We consider omission and misidentification as sources of incorrect responses because such unobservable errors are not only plausible from a psychological perspective, but are also necessary to fully reproduce the patterns of response errors consistently found in empirical data (Wolfe et al., 2010; see also Fig. 9 in Section 4). Specifically, the following patterns have been observed:

(a) the miss rate $\mathbb{P}\left(Y_{k,1} = 0\right)$ increases with increasing set size k;
(b) the false alarm rate $\mathbb{P}\left(Y_{k,0} = 1\right)$ remains approximately constant with a slight downward trend when k increases;
(c) for a given set size, the miss rate tends to be higher than the false alarm rate;
(d) there is evidence for an acceleration in the growth of the miss rate with increasing set size (e.g., also in data of Zenger & Fahle, 1997), resulting in a convex shape of the graph of the miss rate on set size.

As we will show in the following two subsections, with our specifications of l, p_1, and p_2, Assumptions A1 and A2 capture these patterns. In contrast, the premise of an exhaustive search on trials with two or more adjacent set size levels is clearly inappropriate regardless of the processing structure, as the following considerations show. If we assume exhaustive search along with p_1 and p_2 being constant, then the expected frequencies given by the model are contrary to patterns (a), (b), and (c). Otherwise, if we assume straightforward dependencies of p_1 and p_2 on k, then the model either conflicts with pattern (c) or the expected frequencies deviate systematically from patterns (b) and (d) (see Li, 2019, pp. 71–74 for details; cf. Zenger and Fahle, 1997).

Importantly, our model regards omission and misidentification as the only error sources for two reasons. First, excluding

other types of unobservable error keeps the model simpler while maintaining its ability to account for the essential patterns of observed response errors. Second, Assumption A1 specifies a one-to-one correspondence between the termination conditions C1 and C2 and the response types TP and TA. Hence, the condition under which the queue has terminated is implied by the observed response. This allows us to determine the time course underlying either type of response. According to this property, the latent system response time underlying a TP or TA judgment in a k-customer queueing system, denoted by T_k^1 and T_k^0, respectively, is given by the time that elapses between stimulus onset t_0 and termination of the queueing process under C1 or C2. Note that the time each identified item spent in the two processing stages and the waiting room is encompassed in T_k^1 and T_k^0.

We model the RT observed on a trial as the sum of the system response time T_k^q, $q \in \{0, 1\}$, and a residual time. The residual time is defined as the time not spent on collecting and analyzing visual information, such as motor reaction time. It is modeled by an additive constant that depends only on the response types TP and TA, denoted by τ_1 and τ_0, respectively. We assume different residual time parameters for different response types because the execution may be affected by additional factors associated with the decision, such as readiness to either response type (cf. Fisher, 1982). Accordingly, our model entails the RTs observed under different conditions in a visual search experiment as follows:

$$X_{k,q} = \begin{cases} T_k^1 + \tau_1 & \text{if the system terminates under C1,} \\ T_k^0 + \tau_0 & \text{if the system terminates under C2,} \end{cases} \tag{3}$$

where $\tau_1, \tau_0 \geq 0$.

Under Assumption A1, Eq. (3) implies that, for a given level combination (k, q), the RT distribution is determined once the value of the associated response is known. More specifically, if $Y_{k,q} = 0$, the queue must have stopped at the lth departure epoch, be the response a correct rejection or a miss. If $Y_{k,q} = 1$, the queue may have stopped at any of the first l departure epochs, depending on the position of the item classified as a target in the departure order. Hence, T_k^1 follows a finite mixture distribution consisting of the distributions of the first l departure epochs, denoted by T_j, $j = 1, \ldots, l$, respectively. The associated probabilities (i.e., the mixture weights of the l components) depend on the condition (k, q) and the auxiliary parameters l, p_1, and p_2, according to Assumptions A1 and A2.

As the distribution of any departure epoch T_j is fully specified by the parameters (λ, μ, c), the model implies the conditional cumulative distribution function (CDF) of the RT, $X_{k,q}$, given the response, $Y_{k,q}$, to be

$$F_{X_{k,q}|Y_{k,q}}(t \mid y; l, p_1, p_2) = \begin{cases} \sum_{j=1}^{l} v_{j,k,q}(l, p_1, p_2) F_{T_j}(t - \tau_1; \lambda, \mu, c) & \text{if } y = 1, \\ F_{T_l}(t - \tau_0; \lambda, \mu, c) & \text{if } y = 0, \end{cases} \tag{4}$$

where $y \in \{0, 1\}$ denotes the values of the random variable $Y_{k,q}$ and F_{T_j} is the CDF of the jth departure epoch T_j, $j \in \{1, \ldots, l\}$. The weights $v_{j,k,q}$ of the mixture components are given by

$$v_{j,k,0}(l, p_1, p_2) = (1 - p_1)^{j-1} p_1 / \pi_1(k; l, p_1, p_2),$$

$$v_{j,k,1}(l, p_1, p_2) = \frac{(j-1)p_2(1-p_1)^{j-2}p_1 + (1-p_1)^{j-1}(1 - p_2 + (k-j)p_1)}{k(1 - \pi_2(k; l, p_1, p_2))},$$

where π_1 and π_2 are given by Eqs. (1) and (2). The weights $v_{j,k,1}$ of the target-present case are derived by calculating the conditional probability that the jth departing item is classified as a target given that the actual target appears in the ith position in the departure order, $i \in \{1, \ldots, k\}$. For details of the derivation of $v_{j,k,q}$, see Appendix A.2.

From Eqs. (1), (2), and (4), it is apparent that once l, p_1, and p_2 are determined, we obtain the marginal distribution of $Y_{k,q}$ as well as the conditional distribution of $X_{k,q}$ given $Y_{k,q}$. These three quantities are auxiliary parameters because their dependencies on set size k need to be further specified to account for patterns in the data in a psychologically meaningful manner. As detailed in the following two subsections, these specifications follow from the assumptions of incomplete search and imperfect processing. The charm of these assumptions is the simplicity of both the approach and the resulting specification of l. At the same time the latter accounts for many empirical findings, such as the patterns in accuracy data.

2.3. Incomplete search

We consider efficiency optimization as the principle underlying premature search termination and model it by means of the foraging theory (e.g., Stephens & Krebs, 1986). A human observer performing a visual search task is assumed to decide when to quit search analogous to a forager searching for food in patches of various sizes. In accordance with Assumption A1, we define the set of possible stopping policies by counting, that is, "stop and respond TP once the target is found, otherwise stop and respond TA after examining l items". Since this set is bijectively mapped to the set of all possible $l \in \{0, \ldots, k\}$, we simply use l to indicate the corresponding stopping policy. According to foraging theory, for all possible stopping policies, the long term mean reward rate is calculated considering the environment structure; then the best policy is chosen (see e.g., McNamara, 1982). In addition, we assume that participants impose a high subjective standard z of acceptable performance and try to maximize the mean reward rate of the entire experiment while achieving an average accuracy rate not below this standard. This is reasonable for two reasons. First, laboratory visual search tasks are usually designed to be easy for most people, such that obtaining a low accuracy rate would make participants feel uncomfortable. Second, participants usually understand that their search behavior is of interest and aim at showing high compliance.

Given set size k, we define the long term mean reward rate $\gamma_k(l)$ under a stopping policy l as the ratio of the expected number of correct responses to the expected number of items examined until search termination (cf. McNamara, 1982). For the sake of simplicity, our calculations in Appendix A.3 assume a balanced design and ignore the very rare misidentifications. Assume now that the experiment has m set size levels $\boldsymbol{k} = (k_1, \ldots, k_m)$, each appearing with equal probability. We assume without loss of generality that $k_1 < k_2 < \cdots < k_m$. Assume further that the observer adopts stopping policies $\boldsymbol{l} = (l_1, \ldots, l_m)$ for the respective levels. On any random trial, let $R_{\boldsymbol{k}}(\boldsymbol{l})$ denote the random variable that indicates a correct response, and $N_{\boldsymbol{k}}(\boldsymbol{l})$ the random variable of the number of items examined until termination. Then the long term mean reward rate is given by

$$\gamma_{\boldsymbol{k}}(\boldsymbol{l}) = \frac{\mathbb{E}(R_{\boldsymbol{k}}(\boldsymbol{l}))}{\mathbb{E}(N_{\boldsymbol{k}}(\boldsymbol{l}))} = \frac{\sum_{i=1}^{m} \frac{l_i + k_i}{2k_i}}{\sum_{i=1}^{m} \frac{4k_i l_i - l_i^2 + l_i}{4k_i}}. \tag{5}$$

The optimization problem can now be formulated as

$$\gamma_{\boldsymbol{k}}(\boldsymbol{l}) = \max_{l_1, \ldots, l_m}{}! \tag{6}$$

under the natural constraints $0 \leq l_i \leq k_i$ for $i = 1, \ldots, m$ and the accuracy constraint $\mathbb{E}\big(R_{\boldsymbol{k}}(\boldsymbol{l})\big) \geq z$. Here, max! indicates maximization of the function on the left side with respect to $\boldsymbol{l}$, subject to the specified constraints. Without loss of generality, we

may assume that $z \in [1/2, 1]$; in practical applications, we have $z \in (1/2, 1)$ with z typically close to 1.

In case $z \geq 1-(2m)^{-1}$, the optimal solution to Eq. (6) is given by

$$\left(l_1^*, l_2^*, \ldots, l_m^*\right) = \left(k_1, k_2, \ldots, k_{m-1}, k_m\,(1-2m(1-z))\right), \tag{7}$$

implying the search proportions

$$\left(r_1^*, \ldots, r_m^*\right) := \left(l_1^*/k_1, \ldots, l_m^*/k_m\right) = (1, \ldots, 1, 1-2m(1-z)).$$

The above two equations formalize the following rule: Unless the target is found, perform an exhaustive search on the smaller set sizes to achieve maximum accuracy; only on the trials with the largest set size, terminate the search so early that the minimum required accuracy to reach an average accuracy of z can be obtained. For the derivation of the optimal stopping policy given by Eq. (7) under rather general assumptions, see Appendix A.3. The inequality $z \geq 1-(2m)^{-1}$ is, for instance, satisfied for $m = 4$ and $z \geq .875$, or for $m = 5$ and $z \geq .9$. Since most visual search tasks use less than five set size levels and are designed in a way that high accuracy is facilitated, the case $z \geq 1-(2m)^{-1}$ is most relevant for empirical applications.

In reality, of course, participants hardly perform such a formal analysis deliberately and implement the optimal stopping policy exactly as given by Eq. (7). Moreover, an exhaustive search under two or more adjacent set sizes conflicts with patterns in empirical accuracy data, as discussed in Section 2.2. Nevertheless, it seems plausible that the actual search termination behavior is largely shaped by the optimal stopping policy, but with some added flexibility (i.e., performing nearly exhaustive searches on trials with smaller set sizes and reducing search proportions to some degree when set size is large). Accordingly, we model the *actual* termination by an approximation of the optimal stopping policy using a smoothing approach, see Section 2.5.

2.4. Imperfect processing

Incomplete search alone does not suffice to account for all the patterns (a) to (d) in accuracy data (see pattern (b) in Section 2.2). An adaptation of the identification threshold to the set size is necessary to fully capture these patterns. For this purpose, the dependencies of p_1 and p_2 on k need to be specified.

Thanks to Eq. (1), pattern (b) reveals unique information with regard to p_1. Recall that the empirically observed false alarm rates, that is, the realizations of π_1, remain approximately constant with a slight decreasing trend as k increases. If π_1 were constant, we could solve Eq. (1) for p_1. Taking the slight decline into account, we conclude in Appendix A.4 that it is reasonable to specify the dependency of p_1 on k by

$$p_1(k) = a_1 k^{-b} \tag{8}$$

for some $a_1 \in [0, 1]$ and $b > 0$ with $a_1^{1/b} < k_1 = \min\{k_1, \ldots, k_m\}$. The derivation in Appendix A.4 also implies that, more specifically, a_1 should be close to 0 and b slightly larger than 1.

This power law relationship is convincingly supported by an observation based on the two sets of accuracy data from Wolfe et al. (2010). A naive estimate of p_1 appears to follow indeed a power law of set size k for each data set, as Fig. 6 shows. See Appendix A.4 for details.

