Check for
updates

A theoretical attempt to revive the serial/parallel-search dichotomy

Heinrich René Liesefeld [1,2] · Hermann J. Müller [1]

Abstract
A core distinction in Anne Treisman's feature-integration theory (FIT) is in that between parallel and serial search. We outline this dichotomy and selectively review the reasons why it has largely been abandoned in the visual-search community—namely, its theoretical dispensability, failure to find reliable yardsticks for differentiating parallel and serial search, and falsification of core predictions. We then go on to introduce a new theoretical framework that, we argue, clears up some of the theoretical confusion by merging FIT with various competing theories. This framework's core feature is the distinction between and characterization of two fundamentally different search modes: one in which attention is guided to a single item via a priority map (*priority guidance*), and one in which clumps of multiple items are scanned in parallel in a spatially systematic order (*clump scanning*). Finally, we will elaborate how this new theoretical framework can resolve current controversies in the literature and how it relates to other existing theories. We (somewhat optimistically) believe that the outcome of this theoretical exercise is a unification of theories of visual search that can explain, or at least is consistent with, all phenomena reported in the visual-search literature that have previously been accounted for by various conflicting theories.

Keywords Selective attention · Theoretical unification · Serial vs. parallel processing · Visual search · Dimension-weighting account

A core assumption of Treisman's feature-integration theory (FIT; Treisman & Gelade, 1980) is that visual search is either parallel or serial. That is, the presence of a certain, distinguishing target feature is registered via parallel processing of all objects in view (*parallel search*), or the focus of spatial attention (akin to a "spotlight") is shifted from one (clump of) object(s) to the next until the target is found or until all objects have been inspected and rejected (*serial search*). The idea of these two types of search has strong intuitive appeal because both are self-evidently part of everyday life: One often finds a sought-for object by just looking into its rough direction, without investing much effort (*parallel search*), such as the single red-cased book on one's bookshelf full of black-cased books. But then there are also those searches for the proverbial needle in the haystack, where we have to scrutinize one piece on the stack after the other (*serial search*), producing massive costs in terms of effort and time (even if the needle lies in plain view on top of the haystack). As detailed in turn, both strategies have largely complementary advantages and disadvantages, so that, contingent on the specific search conditions, the observer is well advised to resort to one or the other strategy.

According to Treisman, *parallel search* is highly efficient: The presence of a target feature (e.g., is there a red item anywhere in the display?) can be established directly, without allocating attention to any specific item. However, before attention is allocated, feature representations are "free floating" in dimensionally organized modules—that is, features are not yet assigned to specific objects. Accordingly, observers cannot tell which features belong to the same or different objects and they might perceive illusory conjunctions (e.g., a red cross when there are actually only red hearts and blue crosses). This renders parallel search overly prone to error when the target is defined by a conjunction of features (e.g., red *and* cross), because the observer cannot know whether the percept of a red cross reflects reality or is just an illusory conjunction.

This issue is resolved in *serial search*, where spatially focused attention ensures that features are bound into localized objects. On the down side, in serial search, only one (Treisman

✉ Heinrich René Liesefeld
  Heinrich.Liesefeld@psy.lmu.de

[1] Department Psychologie, Ludwig-Maximilians-Universität, Leopoldstr. 13, D-80802 Munich, Germany

[2] Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität, München, Germany

& Gelade, 1980) or a few (Treisman & Souther, 1985) objects are processed at once, rendering search effortful and slow (inefficient).

One hallmark of inefficient search is that an increase in the number of presented objects yields an increase in search times. This set-size effect can be quantified as the slope of the function relating search RT to the number of items in the display. Its steepness has become the main empirical criterion for discriminating serial from parallel search. Accordingly, the main empirical pattern explained by (and taken to support) FIT was that search for feature targets produced flat search slopes and search for conjunction targets steep slopes.

As detailed below, while the serial/parallel dichotomy is useful for describing everyday experience and had some success in explaining certain data patterns, it has largely been dispensed with as an explanatory conceptualization of visual search, for three main reasons: (a) theoretically more parsimonious theories were successful in explaining the same data patterns; (b) core predictions from FIT were falsified; and (c) search slopes turned out to be unreliable yardsticks for classifying searches as serial versus parallel. Nevertheless, we suggest a new look at the serial/parallel dichotomy that integrates competing theories, and we will attempt to demonstrate that this unified theoretical framework can resolve several current controversies in the literature. Given the (hidden) indication of two qualitatively different types of search reviewed (and reinterpreted) here, we conclude that it would be unwise to ultimately reject such a dichotomy.

Challenges to the dichotomy

Theoretical challenge: The dispensability of a search dichotomy

"To the same natural effects we must, as far as possible, assign the same causes" (Newton, 1846, p. 384). This general scientific heuristic, which has become known as *Ockham's razor*, reminds us to adhere to the least complex theory that can still explain all observed data patterns within the theory's scope. Accordingly, theories that can explain all kinds of searches (same scope) without assuming any dichotomy (less complexity) cast considerable doubt on the serial/parallel dichotomy. We will briefly sketch the currently most influential of these (potentially) more parsimonious theories next. Of note, none of these theories assumes a search dichotomy along the lines of Treisman's FIT, though some of them assume two or more successive *stages* of processing.

In guided search (GS; Wolfe, 1994, 2007; Wolfe, Cave, & Franzel, 1989), an initial, parallel processing of the visual scene yields a map-type representation coding for the relative importance of each item in the scene (while Wolfe uses the term *activation map*, *priority map* has recently become more prevalent). This map *guides* subsequent allocations of attention, so that spatial attention is not deployed at random, but selectively only to promising items. The quality of guidance varies depending on the discriminability of the target, so that smaller or larger subsets of all items are inspected. Based on this assumption, GS can produce all kinds of search slopes without assuming a search dichotomy. Crucially, despite the strong serial aspect (which shares critical features with serial-search conceptions in later versions of FIT; Treisman & Souther, 1985, see below), GS can also explain flat search slopes: If the priority of the target is so high that it is always selected first for focal-attentional inspection, the number of nontargets has no effect on search times. Subjectively, the target *pops out* of the display. According to a more recent version of GS, a few objects can be processed in parallel after having (sequentially) passed the attentional bottleneck (carwash model; Moore & Wolfe, 2001; Wolfe, 2007).

Theories based on a strategically modifiable *attentional window* (e.g., Humphreys & Müller, 1993; Theeuwes, 1994; Treisman & Souther, 1985) take, in a way, the opposite approach to explain search slopes: They assume that only the subset of items within the attentional window is processed in parallel and the size of the window adapts to target discriminability to maintain an acceptable signal-to-noise-ratio (see below). Smaller windows require more reallocations, in particular with larger displays, thus yielding steep search slopes. With larger windows, search slopes become flatter until the window encompasses the whole display, in which case all objects are processed at the same time. The size of the window is—similar to the guidance in GS—determined by the difficulty of discriminating the target from the nontargets.

Related, but conceptually different from the attentional window, is the *functional viewing field* (FVF) in the model of Hulleman and Olivers (2017; see also Engel, 1977; Geisler & Chou, 1995). Rather than being strategically modifiable, its size is an outcome of the inherent limitations of the system. The model assumes that multiple objects are processed in parallel within one eye fixation, with the number of objects that are processed to a sufficiently high degree being limited by declining retinal resolution and increasing neural competition towards the periphery. Targets that are harder to discriminate from distractors require higher resolution; in other words, the functional viewing field is smaller for less discriminable targets. With a smaller field, more fixations are needed and, consequently, search slopes increase. Arguably, this theory is theoretically more elegant than attentional-window models, in that the size of the functional viewing field emerges "naturally" rather than being set deliberately by the observer.

Note that GS, attentional-window models, and FVF models incorporate (various) parallel as well as serial aspects. In all cases, the movement of attention (or the eye) is serial. GS has a preattentive (and, in later versions, also an additional postselective) parallel stage, and attentional-window/FVF

models have a postselective and (might need; see Itti, 2017; Müller, Liesefeld, Moran, & Usher, 2017) an additional preattentive parallel stage. In fact, the phenomenon of pop-out cannot be explained without assuming some parallel processing across the whole display, and there is no doubt that visual information is initially processed in parallel in early visual areas.

Consequently, the most parsimonious accounts (in terms of the number of theoretical assumptions, though not necessarily in terms of the number of model parameters) must derive from theories of visual search that assume parallel processing of all objects in the display. These purely parallel theories come in two main flavors: with limited or unlimited capacity. With limited capacity, distribution of the limited resource across all items in the display results in slower processing of each individual item (Snodgrass & Townsend, 1980; Thornton & Gilden, 2007; Ward & McClelland, 1989). Unlimited-capacity parallel models are typically based on signal-detection theory, assuming that the signal (and therefore processing time and/or error rate) depends on target discriminability. Decisions might be based on individual items (Eckstein, Thomas, Palmer, & Shimozaki, 2000; Gardner, 1973; Eriksen & Spencer, 1969; Moran, Zehetleitner, Liesefeld, Müller, & Usher, 2016; Verghese, 2001), on a signal pooled across items (Kinchla, 1974), or on the overall-gist of the scene (*summary statistics*; Rosenholtz, Huang, Raj, Balas, & Ilie, 2012; Treisman & Souther, 1985).

The latter two possibilities have the potential disadvantage that information about the target's location necessarily gets lost during integration and would have to be recovered by some additional mechanism if the task requires localization of the target. This would imply that localizing the target takes longer than detecting the target, which is exactly the opposite of what Zehetleitner and Müller (2010) found in a direct comparison of localization and detection speed. Accordingly, parallel processing of individual items appears to be the more comprehensive parallel account of visual search, because—provided that information on the location of the individual items is preserved (e.g., assuming that each parallel processor is associated with a particular location, as in modern graphic cards)—it would most readily also account for localization performance.