Although a similar approach does not work for p_2 (see again Appendix A.4 for details), it is reasonable to also allow p_2 to depend on k for two reasons. First, for a binary classifier in general, suppressing the Type I error rate $\mathbb{P}(\text{“}\mathcal{T}_i = 1\text{”} \mid \mathcal{T}_i = 0)$ is associated with an increase of the Type II error rate $\mathbb{P}(\text{“}\mathcal{T}_i = 0\text{”} \mid \mathcal{T}_i = 1)$ and vice versa. Second, assuming a constant p_2 predicts a slightly slower growth of the miss rate than observed in empirical data. This discrepancy diminishes when we assume that p_2 varies

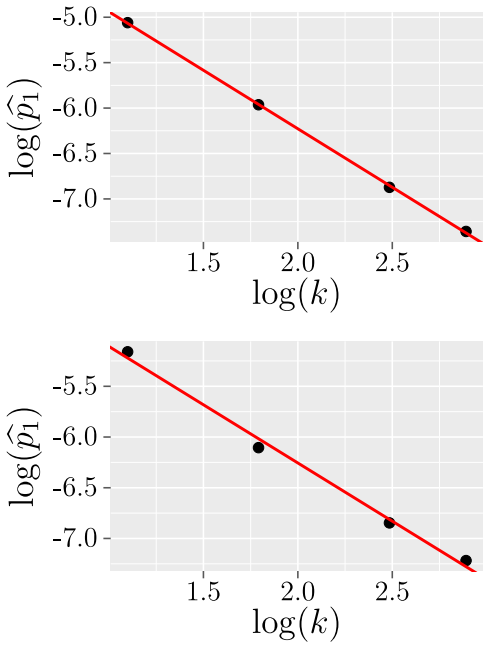


Fig. 6. Log–log plot of the naive estimates of p_1 on set size k for conjunction search data (top) and spatial configuration data (bottom) collected by Wolfe et al. (2010).

as a power function of k, resulting in a better approximation of pattern (d). More precisely, we assume that $p_2(k) = a_2 k^{b_2}$ for some $b_2 > 0$ and $a_2 > 0$. We further infer from an analysis of the Receiver Operating Characteristic curve that p_2 most likely increases with k at a similar speed as p_1 falls (i.e., $b_2 = b$), and accordingly,

$$p_2(k) = a_2 k^b, \quad \text{where } a_2, b > 0 \quad \text{and} \quad a_2^{-1/b} > k_m. \tag{9}$$

2.5. Full model accounting for accuracy and latency jointly

As mentioned in Section 2.3, we model the termination by a smooth approximation of the optimal stopping policy. We suggest to approximate l_k^* by

$$\tilde{l}(k) := k\,(1-\exp(\beta k+\alpha)) \quad \text{for } -\alpha/\beta \geq k \quad \text{and} \quad \beta > 0. \tag{10}$$

The goodness of the proposed approximation is best illustrated by the corresponding search proportions $\left(r_1^*, \ldots, r_m^*\right)$ and the graph of $\tilde{r}(k) := \tilde{l}(k)/k = 1-\exp(\beta k+\alpha)$, see Fig. 7. The parameter β represents the steepness of the curve and can be interpreted as the extent to which the search proportions $\tilde{r}(k_i)$ for different set size levels k_i are polarized. Here, “polarization” indicates that parameters $\tilde{r}(k_1), \ldots, \tilde{r}(k_{m-1})$ are close to 1 while $\tilde{r}(k_m)$ is close to 0. Specifically, the larger β, the steeper the curve of $\tilde{r}(k_i)$, thus the larger the jump at the largest set size level.

We choose this specific form for two reasons. First, it results in decreasing search proportions $\tilde{r}(k)$ for increasing k, which is necessary to capture pattern (a). Second, it leads to a convex form of the miss probability $\mathbb{P}(Y_{k,1} = 0)$, in line with pattern (d) found in empirical data.

Combining Eq. (10) with Eqs. (8) and (9), we obtain the following explicit equations for the error probabilities of responses:

$$\pi_1(k) = 1-\left(1-a_1 k^{-b}\right)^{k(1-\exp(\alpha+\beta k))}; \tag{11}$$

$$\pi_2(k) = \left[\left(1-a_1 k^{-b}\right)\exp(\alpha+\beta k) + a_2 k^b\,(1-\exp(\alpha+\beta k))\right] \cdot \left(1-a_1 k^{-b}\right)^{k(1-\exp(\alpha+\beta k))-1}. \tag{12}$$

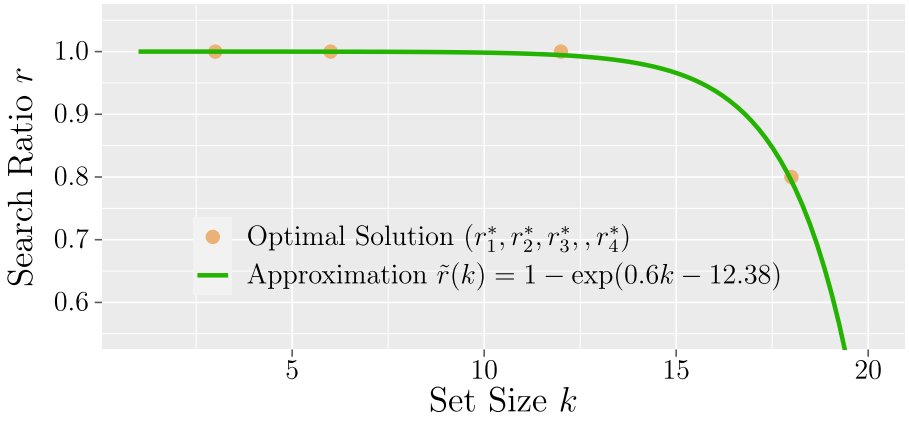


Fig. 7. The search proportions $(r_1^*, r_2^*, r_3^*, r_4^*) = (1, 1, 1, 0.8)$ that correspond to the theoretically optimal stopping policy under the accuracy constraint $z = 0.95$ in an experiment with set size levels $(3, 6, 12, 18)$. The search proportions are approximated by the function $\tilde{r}(k) = 1 - \exp(0.6k - 12.38)$.

Given our specification of the emergence of response errors, the model hypothesizes the marginal distribution of the response, $Y_{k,q}$, as

$$P_{Y_{k,q}}(y) = (1-q)\pi_1^y\,(1-\pi_1)^{1-y} + q\,(1-\pi_2)^y\,\pi_2^{1-y}, \tag{13}$$

where $q, y \in \{0, 1\}$. The model parameters $(\alpha, \beta, a_1, a_2, b)$ in π_1 and π_2, which we refer to as accuracy-related parameters, as well as set size k, are omitted for clarity.

As discussed in Section 2.2, by determining l, p_1 and p_2, we also obtain the distributions of T_k^1 and T_k^0 in Eq. (3), and thus the conditional distribution of $X_{k,q}$ given $Y_{k,q}$. However, there is a technical issue that needs to be resolved: While Eq. (10) introduces non-integer values of l, Assumption A1 implicitly requires l to be a natural number in order to model $X_{k,q}$. Simply rounding the value of $\tilde{l}$ obtained from Eq. (10) to a natural number leads to a considerable bias. Therefore, we adopt a probabilistic approach: the termination criterion in Assumption A1 functions as a random variable that takes the value $\lceil\tilde{l}\rceil$ with probability $\tilde{l} - \lfloor\tilde{l}\rfloor$ and the value $\lfloor\tilde{l}\rfloor$ with probability $1 - (\tilde{l} - \lfloor\tilde{l}\rfloor)$. In other words, the termination criterion is modeled by a two-point random variable with the floor $\lfloor\tilde{l}\rfloor$ and the ceiling $\lceil\tilde{l}\rceil$ of $\tilde{l}$ as possible outcomes and $\tilde{l}$ as the expected value. The distributions of T_k^1 and T_k^0 can thus be determined as finite mixture distributions. In particular,

$$\begin{aligned}\tilde{F}_{X_{k,q}|Y_{k,q}}(t \mid y) = {} & (1 + \lfloor\tilde{l}\rfloor - \tilde{l})F_{X_{k,q}|Y_{k,q}}(t \mid y; \lfloor\tilde{l}\rfloor, p_1, p_2) \\ & + (\tilde{l} - \lfloor\tilde{l}\rfloor)F_{X_{k,q}|Y_{k,q}}(t \mid y; \lceil\tilde{l}\rceil, p_1, p_2),\end{aligned} \tag{14}$$

where $F_{X_{k,q}|Y_{k,q}}$ is given by Eq. (4) with $\tilde{l}$, p_1, p_2 as specified in Eqs. (10), (8), and (9). Again, the parameters $(\alpha, \beta, a_1, a_2, b)$ are omitted for clarity. In addition to these accuracy-related parameters, the model RTs also depend on $(\lambda, \mu, c, \tau_1, \tau_0)$, which can be inferred solely from the RT data, cf. Section 2.2. We therefore call them RT-related parameters.

Hence, our model explicitly and completely specifies the respective dependencies of $P_{Y_{k,q}}$ and $\tilde{F}_{X_{k,q}|Y_{k,q}}$ on set size k. As a result, regardless of the specific set size k_i on a visual search trial or the total number m of set size levels studied in the experiment, the number of free parameters — both RT-related and accuracy-related — remains constant.

3. Parameter estimation

The proposed queueing model has 10 free parameters in total: $\alpha, \beta, a_1, a_2, b, \lambda, \mu, c, \tau_1, \tau_0$. Since our model specifies an analytical closed form for π_1 and π_2 in Eq. (13), we estimate the accuracy-related parameters $\alpha, \beta, a_1, a_2, b$ using a maximum likelihood (ML) estimation approach. An estimator with explicit analytical form is not available for F_{T_i}, hence also not for $\tilde{F}_{X_{k,q}|Y_{k,q}}$, due to the non-standard assumptions in our model, e.g., the non-constant arrival intensity. Therefore, with respect to modeling the RTs on a distributional level, we approximate $\tilde{F}_{X_{k,q}|Y_{k,q}}$ by Monte-Carlo simulation. Accordingly, we estimate the RT-related parameters $\lambda, \mu, c, \tau_1, \tau_0$ using a minimum distance (MD) estimation approach. In accordance with the respective dependencies of π_1, π_2 and $\tilde{F}_{X_{k,q}|Y_{k,q}}$ on the parameters, the accuracy-related parameters are estimated ahead of the RT-related parameters and their ML estimates are used as input for the MD estimation procedure. For technical details of the complete parameter estimation procedure, we refer to Chapter 6 in Li (2019). The parameter estimation has been implemented in R (R Core Team, 2021) by Schlather and Li (2022). The full reproducible R scripts for data fitting are available at the Open Science Framework (https://osf.io/k9573/).

3.1. ML estimation for accuracy-related parameters

Assume a balanced two-factorial within-subject design of m set size levels $(k_1, \ldots, k_m)$ crossed with target presence versus absence in a display ($q = 1$ versus $q = 0$, respectively). Let $n_{k,q,y}$ denote the number of observations of $Y_{k,q} = y$, $y \in \{0, 1\}$, on trials with set size k and objective target status q. Based on Eq. (13), the log-likelihood ℓ of the response pattern $\left(n_{k_i,1,1}, n_{k_i,1,0}, n_{k_i,0,1}, n_{k_i,0,0}\right)$ is given by

$$\begin{aligned}\ell(\alpha, \beta, a_1, a_2, b) = \sum_{i=1}^{m} \Big[& n_{k_i,1,1} \ln(1 - \pi_2(k_i)) + n_{k_i,1,0} \ln(\pi_2(k_i)) \\ & + n_{k_i,0,1} \ln(\pi_1(k_i)) + n_{k_i,0,0} \ln(1 - \pi_1(k_i)) \Big] \\ & + C,\end{aligned}$$

where C is a constant, and π_1 and π_2 are given by Eqs. (11) and (12).

To ensure that estimated probabilities are in $[0, 1]$, we reparameterize the model by $\tilde{\beta} = -\beta k_m/\alpha$ and $\tilde{b} = -b\ln(k_m)/\ln(a_2)$. Note that, by assumption, k_m is the greatest element of $\{k_1, \ldots, k_m\}$. To refine the estimation, we apply a technique of profiling (see e.g., Severini & Wong, 1992) to $\tilde{\beta}$.