To keep the proportion of erroneous responses low despite increases in the number of nontargets, the decision criterion must be adapted, which is paid for by an increase in decision time (Moran et al., 2016; Palmer & McLean, 1995). In these models, if (due to time constraints) decision criteria are not adapted or accumulated information is insufficient, error rates increase (Eckstein et al., 2000; Eriksen & Spencer, 1969; Gardner, 1973; Kinchla, 1974; Verghese, 2001). Of note, even such purely parallel models can well account for steep RT search slopes and other aspects of the visual-search data in various ways (see below; Humphreys & Müller, 1993;

Moran et al., 2016). Furthermore, purely parallel models can theoretically even produce the subjective experience of seriality: Potentially an item reaches consciousness only after processing is completed, and even synchronously starting parallel identifications result in asynchronous completion times due to (random) variation in processing time.

Empirical challenge: Conjunctions can pop out and feature search can be inefficient

The most central and characteristic idea of FIT (Treisman & Gelade, 1980) is that single features can be processed preattentively, whereas the processing of conjunctions of features requires focal attention. Accordingly, searching for a target standing out in one feature dimension among any number of homogenous nontargets (*feature search*; e.g., a red bar among green bars) is easy and produces the subjective experience of pop-out, whereas searching for a target differing from the nontargets only in its specific combination of features (*conjunction search*; e.g., a red vertical bar among red horizontal and green vertical bars) is hard and feels effortful. The critical finding was that the number of nontargets did not influence response times in feature search, but did have a huge effect in conjunction search, producing steep search slopes.

However, it soon turned out that conjunction searches can produce flat search slopes as well (McLeod, Driver, & Crisp, 1988; Nakayama & Silverman, 1986; Steinman, 1987; Theeuwes, 1996; Theeuwes & Kooi, 1994; Wolfe et al., 1989) if the target features differ sufficiently from the nontarget features (see von Mühlenen & Müller, 2000). Additionally, even feature searches can produce steep slopes (Duncan & Humphreys, 1989; Liesefeld, Moran, Usher, Müller, & Zehetleitner, 2016; Nagy & Sanchez, 1990; Nothdurft, 1993; Roper, Cosman, & Vecera, 2013; Treisman & Gormican, 1988; Wolfe, Klempen, & Shulman, 1999) if the target feature is similar enough to the nontarget features (i.e., if the target is not salient). Thus, feature and conjunction search are not so different after all, but searches might differ on a continuum, depending on target–nontarget discriminability (Duncan & Humphreys, 1989; Liesefeld et al., 2016; Wolfe, 1998).

Practical challenge: Search slopes do not discriminate

The prime evidence for the existence of two search modes was the observation of flat (parallel) versus steep (serial) search slopes. However, there are two serious problems with this evidence: (a) It soon turned out that there is a continuum of search slopes in between flat and steep slopes (Liesefeld et al., 2016; Wolfe, 1998), invalidating the inference of a search-mode dichotomy from a search-slope dichotomy; (b) steep search slopes may be produced by all kinds of mechanisms, including some that are clearly serial or, respectively, parallel

in nature, and some that can involve aspects of either (for general issues in discriminating serial vs. parallel processes based on [mean] set-size effects and more promising alternative suggestions, see Townsend, 1971, 1972, 1976, 1990). The influences that we believe are most relevant to the debate are systematized below (though see also Palmer, 1995).

In (partly) serial-search models, both the number of inspections and the duration of each inspection (*dwell time*) can theoretically influence search slopes. The number of inspections, in turn, can depend on (a) how informed the decision to inspect a certain item is (*guidance*); (b) the criterion/ mechanism to stop searching and guess instead (*quitting*); (c) how many objects are processed in parallel (*window size*); and (d) how well the observer keeps track of already inspected locations (*memory*). Dwell time can depend on how long it takes (e) to recover control over the focus of attention after an incorrect allocation (*disengagement*); (f) to program and execute a shift of attention from A to B (*shifting*) and/or to identify and reject nontargets (*identification*). Regarding the latter factor, nontarget identification can take longer (g) because evidence accumulation is slowed (*accumulation rate*), (*h*) because more evidence is required to identify an object as a nontarget (*criterion*), or (i) because the tendency towards making a nontarget decision decreases (*bias*). This bias would likely decrease with the probability of actually processing nontargets, which, in turn, should decrease with increasing guidance, so that a change in decision bias might somewhat *attenuate* set-size effects. For illustrations of those influences (enumerated above) that were most crucial to the debate— namely, dwell time, guidance, quitting, and window size— we created an interactive spreadsheet available at https:// github.com/Liesefeld/searchSlopes.

If subsamples of items are processed in parallel, window size would likely be strategically adapted in order to maintain an acceptable signal-to-noise ratio within each window (Humphreys & Müller, 1993; Theeuwes, 1994; Treisman & Souther, 1985). In particular, less discriminable targets would require smaller windows because each nontarget contributes an independent amount of decision noise, so that the signal-to-noise ratio decreases with the number of nontargets within a window. Items might be processed individually, but in parallel, so that the probability of at least one nontarget identification process accidentally hitting the target boundary increases with set size (Moran et al., 2016). Alternatively, information is pooled across all objects and summary statistics within each window are evaluated (Hulleman & Olivers, 2017; Treisman & Gormican, 1988; Treisman & Souther, 1985). These statistics are more influenced by the target (and thus indicative of target presence) the smaller the number of nontargets that are included in the summary statistic. The disadvantage of smaller window sizes is that the window has to shift more often to process the same total display area/number of objects. Furthermore, if a limited processing resource is shared among

all items, processing more items in parallel would decrease processing speed for each individual item, including the target. Thus, steeper search slopes can be explained by smaller windows and by slower processing within each window.

For purely parallel models, similar principles apply as for parallel processing within attentional windows, only that the window size would be fixed at display size. Limited-capacity parallel models assume an increase in decision time when the resource is shared among an increasing number of items. Unlimited-capacity parallel models, instead of adapting window size, adapt the decision criterion (i.e., wait for the accrual of more information). That is, search slopes in purely parallel models can be explained by limited capacity and by increases in decision criteria. Relatedly, the time required for the spatially parallel computation of priority maps at the first stage of partly serial models might, too, depend on the number of nontargets (Buetti, Cronin, Madison, Wang, & Lleras, 2016).

Matters are further complicated by the fact that target discriminability, which is considered crucial by virtually all contestants in the debate, is influenced not only by features of the target and the nontargets but also by their spatial arrangement: Targets presented further in the periphery are processed with lower resolution, so that one would have to assure that large set sizes do not come along with a higher incidence of peripheral targets than small set sizes do. Furthermore, target discriminability depends on local feature contrast (Julesz & Bergen, 1983; Nothdurft, 1993; Tsotsos et al., 1995), so that one would also have to ensure that local density is kept constant across set sizes. In an attempt to meet these criteria, Liesefeld et al. (2016) implemented densely packed, circular search arrays to keep local feature contrast constant and manipulated set size by adding rings. With such arrays centered on fixation, targets would, on average, be farther from fixation in larger displays. To compensate for this confound (of set size with target eccentricity), Liesefeld et al. positioned the center of their circular arrangements unpredictably across trials in a way that kept the average target eccentricity constant across set sizes. To achieve the same ends, Palmer (1995) presented identically arranged search arrays and manipulated set size by cueing smaller or larger display regions as relevant for the upcoming search task; *visual marking* (Watson & Humphreys, 1997), based on previewing a set of the nontarget items in advance of the whole search display (the latter including the target), might be used to manipulate set size in a similar way. However, these techniques complicate the task design, might induce dynamics of their own, and were not implemented in most studies that examined set-size effects.

Interpreting slopes becomes even more difficult when considering that participants are likely sometimes to guess that a target is present before they have found it and quit search with a target-absent response without having scanned all objects in the display. Guessing rates likely depend on various task characteristics, with stimulus properties that make displays more

difficult to search (in particular, greater set sizes and less discriminable targets) engendering higher guessing rates. A well-studied effect is that more targets are missed when targets are encountered rarely, indicating that target prevalence influences the disposition to answer/guess "target present" versus "absent" (Ishibashi, Kita, & Wolfe, 2012; Wolfe, Horowitz, & Kenner, 2005; Wolfe & Van Wert, 2010).

Early quitting poses a special problem for another criterion highlighted by Treisman: serial, self-terminating search should yield a 1:2 ratio of target-present:absent slopes because the target is found on average after inspecting half the items, but to be certain that no target is present requires scanning of all objects. This is indeed a very strong and unique prediction from the serial/parallel dichotomy envisaged by FIT. However, empirically, present:absent slope ratios take various values (Liesefeld et al., 2016; Pashler, 1987a; Wolfe, 1998). In a certain range of efficiency, target-present slopes become flat while target-absent slopes remain steep, thus producing slope ratios of 1:6.90 and higher (Liesefeld et al., 2016).

A final, more advanced approach is to take more information from the data into account, such as the variability or the whole distribution of response times across trials (Rangelov, Müller, & Zehetleitner, 2017; Townsend, 1972; Wolfe, Palmer, & Horowitz, 2010). However, at least regarding the distinction between serial and parallel search, even these are not particularly diagnostic: both partly serial and purely parallel models account almost equally well for various kinds of searches (with a slight advantage for our serial model so far; Moran, Liesefeld, Usher, & Müller, 2017; Moran et al., 2016; Moran, Zehetleitner, Müller, & Usher, 2013; Narbutas, Lin, Kristan, & Heinke, 2017).

Treisman's response

From the above ingredients (and potentially more), one can chose ad libitum to explain variations in search slopes (we encourage you to try this out for yourself using the interactive spreadsheet at https://github.com/Liesefeld/searchSlopes). Somewhat surprisingly, rather than explaining the challenging findings of steep search slopes in feature search within a parallel framework (e.g., by assuming limited capacity), Treisman opted to drop the assumption that all feature searches are parallel. Doing so required a revision of FIT. Treisman and Gormican (1988) argued that attention is needed not only to process conjunctions but also to discriminate a target from nontargets that share the same feature differing only in degree. Attention is needed because it enhances resolution (e.g., Yeshurun & Carrasco, 1998). Thus, feature search becomes serial when targets and nontargets do not differ qualitatively, but only quantitatively. From the bouquet of options outlined above, Treisman considered three and finally

decided on the latter two: variation in dwell time, window size, and guidance.