3.2. MD estimation for RT-related parameters

The MD estimation approach searches for the parameter values that minimize the distance between the CDFs of the model

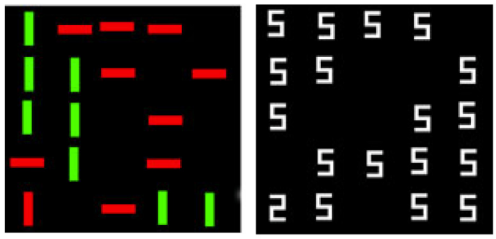


Fig. 8. Two examples of search displays used in Wolfe et al. (2010), a conjunction search task (left, target = vertical red bar) and a spatial configuration search task (right, target = 2). From "Serial_Conj", by Visual Attention Lab, 2010, Retrieved from http://search.bwh.harvard.edu/new/data_set_files.html. Copyright 2010 by Visual Attention Lab.

RT and the empirical data (see e.g., Basu, Shioya, & Park, 2011). The distance is measured by the Wasserstein metric of order 1 (Vallender, 1974):

$$W(G_n, F_\theta) = \int_{-\infty}^{\infty} |G_n(x) - F_\theta(x)|\, \mathrm{d}x,$$

where F_θ is a parametric model CDF with parameter θ and G_n the empirical CDF of the observed data with sample size n.

As stated above, the conditional CDFs of RTs given by Eq. (14) are approximated numerically using Monte-Carlo simulation. We developed an algorithm (see Appendix B) for the simulation of a queueing process as specified in Section 2. Each simulation run returns a TP or TA response and the associated RT based on Eq. (3) for a given level combination and a given parameter vector. We call the data simulated or collected under a level combination (k, q) a sub-data set. Each simulated sub-data set consists of 10^5 simulation runs, which are batched into a block. To ensure that the distance calculation returns the same value for the same set of parameter inputs, the random seed is reset to 0 once at the beginning of the block of simulation runs for a sub-data set.

Taking RTs of both correct and incorrect responses on both target-present and target-absent trials into account means that a data set with m set size levels results in $4m$ comparisons between the model and empirical CDFs and thus $4m$ distances for the entire experiment. From a psychological viewpoint, it is reasonable to assume that the same parameter values underlie the data observed under different conditions in the same visual search experiment. This constraint requires to search for common parameter values for all sub-data sets that minimize a global distance, which is best defined as the weighted sum of involved distances, using the relative informativeness of the data as weights, as detailed below. However, simply fitting the model to all sub-data sets at once leads to overparameterization and causes identifiability issues, making parameter estimates unstable.

In order to combine the information across sub-data sets in an efficient way and to cope with the identifiability issues, we develop a hierarchical, iterative parameter estimation procedure (see Appendix C). This procedure integrates information from individual sub-data sets to find better estimates step by step along a hierarchical structure. First, each sub-data set is fitted separately with the profiling technique (see e.g., Severini & Wong, 1992) applied to λ^{-1}, the parameter for interarrival times. The estimates of individual sub-data sets are then averaged to obtain the initial values for the joint fitting of sub-data sets. The distances involved in the calculation of each step are weighted by the relative informativeness of the data they are based on. When the model is fitted to correct and incorrect RTs of a sub-data set jointly, either distance is weighted by the relative frequency of the respective responses within this sub-data set. For the sub-data set obtained in condition (k, q), we denote the frequency-weighted sum of the resulting distances by $W_{k,q}$. When different sub-data sets are fitted jointly, each $W_{k,q}$ is additionally divided by the lowest value (i.e., the best individual fit) obtained in the first step for the corresponding sub-data set. The global distance in the second step is then the sum of the relativized $W_{k,q}$. This relativization prevents sub-data sets with large individual distances, which indicate comparatively bad fits, from dominating the optimization. To deal with over-parameterization, a subset of parameters is held fixed when the complementary set is estimated. The process is iterated alternatingly with exchanged roles of the two subsets. Because c, the number of servers, has discrete values, it is held fixed when using this procedure to estimate other parameters. The entire procedure is applied for each $c = 1, \ldots, 10$ to find the best fitting c.

Table 1
ML estimates of accuracy-related parameters and corresponding log-likelihoods for the conjunction search and spatial configuration search data of Wolfe et al. (2010).

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{a_1}$	$\hat{a_2}$	$\hat{b}$	log-likelihood
CS	−4.01	.0531	.0262	1.00×10^{-5}	1.28	−27.69
SC	−4.41	.114	.0182	1.21×10^{-5}	1.11	−31.76

Note. CS = conjunction search; SC = spatial configuration search. The ML estimates are rounded to three significant digits.

4. Data fitting

4.1. Brief description of the experiments by Wolfe et al. (2010)

Our model was fitted to data aggregated across individual participants collected in two standard visual search experiments by Wolfe et al. (2010). Ten and nine participants performed a conjunction search task and a spatial configuration search task, respectively. Each participant completed a total of 4000 experimental trials. The experiments followed a factorial within-subject design of set size (3, 6, 12, or 18) and the objective status of the target (absence or presence). On each trial, the levels of both factors were selected randomly and independently with equal probabilities of 1/4 and 1/2, respectively. Fig. 8 illustrates two examples of the search displays. With a delay of 500 ms, the search display appeared and remained visible and unchanged until the participant responded. Whenever a key was pressed, indicating either a target-present or a target-absent response, a feedback was shown for 500 ms. Participants were instructed to respond as quickly and accurately as possible. Following Palmer et al. (2011) and Wolfe et al. (2010), we excluded trials with RT < 200 ms.

4.2. Results

For both task types, the ML estimates of the accuracy-related parameters (Table 1) resulted in predicted relative frequencies very similar to the empirically observed error rates, see Fig. 9.

Based on the accuracy-related parameters in Table 1, we obtained MD estimates for the RT-related parameters using the Monte-Carlo-based procedure described in Section 3.2. For the conjunction search response latencies, a generally good fit was already obtained when all parameters were kept invariant across experimental conditions. For the spatial configuration search latencies, this restriction led to a sufficiently good fit only for some of the sub-data sets. However, allowing μ^{-1} to depend on set size improved the fit markedly. For both task types, the best overall fit was observed for $c = 5$ with the corresponding MD estimates

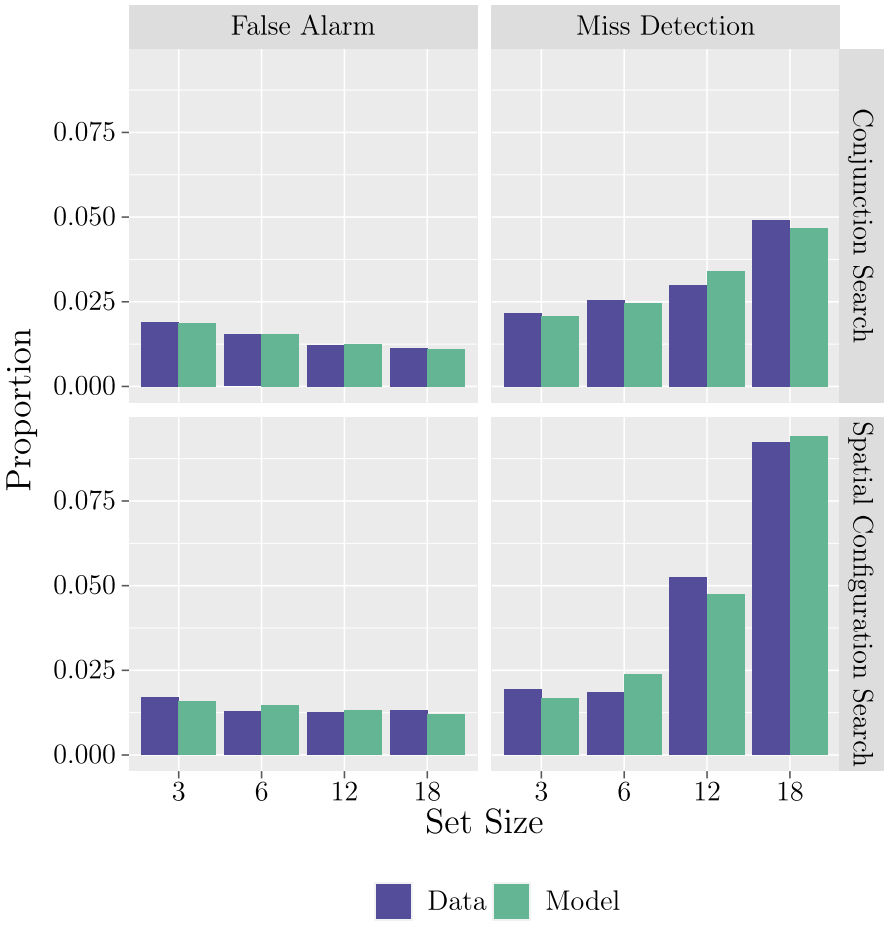


Fig. 9. Comparison of observed false alarm (left column) and miss (right column) rates and the corresponding frequencies expected from the model, separately for conjunction search data (upper row) and spatial configuration search data (lower row) of Wolfe et al. (2010).

Table 2
MD estimates of RT-related parameters in milliseconds and corresponding sums of distances for the conjunction search and spatial configuration search data sets from Wolfe et al. (2010).

	$\hat{c}$	$\widehat{\lambda^{-1}}$	$\widehat{\mu_3^{-1}}$	$\widehat{\mu_6^{-1}}$	$\widehat{\mu_{12}^{-1}}$	$\widehat{\mu_{18}^{-1}}$	$\widehat{\tau_0}$	$\widehat{\tau_1}$	$\sum_i W_{k_i,0}$	$\sum_i W_{k_i,1}$
CS	5	23	151	151	151	151	245	353	407	144
SC	5	81	235	361	493	501	285	388	453	158

Note. CS = conjunction search; SC = spatial configuration search. The subscript next to μ indicates the set size level it applies to. The constraint $\mu_3 = \mu_6 = \mu_{12} = \mu_{18}$ was applied to the data fitting for conjunction search. The numbers are rounded to the nearest integer.

displayed in Table 2. As illustrated in Figs. 10, 11, 12, and 13, the conditional RT distribution expected under the model deviated less from the empirical one for small set sizes than for large ones. The same was true for target-present versus target-absent trials, correct versus incorrect responses, and conjunction search versus spatial configuration search tasks, respectively.

4.3. Discussion

Fig. 9 demonstrates that the model is able to reproduce marginal error rates of responses in both task types almost perfectly. The minor deviations do not exhibit any systematic pattern. Figs. 10 to 13 show that the model is able to capture the conditional RT distributions quite well, even for the rare incorrect responses. However, the deviations show certain patterns. The most noticeable one is that for large set sizes, the empirical RTs of correct rejections and misses have bigger tails than the model RTs. If this was the result of a larger variability in processing times of single stimuli, the same pattern should be observed for hits and false alarms. Yet this is not the case. Thus, the underestimation of extreme latencies may arise from larger variability in empirical search terminations for target-absent responses than implied by the model.

Both task types share the same estimate for the number of servers in the queueing system, that is, the maximum number of items allowed in the identification stage at the same time ($\hat{c} = 5$). Notably, this estimate of the number of parallel processing channels available for controlled processing is roughly in line with previous studies and alternative models of visual information processing. For example, in the time-dependent limited-channel model (Fisher, 1982), the number of channels that best fitted the empirical mean visual search RTs ranged between three and five.