Treisman and Souther (1985) found explaining slopes via variations in dwell time problematic because they observed target-absent slopes as shallow as 13 ms/item (see also Liesefeld et al., 2016, who found statistically significant target-absent slopes down to 2.9 ms/item), whereas traditional estimates of dwell time/shifts of attention are around 150–300 ms (Eriksen & Hoffman, 1972; Duncan, Ward, & Shapiro, 1994; Raymond, Shapiro, & Arnell, 1992; Theeuwes, Godijn, & Pratt, 2004; Ward, Duncan, & Shapiro, 1996, 1997). More recent evidence is indicative of shorter dwell times, of the order of 50–60 ms (Grubert & Eimer, 2016; Jenkins, Grubert, & Eimer, 2018), which might still be too long to explain some (barely) inefficient searches.

Given this, Treisman preferred an explanation in terms of window size and summary statistics: "subjects might check groups of items in parallel, with group size depending on the discriminability of the pooled feature response to groups containing only distractors and to groups in which the target replaced one of the distractors" (Treisman & Gormican, 1988, p. 18; see also Treisman & Souther, 1985). As elaborated above, discriminability of clumps containing versus not containing a target based on summary statistics depends on both the discriminability of target versus nontarget features and the number of items included in each clump (window size). Thus, with decreasing target-nontarget discriminability, observers go for smaller windows in order to keep target-present and target-absent clumps discriminable based on summary statistics.

Additionally, Treisman and Sato (1990) explained flat slopes in conjunction search by assuming that a master map can guide serial search (effectively adopting a version of GS): They argued that when target and nontarget features are highly discriminable in conjunction search, people might inhibit one or multiple nontarget features so that the target is the only object for which activity on the master map survives. The serial scan then always starts at the target so that no other items need to be inspected and, consequently, search slopes are flat. More items need to be scanned in conjunction search only if the target does not stand out sufficiently—that is, if random noise makes it difficult to discriminate the target priority signal from the priority signals induced by nontargets.

Although these additional assumptions could rescue the assumption of a dichotomy that is so central to FIT, they also weakened its theoretical appeal: with each additional assumption, the theory became less parsimonious. This is particularly problematic as theories are available that are already more parsimonious than the original version of FIT (see above). Furthermore, the various post hoc assumptions added over time to account for new findings and the intermingling of different theoretical accounts for the same phenomena make this theory somewhat hard to digest for the "'theoretician." Some complexities are probably owing to the adherence to

the problematic yardstick of steep versus flat search slopes. For this reason, we will next propose a different yardstick—along with a new theoretical framework that adopts many assumptions from FIT and other theories, including the idea of a search dichotomy, but rearranges these in a way that appears more orderly to us.

A new search dichotomy

In the theoretical framework proposed here, the core dichotomy concerns whether focal attention is guided by the priority map or, alternatively, items are scanned in a spatially systematic way (e.g., left-to-right or clockwise). To reiterate, the priority map is a (somewhat noisy) spatial representation of the visual scene that, rather than conveying information about stimulus features, simply provides a summary code of the relative importance of each item in the scene. Although it is derived from visual input, it is not involved in transmitting visual feature information (see Wolfe, 2007; cf. Li, 1999, 2002). We shall argue that the priority map is useful during search only if it reliably guides attention to the target and if this strategy is more efficient than its alternative—namely, inspecting clumps of adjacent objects in parallel in a spatially systematic fashion. Both strategies (sketched in Fig. 1) share certain computational principles, and both can be used to solve most search tasks; but they also differ in certain features, making one or the other more efficient in a given search scenario. We will refer to these strategies by their most distinguishing characteristic: *priority guidance* versus *clump scanning*. In an attempt to flesh out a working theory, we will outline below some mechanisms underlying these search modes that appear plausible based on our knowledge of the literature. However, the general idea does not depend on all of these details, and some will likely be subject to modification in future work.
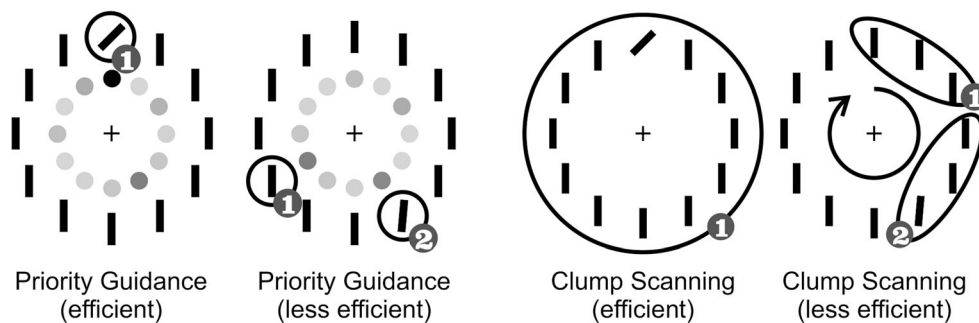
Clump scanning

For clump scanning, we adopt the assumption that subsets of items within an attentional window (or focus) can be processed in parallel (Humphreys & Müller, 1993; Pashler, 1987a; Theeuwes, 1994; Treisman & Souther, 1985), but only spatially contiguous items can be in the window at any time (Posner, Snyder, & Davidson, 1980; for more recent evidence against split attentional foci, see VanRullen, Carlson, & Cavanagh, 2007). In other words, if two items are in the attentional window, all items in between are in the window, too. The restriction to spatially adjacent items (a geographical clump) is crucial, because it typically renders selection of multiple potentially relevant objects based on the output of the priority map inefficient. This is because the priority map would typically highlight nonadjacent items, either because salient objects are positioned at random (not necessarily adjacent) locations or because priority noise in homogeneous as well as heterogeneous displays is distributed randomly without any strong (or even a negative) bias for adjacent positions.