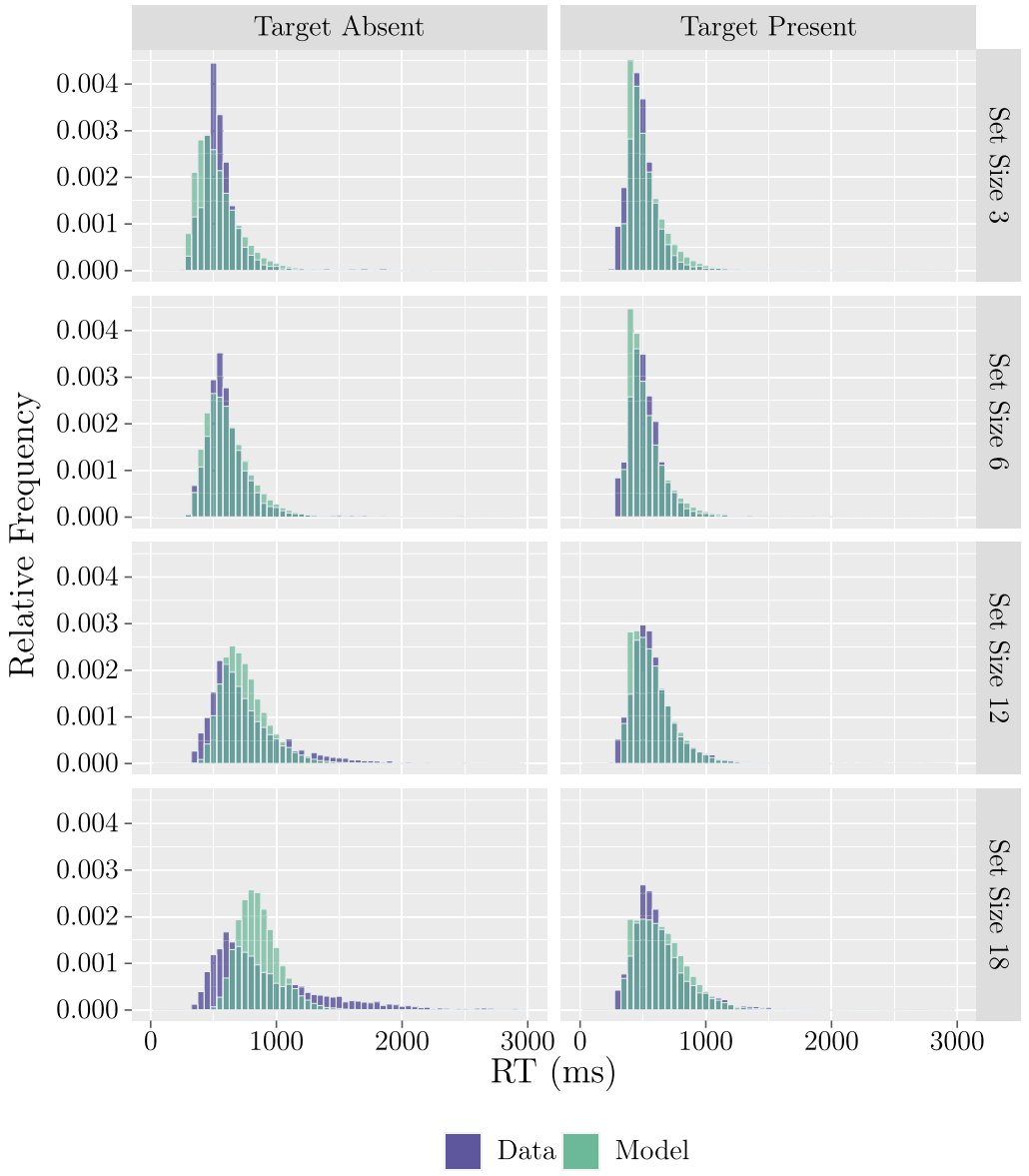


Fig. 10. Comparison of the empirical and the model-based conditional distributions of the RT given the response is *correct* in the *conjunction search* data from Wolfe et al. (2010) for $c = 5$, separately for target presence vs. absence and different set sizes.

Fisher (1984) obtained the same range based on the results of fitting the steady-state limited-channel model to data collected with the multiple-frame paradigm. The Theory of Visual Attention (Bundesen, 1990) assumes a capacity limit of four items. In the same vein, based on different empirical procedures, visual working memory has been argued to have a capacity limit of three to five items in young adults (Cowan, 2001; Isbell, Fukuda, Neville, & Vogel, 2015; Luck & Vogel, 1997; Pashler, 1988). Thus, the proposed queueing model provides converging evidence concerning the number of available attentive processing channels for visual information.

Moreover, the estimates for the residual times τ_1 and τ_0 are remarkably similar across both task types as well. The finding $\widehat{\tau_1} \approx \widehat{\tau_0} + 100$ may reflect a stronger readiness for the target-absent response regardless of task type. This finding is also in line with the viewpoint that deciding to respond TP might induce additional processes, such as saccade planning, compared to simply deciding to respond TA (Wolfe et al., 2010).

The estimates of the remaining parameters are consistent with the subjectively experienced higher difficulty of the spatial configuration task compared to the conjunction task. The clearly larger values of $\widehat{\lambda^{-1}}$ and $\widehat{\mu^{-1}}$ for the former indicate longer processing times in both the preattentive stage and the attentive stage. An intriguing finding is that relaxing the constraint on μ^{-1} improves model fit noticeably for spatial configuration search data, and that $\widehat{\mu^{-1}}$ increases with set size (see Table 2). One possible explanation is that adaptation of the identification threshold to set size demands more attentional resources for spatial configuration stimuli than for conjunction stimuli. If this is the case, the adaptation should be more restricted due to the higher costs. This reasoning is in line with the finding that the value of $\hat{b}$ is smaller for the spatial configuration task, indicating that p_1 and p_2 are less sensitive to set size. In sum, compared to conjunction stimuli, identification of individual spatial configuration stimuli appears to require more cognitive resources. In addition, its adaptation to set size is possibly more costly. As a

Fig. 11. Comparison of the empirical and the model-based conditional distributions of RT given the response is *incorrect* in the *conjunction search* data from Wolfe et al. (2010) for $c = 5$, separately for target presence vs. absence and different set sizes.

counterbalance, the termination criterion more strongly depends on set size, as indicated by the larger value of $\hat{\beta}$. That is, during spatial configuration search, participants search more thoroughly on small trials but more hastily on large trials.

5. Model uncertainty analysis

Since c represents the number of parallel, independent processing channels in the attentive stage, it is the most significant structural parameter of the proposed queueing model. Thus, we used a bootstrap approach to assess the extent to which the model fit, as measured by the Wasserstein metric, can reflect a misspecification regarding c. We compared the proposed hybrid model with a standard serial model, focusing on the most relevant aspect: the ability to distinguish between a single-server system, $c = 1$, which represents strictly serial attentive processing, and a multiserver system, $c = 5$, which fits Wolfe et al.'s (2010) data best. The R-scripts and data for the uncertainty analysis are also available at the Open Science Framework (https://osf.io/k9573/).

We used the best-fitting parameter values to the conjunction search data given $c = 1$ and $c = 5$ to simulate 600 samples in each case. Each sample contained 40,000 trials, assuming the same experimental design as Wolfe et al. (2010). We then fitted the simulated data sets to the queueing model with the constraint $\mu_3 = \mu_6 = \mu_{12} = \mu_{18}$. Similarly, for the model variant without the constraint of a common μ, we simulated 618 samples for $c = 1$ and $c = 5$ each, using the corresponding best-fitting parameter values to the spatial configuration search data.

As displayed in Table 3, the model with the correct number of servers was selected in more than 90% of the cases. As shown in Figs. D.14 and D.15 in Appendix D, fitting the data originating from a single-server system to any multiserver system results in a clearly larger misfit that increases with increasing $\hat{c}$, with or without the constraint of a common μ. Similarly, fitting the data

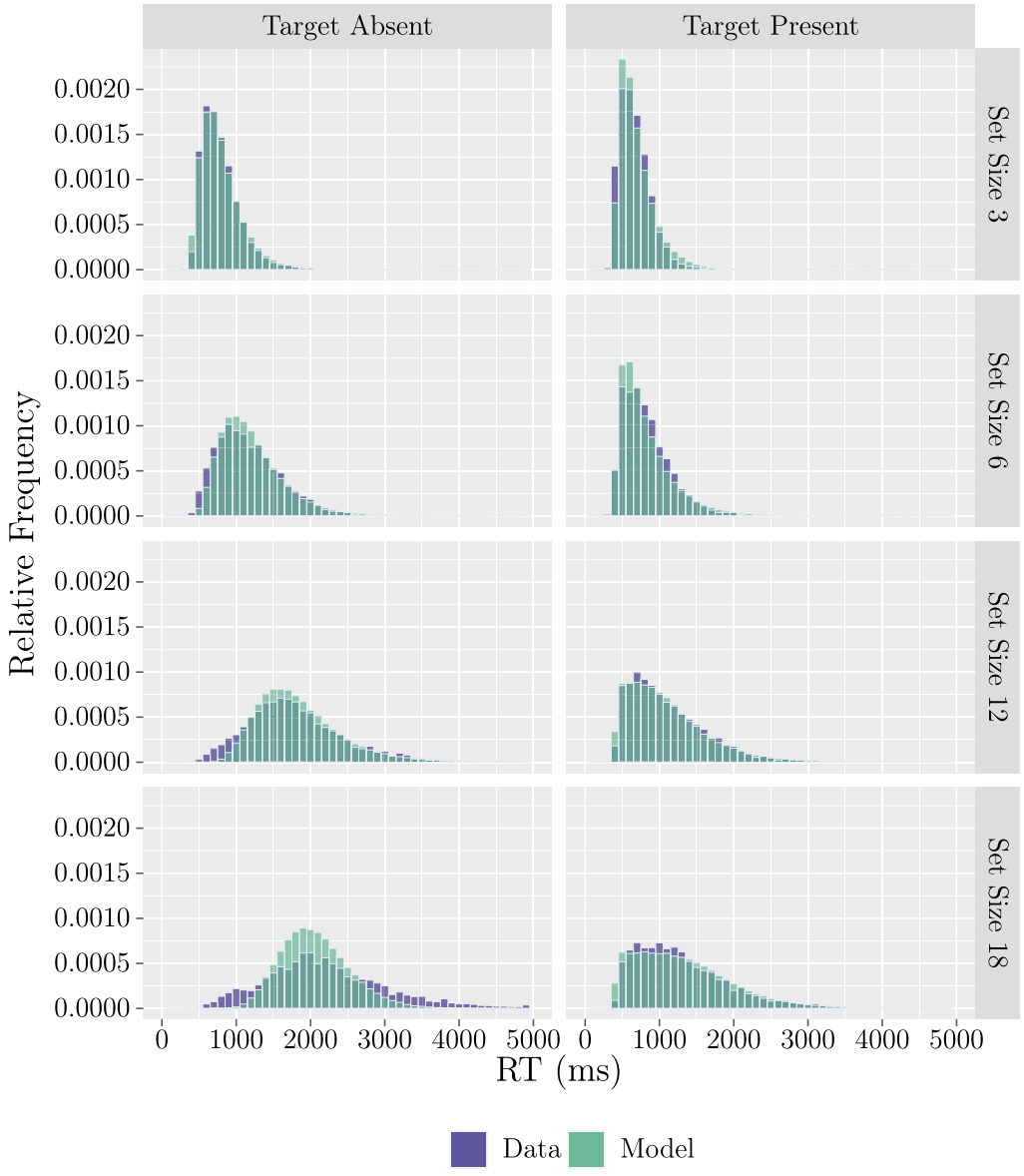


Fig. 12. Comparison of the empirical and the model-based conditional distributions of RT given the response is *correct* in the *spatial configuration search* data from Wolfe et al. (2010) for $c = 5$, separately for target presence vs. absence and different set sizes.

originating from a five-server system to a single-server system produces a notably larger misfit. Again, fitting multiserver models with $1 < \hat{c} < 5$ results in somewhat smaller but still obvious misfit. These results indicate that a model misspecification regarding c can be well reflected in the total distance and identified correctly in almost all cases. Notably, we did not observe a single sample generated from a single-server system that was classified as best-fitting a five-server system.

Table 3
Frequencies of bootstrap samples generated from a c server system and best-fitting a $\hat{c}$ server system.

	With constraint on μ				Without constraint on μ			
$\hat{c} =$	1	2	5	6	1	4	5	6
$c = 1$	598	2	0	0	618	0	0	0
$c = 5$	0	0	595	5	0	58	559	1

6. Conclusion

The existing empirical findings on visual search do not allow for a coherent theoretical conclusion on the serial/parallel debate. Serial models are in line with some findings and in conflict with others; the same applies to parallel models. Moore and Wolfe (2001) explained how a hybrid structure that combines serial and parallel features could accommodate the apparently conflicting findings. We formalize and implement this notion as a single-line, multiserver queueing model in continuous time. This model is well embedded in the theoretical framework of two processing stages in visual search (Treisman, 1986; Treisman & Gelade, 1980) and accounts for responses and RTs jointly on a distributional level. We incorporate two psychological mechanisms to explain the emergence of incorrect responses in the

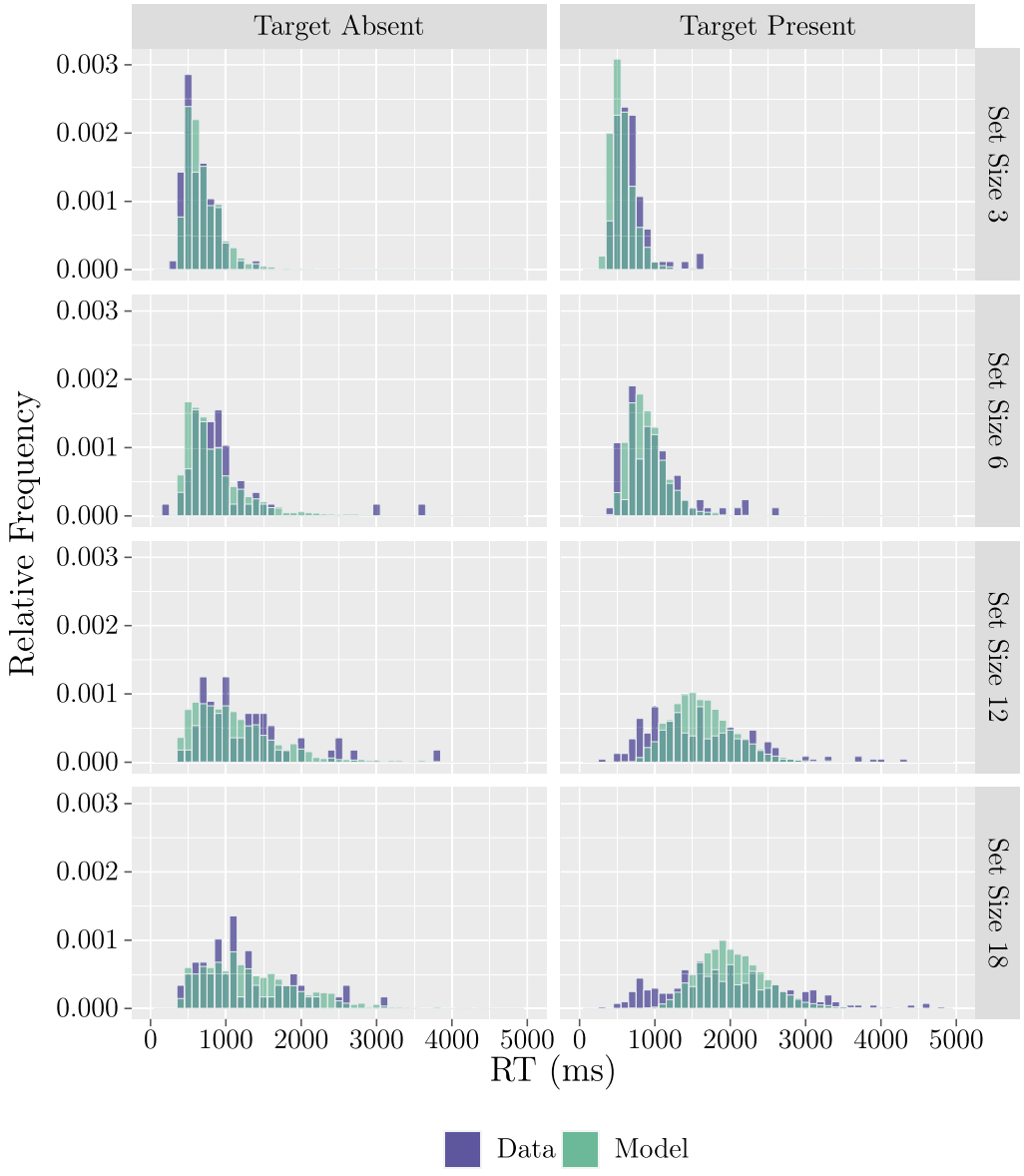


Fig. 13. Comparison of the empirical and the model-based conditional distributions of RT given the response is *incorrect* in the *spatial configuration search* data from Wolfe et al. (2010) for $c = 5$, separately for target presence vs. absence and different set sizes.

queueing system, namely, a threshold adaptation that accounts for genuine processing errors and a quasi-optimal stopping rule that accounts for omission errors. In addition to proposing a maximum-likelihood method for fitting accuracy data, we solve the problem of the lack of an explicit analytical form for the conditional CDF of the model RT by developing a numerically stable, well-performing minimum-distance parameter estimation procedure based on Monte-Carlo simulation.

We fitted the model to two prototypical empirical data sets of Wolfe et al. (2010), obtained with conjunction search and spatial configuration search tasks, respectively. The proposed queueing model captures the accuracy rates almost perfectly for both data sets. Regarding the RT distributions, a generally good fit is found for the model with five servers. The model uncertainty analysis clearly shows that such a good model fit cannot be achieved when the data originate from a strictly serial processing system. Thus, the results provide convincing empirical support for the proposed hybrid structure beyond the dichotomy of serial and parallel processing.

Our results show that a mechanism that exhibits both serial and parallel aspects underlies selective attention in standard visual search tasks with different stimulus types. The serial character of our model rests on the single waiting line that represents an attentional bottleneck, while subsequent attentive processing of items can be seen as asynchronous-parallel with limited capacity. Furthermore, the results also demonstrate that accuracy data and the RTs of both correct and incorrect responses are very important for a comprehensive account of visual search performance.

On the theoretical level, our model provides a novel approach to account for the "bottleneck" phenomenon in visual attention data. Since most theories and models of visual attention acknowledge the existence of at least one parallel stage in visual processing, the core of the serial/parallel debate essentially is

how to explain the positive mean RT × set size slopes found in empirical data. A serial model explains them as a direct consequence of a mandatory serial processing stage, whereas a limited-capacity parallel model proposes sequential sampling of visual stimuli. In contrast, our model explains the positive slope by a serial transition between two parallel processing stages with different capacities. Accordingly, the attentional "bottleneck" arises as an emergent property of a single-line queueing system that enables efficient information flow between a massively parallel preattentive stage and a limited-capacity parallel attentive stage of processing.

Several aspects need to be considered for further developments and extensions of the model. First, possible extensions of the model should take guidance in visual search (e.g., Wolfe, 1994, 2021) into account. For instance, the assumption of equiprobable arrival orders can be replaced by a prioritized arrival order, so that an item with a salient feature has a higher probability of being among the first to get through the preattentive stage (bottom-up guidance). Likewise, the FCFS discipline can be replaced with a priority queueing discipline, allowing an item with a goal feature to be served ahead of other items in the waiting room (top-down guidance). Second, inhibition of return should be modeled in a more realistic manner. Wolfe (2012) pointed out that there is sufficient inhibition of return to locations searched previously but no perfect memory for rejected items. Taking non-perfect inhibition of return into account may improve the fit at the right tail of RT distributions for large set sizes. Third, the calculation of the reward rate should probably incorporate expenditure resulting from key pressing and waiting for the next display as factors that affect effort. Fourth and most importantly, the generalizability of the model needs to be investigated. Fisher (1982, 1984) specified different queueing models to account for RT data obtained by different RT methods (see Wolfe, 2018) and accuracy data obtained by several accuracy methods, respectively. Ultimately, the question arises whether data from different visual attention paradigms can be explained within a unified model framework, especially the attentional dwell time findings combined with visual search findings.

In addition, we regard the development of appropriate, well-understood goodness-of-fit tests and model-selection measures for combined accuracy and latency data as the most important statistical advancement for future applications of the queueing model and its variants. Based on such statistical methods, selective manipulations of specific cognitive processes underlying the model parameters could then be conducted to evaluate whether an experimental manipulation maps on the corresponding model parameter correctly. If successful, such analyses would lend further support to the psychological validity of the proposed queueing model of visual search in general and of its parameters in particular.

Data availability

Data and R code related to this article can be found at the Open Science Framework at https://osf.io/k9573/.

Acknowledgment

This work was supported by the Graduate School of Economic and Social Sciences at the University of Mannheim, Germany, the Research Training Group "Statistical Modeling in Psychology" (GRK 2277), and the Research Training Group "Statistical Modeling of Complex Systems and Processes" (GRK 1953), all funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). The first author was also supported by the Independent Research Fund Denmark under Grant DFF 8018-00014B "Foundations of response time measurement". We thank Jonas Brehmer and Christopher Dörr for their valuable comments and suggestions, as well as for proofreading early versions of this article. We are also very grateful to Professor Richard Schweickert and two anonymous reviewers for their insightful and constructive comments and detailed suggestions.

Appendix A. Comments on and derivations of selected equations

A.1. Comments on the error probabilities of responses, Eqs. (1) and (2)

In our model, a false alarm error can arise only from misidentification. To see

$$\pi_1(k; l, p_1, p_2) := \mathbb{P}(Y_{k,0} = 1) = 1 - (1 - p_1)^l,$$

note that we have $\mathbb{P}(Y_{k,0} = 0) = (1 - p_1)^l$, where $1 - p_1$ is the probability of a correct classification of a distractor. The search is stopped immediately after correctly identifying l distractors.

A miss error can arise from omission or misidentification. If the target is not within the first l departing items (which happens with probability $(k - l)/k$), an omission occurs. In this case, the event $Y_{k,1} = 0$ occurs if and only if all the l distractors are correctly detected. That is, the miss error arises from omission with probability $(1 - p_1)^l (k-l)/k$. In the complemental case, the target can be in any of the l positions of the departing sequence. For each of these l subcases, the event $Y_{k,1} = 0$ occurs if and only if the target is misidentified (with probability p_2) and $l - 1$ distractors are correctly identified. That is, the miss error arises from misidentification with probability $(1 - p_1)^{l-1} p_2 l/k$.

A.2. Derivation of the conditional CDF within Eq. (4)

In principle, we should distinguish four cases, $Y_{k,0} = 1$, $Y_{k,1} = 1$, $Y_{k,0} = 0$, $Y_{k,1} = 0$. By Assumption A1, in case of $Y_{k,0} = 0$ or $Y_{k,1} = 0$, the RT is the processing time of l items plus a residual time τ_0, which does not involve collecting and analyzing visual information.

In case of $Y_{k,0} = 1$ or $Y_{k,1} = 1$, the conditional RT given $Y_{k,q} = 1$ is — in contrast to the event $Y_{k,0} = 0$ or $Y_{k,1} = 0$ — a finite mixture distribution consisting of the CDFs of the jth departure epoch, $j \leq l$. The weights $v_{j,k,q}$ are the conditional probabilities that the queue terminates at the jth departure epoch given $Y_{k,q} = 1$.

Consider first the case $Y_{k,0} = 1$. Since there are only distractors in the display, the response is a false alarm, i.e., $\mathbb{P}(Y_{k,0} = 1) = \pi_1$ by definition of π_1. The search terminates at the jth departure epoch if and only if the first $j - 1$ departing items are correctly classified as distractors and the jth departing distractor is misidentified as the target. Therefore, the probability that the queue terminates at the jth departure epoch given $Y_{k,0} = 1$ is

$$v_{j,k,0} = (1 - p_1)^{j-1} p_1/\pi_1(k; l, p_1, p_2).$$

Now, consider the case $Y_{k,1} = 1$. Here, $\mathbb{P}(Y_{k,1} = 1) = 1 - \pi_2$ by definition of π_2. We have to distinguish three cases:

1. The target is correctly identified at position $j \leq l$ in the departure order.
   Since equiprobable arrival orders imply equiprobable departure orders, the target appears at position j of the departing sequence with probability $1/k$. As the target as well as the $j - 1$ distractors must be classified correctly, the probability that the queue terminates at the jth departure epoch with the event $Y_{k,1} = 1$ is
   $$k^{-1}(1 - p_1)^{j-1}(1 - p_2).$$

2. The target is misidentified at a position $i < j$ in the departure sequence and a distractor is misidentified at position j.
   Then there are $j-1$ possible positions for the target and $j-2$ distractors must have been identified correctly. This gives
   $k^{-1}(j-1)p_2(1-p_1)^{j-2}p_1$.
3. The target is in a position $i > j$ and a distractor is misidentified at position j.
   Then the target may take $k-j$ possible positions and $j-1$ distractors have to be identified correctly. This gives
   $k^{-1}(k-j)(1-p_1)^{j-1}p_1$.