It would, of course, be conceivable that several objects positioned around a high-priority location are processed in parallel within a window. However, this would not considerably increase, or it might even decrease, search efficiency compared with processing only the single object at the high-priority location, because the surrounding low-priority objects are unlikely to be targets anyway and processing them in addition might cost resources. Thus, the parallel-processing advantage of clump scanning and the set-restriction advantage of priority guidance are assumed to be incompatible in typical search scenarios.

In clump scanning, the size of the attentional window would be adapted according to the difficulty of search (in order to maintain a reasonable signal-to-noise ratio), and the window is shifted systematically (e.g., in clockwise direction or along plausible target locations in naturalistic scenes). We conceive of clump scanning as systematic, rather than guided,

Fig. 1 The two proposed search modes applied to the same search displays (the black bars). In this hypothetical task, participants search for a strongly (efficient search) or slightly (less efficient search) tilted target bar among homogenous vertical nontarget bars. With priority guidance (left), items are examined individually, and choice of items is guided by a priority map, as sketched by the grayish dots. Because this map is noisy, attention might sometimes visit a nontarget first (priority guidance, less efficient). With clump scanning (right), several spatially contiguous items within an attentional window are examined in parallel. When the window does not encompass all items (clump scanning, less efficient), it is moved across the display in some spatially systematic fashion (e.g., clockwise, as indicated by the arrow)

in order to avoid postulating two different priority maps (one for priority guidance and one for clump scanning; this will become clearer below) and to tentatively keep the two modes maximally dissimilar. Also note that there is some evidence that observers turn to systematic strategies under some conditions (e.g., von Mühlenen, Müller, & Müller, 2003).

Processing within a window could be described as item-wise two-boundary parallel (probably capacity-limited) template matching (Bundesen, 1990; Bundesen, Habekost, & Kyllingsbaek, 2005; Duncan & Humphreys, 1989, 1992; Humphreys & Müller, 1993; Olivers, Peters, Houtkamp, & Roelfsema, 2011; Soto, Hodsoll, Rotshtein, & Humphreys, 2008; Wolfe, 1994, 2007). Thus, for each item, a diffusor accumulates evidence in favor of a "target" or a "nontarget" decision and search terminates either when the "target"-decision threshold is reached or when all or a sufficiently high proportion of diffusors have reached the "nontarget"-decision threshold (for a detailed description and computational implementation, see Moran et al., 2016). Alternatively, observers might evaluate summary statistics of all items within the attentional window for deciding between a target-present and target-absent display (Rosenholtz, Huang, Raj, et al., 2012; Treisman & Souther, 1985). Search goals (or templates) and previous experience might speed the decision process by optimizing the coding of target features or increasing the target signal (e.g., Chelazzi, Duncan, Miller, & Desimone, 1998; Li, 1999, 2002; Zelinsky & Bisley, 2015). The respective mechanisms likely affect neuronal activation already before search-display onset.

Priority guidance

Proposed workings of priority guidance are adopted from the dimension-weighting account (Found & Müller, 1996; Liesefeld, Liesefeld, & Müller, 2019; Liesefeld, Liesefeld, Pollmann, & Müller, 2019; Liesefeld & Müller, 2019; Müller, Heller, & Ziegler, 1995), which, in essence, is a version/specification of guided search. As sketched in Fig. 2, various maps code for the saliency of each object within each of multiple dimensions. Saliency is determined by local feature contrast computed per feature dimension—that is, an object is the more salient within a particular dimension (i.e., achieves a higher value on that dimension's saliency map) the more it differs featurally (in that dimension) from the objects in its surround (see also the concept of *conspicuity maps* in Walther & Koch, 2006). Of note, dimension-specific saliency maps do not convey information about specific object features, but only about how much each object differs from its surround.

These saliency signals are then integrated into a priority map in a weighted fashion. While saliency is the bottom-up contributor to the priority map, top-down influences (goals, expectations, etc.) are expressed in terms of dimensional weights that modulate the gain from the individual dimensional saliency maps to the priority map (*dimension weighting*; see Liesefeld, Liesefeld, Pollmann, & Müller, 2019). The weighting is likely never "absolute" for one dimension, in order to ensure that unexpected but highly relevant signals generated in other dimensions can pass through and summon attention (see Liesefeld, Liesefeld, & Müller, 2019), and/or because stronger weighting might require more cognitive resources than the observer may be able or willing to expend (e.g., due to other cognitive demands of the task; see, e.g., Irons & Leber, 2018; Lavie, Hirst, de Fockert, & Viding, 2004). Once the weights are set, priority guidance is performed by a relatively autonomous (i.e., reflexive) system, which is why it subjectively feels rather passive and effortless (see Wyble et al., 2018).

By default, the priority map guides attention to the most probable target location(s); alternatively its summary statistics are evaluated (e.g., with homogeneity indicating target absence and heterogeneity target presence in certain search scenarios) to make a "target-present" versus "target-absent" decision without the need to allocate focal attention (see Luck & Ford, 1998; Rangelov et al., 2017; Rosenholtz, Huang, & Ehinger, 2012). When there are multiple priority peaks, attention is allocated sequentially to each peak so as to determine whether the item at the respective position is a target or a nontarget/distractor object (with attention potentially moving from one peak to the next, in descending order, by suppression of activation at each inspected and rejected location; Klein, 1988; Klein, Schmidt, & Müller, 1998; Wang & Klein, 2010). But even if there is just one priority peak (the target), attention might typically be allocated there to further examine the item that gave rise to the peak (e.g., to extract whatever information is required for a classification response) or double-check that this item is indeed the target rather than a peak of noise (Hoffman, 1978, 1979). This close inspection might well involve a two-boundary template matching process as in clump scanning, though with a window size restricted to one item.

Resolving current controversies

We propose that acknowledging these two complementary (albeit, as pointed out above, "redundant") modes of search, instead of categorically rejecting the idea of a search dichotomy, helps resolve various controversies in the literature on visual search and visual attention in general. Given our background, we are most familiar with controversies regarding the dimension-weighting account (DWA) and so will concentrate on these here. Notwithstanding this somewhat limited scope of the following discussion, we expect that the proposed dichotomy will also help to resolve other controversies in future theoretical work.
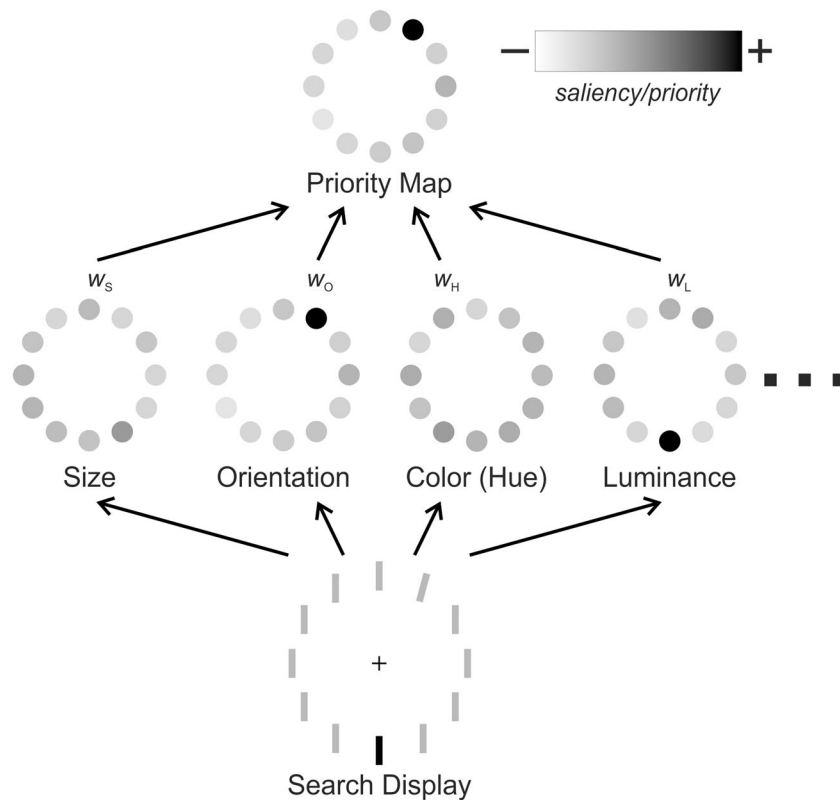
Fig. 2 Simplified sketch of priority computations according to the dimension-weighting account. From the search display, saliency maps are computed for each feature dimension, noisily reflecting how strongly each object differs from its surround. These activations are weighted and integrated at the superordinate priority map, which in turn guides the allocation of focal attention. Thus, the influences of the various saliency maps on priority signaling depend on the respective weight settings (w_S, w_O, w_M, and w_L); in this example, w_o is set to 1, while all other weights are set to 0, reflecting (somewhat unrealistically) perfect implementation of the task goal to find the tilted target bar. What exactly constitutes a dimension in terms of priority computations and whether some of the weights may influence each other remains to be examined (see Liesefeld, Liesefeld, Pollmann, & Müller, 2019)

One recurring challenge leveled at the DWA is that certain experimental effects are not dimension-specific but feature-specific in nature, apparently indicating that top-down influences on priority computations are not dimensionally constrained, but applied already at an earlier, feature-specific stimulus-coding stage (Li, 1999, 2002). We will show that in the reported instances, search displays are constructed in a way that renders clump scanning more efficient than priority guidance. As dimension weighting is assumed to play a role only in the computation of priorities, no dimensional constraints are to be expected during clump scanning, because clump scanning (typically) proceeds in a spatially systematic manner uninfluenced by saliency/priority computations. Instead, the notion of template matching already implies feature-specific processing, as templates are by definition representations of specific features (though likely with an element of imprecision; Geng, DiQuattro, & Helm, 2017).