A.3. Derivation of the mean reward rate, Eq. (5), and the solution to the optimal stopping problem, Eq. (7)

In accordance with the notation usage in the main text, $Y_{k,q} = 1$ and $Y_{k,q} = 0$ denote the TP and TA responses, respectively, observed on trials with set size k and objective target status q. Throughout the derivation, we assume that the misidentification probabilities p_1 and p_2 are zero. This is in accordance with the presumption that observers decide about their optimal stopping policy in ignorance of their own genuine processing errors.

First, consider set size k as given and fixed. Equiprobable arrival orders imply equiprobable departure orders due to the properties of permutations. Hence, on target-present trials, any departure position i of the target is equiprobable. Under the assumption $p_1 = p_2 = 0$ and in case of a target-present trial, the probability of observing a TP response under the stopping policy $l = l(k)$ is

$$\begin{aligned}&\mathbb{P}(Y_{k,1}=1)\\&\quad=\sum_{i=1}^{l}\mathbb{P}\left(\text{the } i\text{th departure is target} \mid q=1\right)=\sum_{i=1}^{l}\frac{1}{k}=\frac{l}{k}\\&\quad=1-\mathbb{P}(Y_{k,1}=0).\end{aligned}$$

Let $N_k(l)$ denote the random variable of the number of items examined until termination, regardless of the outcome response under the stopping policy l on a trial with set size k. For $l \le k$, the expected value of $N_k(l)$ on target-present trials is then

$$\begin{aligned}&\mathbb{E}\left(N_k(l) \mid q=1\right)\\&\quad=\sum_{i=1}^{l}\mathbb{P}\left(\text{quitting after examining } i \text{ items} \mid q=1\right)\cdot i\\&\quad=\sum_{i=1}^{l} i\cdot\frac{1}{k}+l\cdot\mathbb{P}(Y_{k,1}=0)=\frac{2kl-l^2+l}{2k}.\end{aligned}$$

On target-absent trials, due to the assumed absence of genuine processing errors, processing will always terminate with a correct target-absent response after searching l items: $\mathbb{P}(Y_{k,0}=0)=1$ and $\mathbb{P}(Y_{k,0}=1)=0$. Accordingly, the expected value of $N_k(l)$ on target-absent trials is $\mathbb{E}(N_k(l) \mid q=0)=l$. Then the expected value of $N_k(l)$ on any trial can be calculated:

$$\begin{aligned}&\mathbb{E}\left(N_k(l)\right)\\&\quad=\mathbb{E}\left(N_k(l) \mid q=1\right)\cdot\mathbb{P}\left(q=1\right)+\mathbb{E}\left(N_k(l) \mid q=0\right)\cdot\mathbb{P}\left(q=0\right)\\&\quad=\frac{2kl-l^2+l}{2k}\cdot\frac{1}{2}+l\cdot\frac{1}{2}=\frac{4kl-l^2+l}{4k}.\end{aligned}$$

Let the random variable $R_k(l)$ equal 1 if the response is correct under the stopping policy l on a trial with k items. Otherwise, $R_k(l)$ shall be 0. Then the expected value of $R_k(l)$ under a target prevalence of 50% is given by

$$\begin{aligned}&\mathbb{E}\left(R_k(l)\right)\\&\quad=\mathbb{P}\left(Y_{k,1}=1\right)\cdot\mathbb{P}\left(q=1\right)+\mathbb{P}\left(Y_{k,0}=0\right)\cdot\mathbb{P}\left(q=0\right)\\&\quad=\frac{l}{k}\cdot\frac{1}{2}+1\cdot\frac{1}{2}=\frac{l+k}{2k}.\end{aligned}$$

Assume that the experiment has m set size levels $\boldsymbol{k} = (k_1,\dots,k_m)$ with $k_1 < k_2 < \cdots < k_m$. Consider the set size k as a random variable that takes the m values $k_1,\dots,k_m$ with equiprobability $1/m$. Assume further that the observer adopts stopping policies $\boldsymbol{l} := (l_1,\dots,l_m)$ for the respective levels. For a single random trial, we denote by $R_{\boldsymbol{k}}(\boldsymbol{l})$ the indicator function of a correct response and by $N_{\boldsymbol{k}}(\boldsymbol{l})$ the number of items searched until search termination. Then, the expected numbers equal

$$\mathbb{E}\left(R_{\boldsymbol{k}}(\boldsymbol{l})\right)=\frac{1}{m}\sum_{i=1}^{m}\mathbb{E}\left(R_{k_i}(l_i)\right);$$

$$\mathbb{E}\left(N_{\boldsymbol{k}}(\boldsymbol{l})\right)=\frac{1}{m}\sum_{i=1}^{m}\mathbb{E}(N_{k_i}(l_i)).$$

Let $\gamma_{\boldsymbol{k}}(\boldsymbol{l})$ be the mean reward rate in a long series of repeated, independent trials. More precisely, with respect to the weak law of large numbers, define

$$\gamma_{\boldsymbol{k}}(\boldsymbol{l})=\frac{\mathbb{E}\left(R_{\boldsymbol{k}}(\boldsymbol{l})\right)}{\mathbb{E}\left(N_{\boldsymbol{k}}(\boldsymbol{l})\right)}.$$

Hence,

$$\gamma_{\boldsymbol{k}}(\boldsymbol{l})=\frac{\sum_{i=1}^{m}\frac{l_i+k_i}{2k_i}}{\sum_{i=1}^{m}\frac{4k_il_i-l_i^2+l_i}{4k_i}},$$

and our optimization problem reads

$$\frac{\sum_{i=1}^{m}\frac{4k_il_i-l_i^2+l_i}{k_i}}{\sum_{i=1}^{m}\frac{l_i+k_i}{k_i}}=\min_{0\le l_i\le k_i, mz\le\sum_{i=1}^{m}\frac{l_i+k_i}{2k_i}}!,$$

where min! indicates minimization of the function on the left side with respect to $(l_1,\dots,l_m)$, subject to the constraints specified underneath, cf. Eq. (6). With $r_i := l_i/k_i \in [0,1]$, this is equivalent to

$$f(\boldsymbol{r}):=\frac{\sum_{i=1}^{m}(4k_ir_i-k_ir_i^2+r_i)}{m+\sum_{i=1}^{m}r_i}=\min_{0\le r_i\le 1,(2z-1)m\le\sum_{i=1}^{m}r_i}!,$$

where $\boldsymbol{r}=(r_1,\dots,r_m)$.

To solve the optimization problem, we first show that $f(\omega\boldsymbol{r}) < f(\boldsymbol{r})$ for $\omega\in[0,1)$ and $r_i\in[0,1]$ with $\boldsymbol{r}\neq 0$. Let $\overline{\boldsymbol{u}} := m^{-1}\sum_{i=1}^{m}u_i$ for any m-dimensional vector $\boldsymbol{u}$. Then,

$$f(\boldsymbol{r})=\frac{4\overline{\boldsymbol{kr}}-\overline{\boldsymbol{kr}^2}+\overline{\boldsymbol{r}}}{1+\overline{\boldsymbol{r}}},$$

where $\boldsymbol{kr}=(k_1r_1,\dots,k_mr_m)$, $\boldsymbol{kr}^2=(k_1r_1^2,\dots,k_mr_m^2)$, and

$$\begin{aligned}&(4\omega\overline{\boldsymbol{kr}}-\omega^2\overline{\boldsymbol{kr}^2}+\omega\overline{\boldsymbol{r}})(1+\overline{\boldsymbol{r}})<(4\overline{\boldsymbol{kr}}-\overline{\boldsymbol{kr}^2}+\overline{\boldsymbol{r}})(1+\omega\overline{\boldsymbol{r}})\\\Leftrightarrow\;&4\omega\overline{\boldsymbol{kr}}-\omega^2\overline{\boldsymbol{kr}^2}+\omega\overline{\boldsymbol{r}}-\omega^2\overline{\boldsymbol{kr}^2}\cdot\overline{\boldsymbol{r}}<4\overline{\boldsymbol{kr}}-\overline{\boldsymbol{kr}^2}+\overline{\boldsymbol{r}}-\omega\overline{\boldsymbol{kr}^2}\cdot\overline{\boldsymbol{r}}\\\Leftrightarrow\;&0<4(1-\omega)\overline{\boldsymbol{kr}}+(1-\omega)\overline{\boldsymbol{r}}-(\omega-\omega^2)\overline{\boldsymbol{kr}^2}\cdot\overline{\boldsymbol{r}}-(1-\omega^2)\overline{\boldsymbol{kr}^2}\\\Leftrightarrow\;&0<4(1-\omega)\overline{\boldsymbol{kr}}+(1-\omega)\overline{\boldsymbol{r}}-(1-\omega)\omega\overline{\boldsymbol{kr}^2}\cdot\overline{\boldsymbol{r}}\\&\qquad-(1-\omega)(1+\omega)\overline{\boldsymbol{kr}^2}\\\Leftrightarrow\;&0<4\overline{\boldsymbol{kr}}+\overline{\boldsymbol{r}}-(1+\omega+\omega\overline{\boldsymbol{r}})\overline{\boldsymbol{kr}^2}.\end{aligned}$$

As $\overline{\boldsymbol{kr}^2}\le\overline{\boldsymbol{kr}}$, we have

$$4\overline{\boldsymbol{kr}}+\overline{\boldsymbol{r}}-(1+\omega+\omega\overline{\boldsymbol{r}})\overline{\boldsymbol{kr}^2}>4\overline{\boldsymbol{kr}}+\overline{\boldsymbol{r}}-3\overline{\boldsymbol{kr}}=\overline{\boldsymbol{kr}}+\overline{\boldsymbol{r}}>0$$

and the inequality $f(\omega \boldsymbol{r}) < f(\boldsymbol{r})$ is shown. Furthermore, $f(\boldsymbol{r}) \geq 0$ and $f(\boldsymbol{r}) = 0$ if and only if $\boldsymbol{r} = 0$. Now, it follows readily by contradiction, that the equality $\boldsymbol{r} = 0$ or $(2z-1)m = \sum_{i=1}^{m} r_i$ holds in the optimum. Hence, it suffices to consider

$$f^*(\boldsymbol{r}) := \sum_{i=1}^{m} k_i r_i (4 - r_i)$$

instead of $f(\boldsymbol{r})$. Let $\boldsymbol{r}^* := (r_1^*, \ldots, r_n^*) \neq 0$, $r_i^* \in [0,1]$, be a solution of the minimization problem, and $\varphi_i = r_i^*(4 - r_i^*)$. Assume that some $j > i$ exist with $r_j^* > r_i^*$. Then $\varphi_j > \varphi_i$, hence $k_i\varphi_j + k_j\varphi_i < k_i\varphi_i + k_j\varphi_j$, which is a contradiction to $\boldsymbol{r}^*$ being optimal. Assume now, that $0 < r_j^* \leq r_i^* < 1$ for some $j > i$. Then, for any $0 < \varepsilon \leq \min\{r_j^*, 1 - r_i^*\}$, we have

$$k_i(r_i^* + \varepsilon)(4 - r_i^* - \varepsilon) + k_j(r_j^* - \varepsilon)(4 - r_j^* + \varepsilon) < k_i\varphi_i + k_j\varphi_j,$$

which is again a contradiction to $\boldsymbol{r}^*$ being optimal. Consequently,

$$r_i^* = \begin{cases} 1, & i \leq (2z-1)m, \\ (2z-1)m - \lfloor (2z-1)m \rfloor, & i = \lfloor (2z-1)m \rfloor + 1, \\ 0, & i > \lfloor (2z-1)m \rfloor + 1, \end{cases}$$

where $\lfloor x \rfloor$ is the largest integer not greater than x for any $x \in \mathbb{R}$. Note that if $z > 1 - (2m)^{-1}$, then all r_i^* are positive.

A.4. Derivation of p_1 and p_2, Eqs. (8) and (9)

We assume here that the optimal strategy l^* is applied and that z is very close to one. Hence, $l^*(k)$ can be considered to be equal to k for all k.

Now, Eq. (1) states that

$$\pi_1 = 1 - (1 - p_1)^{l^*},$$

replacing l by l^*. According to pattern (b), $\pi_1 \approx h_1(k)$ for some function $h_1(k)$, which is constant in k or slightly decreasing. Solving for p_1 results in

$$p_1(k) = 1 - (1 - h_1(k))^{1/k}.$$

Since π_1, hence $h_1(k)$, is close to zero in empirical data of visual search, the Taylor expansion about the point $h_1(k) = 0$, see Eq. 1.110 in Gradshteyn and Ryzhik (2000), yields

$$(1 - h_1(k))^{1/k} \approx 1 - \frac{h_1(k)}{k}$$

(with only a very small error). Hence,

$$p_1(k) \approx \frac{h_1(k)}{k}.$$

Now, h_1 can be slightly decreasing with increasing k. A simple, natural approach to model this fact is to assume that $h_1(k) = a_1 k^{-\varepsilon}$ for some $\varepsilon > 0$ and $a_1 > 0$. Accordingly, we model p_1 by

$$p_1(k) \approx a_1 k^{-b}$$

for some $b \geq 1$ and $a_1 > 0$.

Fig. 6 confirms greatly our approach and has been obtained as follows. For each k, we first obtain a rough estimate of $\tilde{l}(k)$ as specified in Eq. (10) in Section 2.5 based on the visual fit to the empirical miss rate. Then we substitute this estimate in $\mathbb{P}(Y_{k,0} = 1) = 1 - (1-p_1)^{k(1-\exp(\alpha+\beta k))}$ and solve it for p_1. The naive estimates of p_1 obtained in this way appear to follow a power law of set size k for both data sets, as Fig. 6 shows. Moreover, since the rough estimates of $\tilde{l}(k)$ turn out to satisfy $\tilde{l}(k) < k$, this observation indicates that the explanatory power of the power law relationship is not restricted by the replacement of l with k in the derivation.

Note that a mathematical derivation of the behavior of p_2 is not possible without further assumptions, since, for $l = l^* = k$ in Eq. (2),

$$\pi_2 = \left(1 - a_1 k^{-b}\right)^{k-1} p_2.$$

Numerical evaluations of $\left(1 - a_1 k^{-b}\right)^{k-1}$ show that this term (which is less than, but close to 1) is approximately constant in k if all parameters are in the range of interest (i.e., if a_1 is close to 0, k between 3 and 18 and b not much greater than 1), that is, when $p_1 = a_1 k^{-b}$ is close to 0. Note that the ranges considered here are in line with typical data in visual search experiments. It follows that $\pi_2 = h_2(k)p_2$ for some $h_2(k)$ which is roughly constant in k, and $h_2(k)p_2$ is roughly independent of p_1. Hence, based on this alone, it is not possible to infer the form of p_2 uniquely because several functions can produce pattern (a), (c), and (d), which involve characteristics of π_2.

Appendix B. Algorithm for the simulation of a queueing process

Algorithm 1 (part 1): Simulation of a system response

Input: $(\alpha, \beta, a_1, a_2, b, k, q)$
Output: (response, termination)
set $l = k(1 - \exp(\alpha + \beta k))$;
generate $\Delta L \sim \text{Bernoulli}\left(l - \lfloor l \rfloor\right)$;
set $L = \lfloor l \rfloor + \Delta L$;
set $p_1 = a_1 k^{-b}$, $p_2 = a_2 k^b$;
generate $G_{1,\ldots,k} \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p_1)$, $H \sim \text{Bernoulli}\left(p_2\right)$;
if q = FALSE **then**
  if $G_i = 0 \; \forall i = 1, \ldots, L$ **then**
    set termination $= L$;
    set response = FALSE ;
  else
    set termination $= \min\{j \mid G_j = 1\}$;
    set response = TRUE ;
  end if
else
  generate $P_{\text{target}} \sim \text{unif}\left\{1, \ldots, k\right\}$;
  set $G_{P_{\text{target}}} = 0$;
  if $G_i = 0 \; \forall i = 1, \ldots, L$ **then**
    if $H = 0$ *and* $P_{\text{target}} \leq L$ **then**
      set termination $= P_{\text{target}}$;
      set response = TRUE ;
    else
      set termination $= L$;
      set response = FALSE ;
    end if
  else
    set response = TRUE ;
    if $H = 0$ **then**
      set termination $= \min\{P_{\text{target}}, \min\{j \mid G_j = 1\}\}$;
    else
      set termination $= \min\{j \mid G_j = 1\}$;
    end if
  end if
end if

Algorithm 1 (part 2): Simulation of a system response time (main body)

Data structures:
input = $(\lambda, \mu, c, \tau_0, \tau_1, k,$ response, termination$)$
state = (t, n_A, n_D, n_S)
out = $\begin{pmatrix} A_1 & \dots & A_{n_D} & A_{n_D+1} & \dots & A_{n_A} \\ D_1 & \dots & A_{n_D} & +\infty & \dots & +\infty \end{pmatrix}$
event_list = $(t_A, t_{D_{S1}}, t_{D_{S2}}, \dots, t_{D_{Sc}})$
service_id = $(x_1, \dots, x_c)$
Output: (system_rt, response)

initialization
  set state = $(0, 0, 0, 0)$;
  set out = $\begin{pmatrix} \cdot \\ \cdot \end{pmatrix}$;
  set service_id = $(0, \dots, 0)$;
  set dep_id = 0, dep_server= 0 ;
  generate $I_1 \sim \text{Exp}(k\lambda)$;
  set $t_A = I_1$;
  set event_list = $(t_A, +\infty, \dots, +\infty)$;
  set STOP = FALSE ;
simulation loop
  while *not* STOP **do**
    if $t_A = \min\{ t_A, t_{D_{S1}}, t_{D_{S2}}, \dots, t_{D_{Sc}} \}$ **then**
      execute subroutine UPDATE_ARRIVAL ;
    else
      execute subroutine UPDATE_DEPARTURE ;
    end if
    if $t_A = t_{D_{S1}} = \cdots = t_{D_{Sc}} = +\infty$
    or (response = FALSE *and* $n_D \geq$ termination)
    or (response = TRUE *and* dep_id = termination)
     then
      STOP = TRUE ;
    end if
  end while
set system_rt= $t + \tau_0 +$ response $\cdot (\tau_1 - \tau_0)$;
set output = (system_rt, response) ;

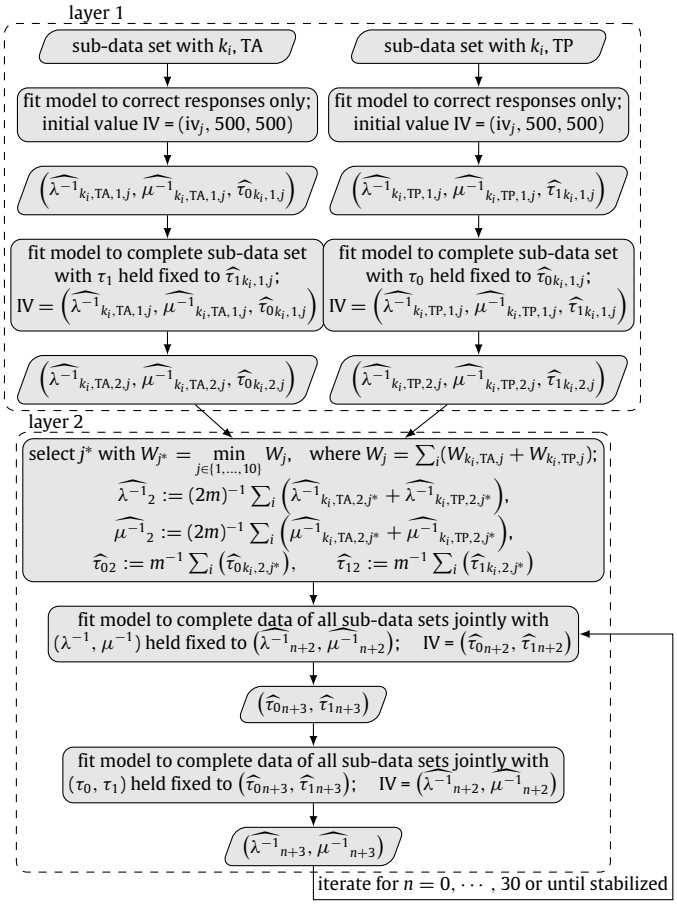
Procedure: UPDATE_ARRIVAL

set $t = t_A$, update n_A by $n_A + 1$, update n_S by $n_S + 1$;
set $A_{n_A} = t$, out = $\left(\text{out}, \begin{pmatrix} A_{n_A} \\ +\infty \end{pmatrix}\right)$;
if $n_A < k$ **then**
  generate $I_{n_A+1} \sim \text{Exp}\left((k - n_A)\lambda\right)$;
  set $t_A = t + I_{n_A+1}$;
else
  set $t_A = +\infty$;
end if
set free_servers = $\{ j \mid x_j = 0 \}$;
if $n_S \leq c$ *and* free_servers $\neq \emptyset$ **then**
  generate $S \sim \text{Exp}(\mu)$;
  set $i = \min$(free_servers) ;
  set $t_{D_{Si}} = t + S$, $x_i = n_A$;
end if

Procedure: UPDATE_DEPARTURE

set $t = \min\{ t_A, t_{D_{S1}}, t_{D_{S2}}, \dots, t_{D_{Sc}} \}$, n_D++, n_S-- ;
set $D_{n_D} = t$;
set dep_server = j such that
  $t_{D_{Sj}} = \min\{ t_A, t_{D_{S1}}, t_{D_{S2}}, \dots, t_{D_{Sc}} \}$;
set dep_id = $x_{\text{dep\_server}}$, $\text{out}_{2,\text{dep\_id}} = D_{n_D}$;
set $x_{\text{dep\_server}} = 0$, free_servers= $\{ j \mid x_j = 0 \}$;
if $n_S \geq c$ **then**
  set $i = \min$(free_servers) ;
  set $x_i = n_D + (c - |\text{free\_servers}|) + 1$;
  generate $S \sim \text{Exp}(\mu)$;
  set $t_{D_{Si}} = t + S$;
else
  set $t_{D_{Sj}} = +\infty \quad \forall j \in$ free_servers;
end if

Appendix C. Flow chart of the parameter estimation procedure



Appendix D. Densities of the total distances resulting from model uncertainty analysis

See Figs. D.14 and D.15.

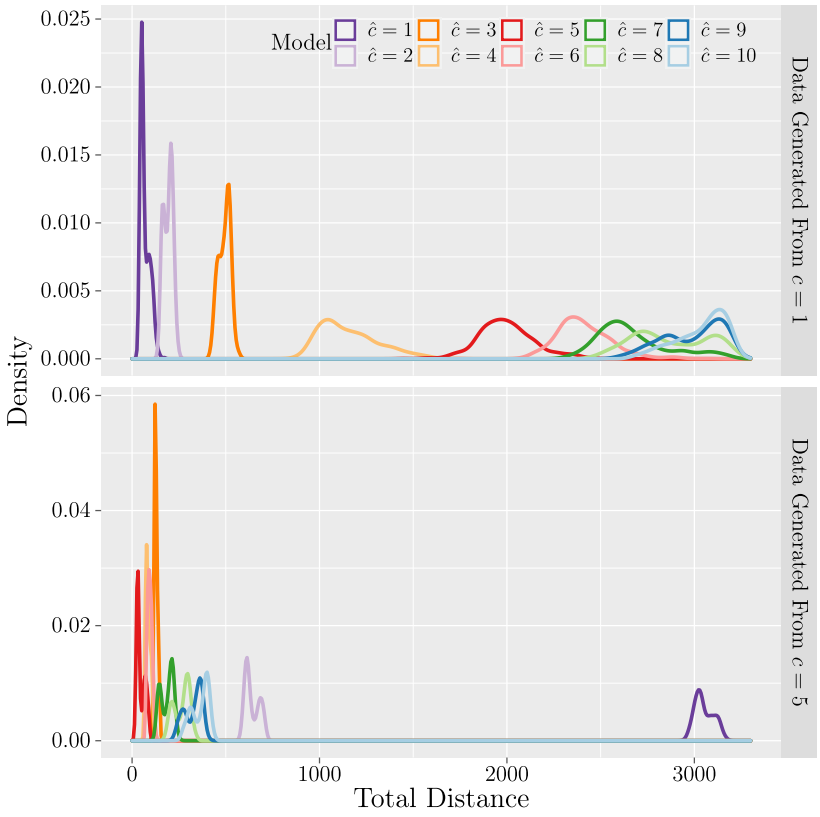


Fig. D.14. Densities of the total distances resulting from fitting the queueing models with $c = 1, \ldots, 10$ to 600 data sets simulated by a single-server system (top) and a five-server system (bottom). The constraint $\mu_3 = \mu_6 = \mu_{12} = \mu_{18}$ was applied to both the models that generated the data and the models that were fitted to the simulated data.

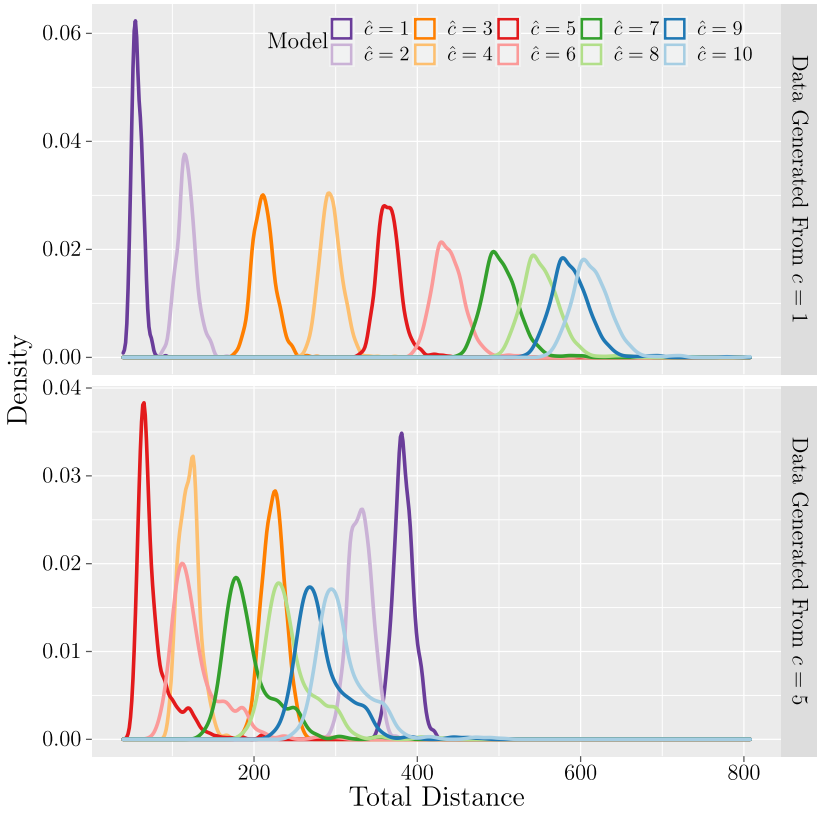


Fig. D.15. Densities of the total distances resulting from fitting the queueing models with $c = 1, \ldots, 10$ to 618 data sets simulated by a single-server system (top) and a five-server system (bottom). The service rate μ was allowed to vary across set sizes for both the models that generated the data and the models that were fitted to the simulated data.

References

Basu, A., Shioya, H., & Park, C. (2011). *Statistical inference: The minimum distance approach* (1st ed.). CRC Press. http://dx.doi.org/10.1201/b10956.

Bhat, U. N. (2015). *An introduction to queueing theory: Modeling and analysis in applications* (2nd ed.). Birkhäuser. http://dx.doi.org/10.1007/978-0-8176-8421-1.

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

Corbett, E. A., & Smith, P. L. (2020). A diffusion model analysis of target detection in near-threshold visual search. *Cognitive Psychology*, *120*, Article 101289. http://dx.doi.org/10.1016/j.cogpsych.2020.101289.

Cousineau, D., & Shiffrin, R. M. (2004). Termination of a visual search with large display size effects. *Spatial Vision*, *17*(4–5), 327–352. http://dx.doi.org/10.1163/1568568041920104.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–185. http://dx.doi.org/10.1017/S0140525X01003922.

Donkin, C., & Shiffrin, R. M. (2011). Visual search as a combination of automatic and attentive processes. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *33*, 2830–2835. https://escholarship.org/uc/item/50r598fx.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. http://dx.doi.org/10.1037/0033-295X.96.3.433.

Duncan, J., Ward, R., & Shapiro, K. L. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, *369*(6478), 313–315. http://dx.doi.org/10.1038/369313a0.

Fisher, D. L. (1982). Limited-channel models of automatic detection: Capacity and scanning in visual search. *Psychological Review*, *89*(6), 662–692. http://dx.doi.org/10.1037/0033-295X.89.6.662.

Fisher, D. L. (1984). Central capacity limits in consistent mapping, visual search tasks: Four channels or more? *Cognitive Psychology*, *16*(4), 449–484. http://dx.doi.org/10.1016/0010-0285(84)90017-3.

Geng, J. J., & Behrmann, M. (2003). Selective visual attention and visual search: Behavioral and neural mechanisms. In D. E. Irwin, & B. H. Ross (Eds.), *The psychology of learning and motivation: Vol. 42. Cognitive vision* (1st ed., pp. 157–191). Academic Press. http://dx.doi.org/10.1016/S0079-7421(03)01005-3.

Gradshteyn, I., & Ryzhik, I. (2000). *Table of integrals, series, and products* (6th ed.). Academic Press.

Gross, D., Shortle, J. F., Thompson, J. M., & Harris, C. M. (2008). *Fundamentals of queueing theory* (4th ed.). Wiley. http://dx.doi.org/10.1002/9781118625651.

Harris, J. R., Shaw, M. L., & Bates, M. (1979). Visual search in multicharacter arrays with and without gaps. *Perception and Psychophysics*, *26*(1), 69–84. http://dx.doi.org/10.3758/BF03199862.

Hoffman, J. E. (1978). Search through a sequentially presented visual display. *Perception and Psychophysics*, *23*(1), 1–11. http://dx.doi.org/10.3758/BF03214288.

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

Isbell, E., Fukuda, K., Neville, H. J., & Vogel, E. K. (2015). Visual working memory continues to develop through adolescence. *Frontiers in Psychology*, *6*, Article 696. http://dx.doi.org/10.3389/fpsyg.2015.00696.

Klein, R. M. (1988). Inhibitory tagging system facilitates visual search. *Nature*, *334*(6181), 430–431. http://dx.doi.org/10.1038/334430a0.

Li, Y. (2019). *Visual search as a queueing process* [Doctoral dissertation, Universität Mannheim]. MADOC. https://madoc.bib.uni-mannheim.de/53040.

Liu, Y. (1996). Queueing network modeling of elementary mental processes. *Psychological Review*, *103*(1), 116–136. http://dx.doi.org/10.1037/0033-295X.103.1.116.

Liu, Y. (2013). Queuing and network models. In J. D. Lee, & A. Kirlik (Eds.), *The Oxford handbook of cognitive engineering* (pp. 449–464). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199757183.013.0031.

Liu, Y., Feyen, R., & Tsimhoni, O. (2006). Queueing network-model human processor (QN-MHP): A computational architecture for multitask performance in human-machine systems. *ACM Transactions on Computer-Human Interaction*, *13*(1), 37–70. http://dx.doi.org/10.1145/1143518.1143520.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. http://dx.doi.org/10.1038/36846.

McNamara, J. (1982). Optimal patch use in a stochastic environment. *Theoretical Population Biology*, *21*(2), 269–288. http://dx.doi.org/10.1016/0040-5809(82)90018-1.

Miller, J. (1993). A queue-series model for reaction time, with discrete stage and continuous flow models as special cases. *Psychological Review*, *100*(4), 702–715. http://dx.doi.org/10.1037/0033-295x.100.4.702.

Moore, C. M., Egeth, H. E., Berglan, L. R., & Luck, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin and Review*, *3*(3), 360–365. http://dx.doi.org/10.3758/BF03210761.

Moore, C. M., & Wolfe, J. M. (2001). Getting beyond the serial/parallel debate in visual search: A hybrid approach. In K. L. Shapiro (Ed.), *The limits of attention: Temporal constraints on human information processing* (pp. 178–198). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198505150.003.0009.

Moran, R., Zehetleitner, M., Liesefeld, H. R., Müller, H. J., & Usher, M. (2016). Serial vs. parallel models of attention in visual search: Accounting for benchmark RT-distributions. *Psychonomic Bulletin and Review*, *23*(5), 1300–1315. http://dx.doi.org/10.3758/s13423-015-0978-1.

Moran, R., Zehetleitner, M., Müller, H. J., & Usher, M. (2013). Competitive guided search: Meeting the challenge of benchmark RT distributions. *Journal of Vision*, *13*(8), 1–31. http://dx.doi.org/10.1167/13.8.24.

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

Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics*, *44*(4), 369–378. http://dx.doi.org/10.3758/bf03210419.

Protter, P. E. (2005). *Stochastic integration and differential equations* (2nd ed.). Springer. http://dx.doi.org/10.1007/978-3-662-10061-5.

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. http://dx.doi.org/10.1037/0033-295X.85.2.59.

Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review*, *92*(2), 212–225. http://dx.doi.org/10.1037/0033-295X.92.2.212.

Schlather, M., & Li, Y. (2022). *qmvs: Queueing model of visual search* (Version 0.1.1) [R package]. CRAN. https://CRAN.R-project.org/package=qmvs.

Severini, T. A., & Wong, W. H. (1992). Profile likelihood and conditionally parametric models. *The Annals of Statistics*, *20*(4), 1768–1802. http://dx.doi.org/10.1214/aos/1176348889.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton University Press. http://dx.doi.org/10.2307/j.ctvs32s6b.

Stewart, W. J. (2009). *Probability, Markov chains, queues, and simulation: The mathematical basis of performance modeling*. Princeton University Press. http://dx.doi.org/10.1515/9781400832811.

Theeuwes, J., Godijn, R., & Pratt, J. (2004). A new estimation of the duration of attentional dwell time. *Psychonomic Bulletin and Review*, *11*(1), 60–64. http://dx.doi.org/10.3758/BF03206461.

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

Townsend, J. T. (1971). A note on the identifiability of parallel and serial processes. *Perception and Psychophysics*, *10*(3), 161–163. http://dx.doi.org/10.3758/BF03205778.

Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*(1), 46–54. http://dx.doi.org/10.1111/j.1467-9280.1990.tb00067.x.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, *39*(4), 321–359. http://dx.doi.org/10.1006/jmps.1995.1033.

Treisman, A. M. (1986). Features and objects in visual processing. *Scientific American*, *255*(5), 114B–125. http://dx.doi.org/10.1038/scientificamerican1186-114b.

Treisman, A. M. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology Section A*, *40A*(2), 201–237. http://dx.doi.org/10.1080/02724988843000104.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. http://dx.doi.org/10.1016/0010-0285(80)90005-5.

Vallender, S. (1974). Calculation of the Wasserstein distance between probability distributions on the line. *Theory of Probability and Its Applications*, *18*(4), 784–786. http://dx.doi.org/10.1137/1118101.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*(2), 202–238. http://dx.doi.org/10.3758/BF03200774.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*(1), 33–39. http://dx.doi.org/10.1111/1467-9280.00006.

Wolfe, J. M. (2012). When do I quit? The search termination problem in visual search. In M. D. Dodd, & J. H. Flowers (Eds.), *The influence of attention, learning, and motivation on visual search* (pp. 183–208). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-4794-8_8.

Wolfe, J. M. (2018). Visual search. In J. T. Wixted, & J. Serences (Eds.), *Steven's handbook of experimental psychology and cognitive neuroscience: Vol 2. Sensation, perception, and attention* (4th ed., pp. 569–623). Hoboken, NJ: Wiley, http://dx.doi.org/10.1002/9781119170174.epcn213.

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin and Review*, *28*(4), 1060–1092. http://dx.doi.org/10.3758/s13423-020-01859-9.

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

Wu, C., & Liu, Y. (2008). Queuing network modeling of the psychological refractory period (PRP). *Psychological Review*, *115*(4), 913–954. http://dx.doi.org/10.1037/a0013123.

Zenger, B., & Fahle, M. (1997). Missed targets are more frequent than false alarms: A model for error rates in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(6), 1783–1791. http://dx.doi.org/10.1037/0096-1523.23.6.1783.