We assume that observers choose either priority guidance or clump scanning to perform a given search task: They might initially try out both strategies, but—constantly monitoring their performance—would quickly settle on the strategy that feels most efficient and least effortful in a given situation. A

theoretical alternative is that both strategies are applied in parallel to each search display, and the strategy that finishes first determines response times. The latter alternative is, arguably, less likely: Running both strategies in parallel would probably waste cognitive resources, and the available evidence suggests that only one distinct search mode persists following a training phase (Leber & Egeth, 2006; Zehetleitner, Goschy, & Müller, 2012) or may be induced at will (Smilek, Enns, Eastwood, & Merikle, 2006; Watson, Brennan, Kingstone, & Enns, 2010; see section on other search dichotomies below). A third theoretical possibility is that, under some conditions, priority guidance is applied first on each trial, and only if it fails to detect a target will the display be processed more thoroughly via spatially systematic clump scanning. This would be in line with steep target-absent slopes together with flat target-present slopes for an intermediate range of target discriminability, where observers potentially switch to inefficient clump scanning once priority guidance has failed to pinpoint the target early on (though this pattern can also be explained by a model that knows only priority guidance; Liesefeld et al., 2016). Thus, before examining the controversies alluded to above in greater detail, it is

useful to consider under which conditions one or the other strategy is more likely to be chosen.

Priority guidance is inefficient when the target produces no or only slightly stronger activity on the priority map relative to other objects in the scene (e.g., because it differs only little from the nontargets and is therefore of low saliency). In this case, the map would often guide attention to "noise" peaks, so that nontargets would be inspected instead of the target. If noise is random and not fixed across attention allocations (e.g., because the priority map is computed anew after each attention allocation; see Wolfe, 2007), the target will still, eventually, be the object achieving the highest priority and thus be selected. If noise is stable or some object other than the target consistently gains a higher priority (Liesefeld, Liesefeld, Töllner, & Müller, 2017), the observer has to keep track of the locations that were already attended focally in order to avoid repeatedly inspecting the same high-priority "distractors." This keeping-track might be limited by working memory resources or by how long space-based inhibition on the priority map persists (Klein, 1988; Klein et al., 1998; Wang & Klein, 2010). Priority guidance would become particularly inefficient when, due to poor memory, the same objects are inspected repeatedly. In the worst case, the same set of high-priority distractors would alternate in summoning attention, so that the target is never found. In any case, in tasks featuring low-saliency targets, priority guidance would always come with uncertainty regarding the absence of the target (in target-absent displays). These would be strong reasons to turn to clump-scanning mode instead.

While the efficiency of clump scanning would also decrease with decreasing target discriminability, a systematic scan would (virtually) guarantee that the target, if present, is eventually found (see the older work on continuous search; e.g., Prinz, 1986) or that all objects are correctly discerned as nontargets on target-absent trials. Furthermore, although previous models have treated the creation of the priority map as essentially cost free or incurring a fixed cost, recent evidence indicates that the time taken to compute a priority map increases with the number of nontargets in the display (Buetti et al., 2016). Such a cost might further reduce the incentive to use priority guidance when it is inefficient.

Strategy choice therefore (loosely) depends on *relative target priority*, which, in turn, is influenced by target saliency, by nontarget saliency, and by how well observers can "tune in" to the specific search targets via setting dimensional weights. As discussed above, (bottom-up) target saliency is influenced by target/nontarget feature contrast, display density, and the target's distance to fixation. The effectiveness of top-down control via dimension weighting depends on factors such as availability of cognitive resources and predictability of and experience with various properties of the search display (e.g., Allenmark, Wang, Liesefeld, Shi, & Müller, 2019). Finally, priority guidance is most beneficial if the display contains many items. If there are only few items that can easily be scanned in one or a few clumps, there is little incentive to hazard the uncertainties involved in priority guidance.

Further (seemingly auxiliary) experimental details might also influence the choice of search mode (ceteris paribus) without or against the intention of the researcher. The type of search task (detection vs. localization/classification; see, e.g., Liesefeld, Liesefeld, Pollmann, & Müller, 2019; Töllner, Rangelov, & Müller, 2012), for example, might affect this choice because knowledge of the location of the target comes at no cost in priority guidance, but might require additional mechanisms for clump scanning (as discussed for parallel target-detection mechanisms in general in the section Theoretical Challenge, above).

As another example, Buetti et al. (2016) presented two types of nontargets: *lures* and *candidates* (see also Pashler, 1987b; Tsotsos, 1990; von Mühlenen & Müller, 2000). Lures are very inconspicuous nontargets that are very unlikely to be selected for further inspection. Candidates are more conspicuous nontargets that might be inspected before the target is found. If lures and candidates are randomly intermixed (spatially) within a display, scanning of a geographical clump would usually include candidates and lures; it might then be more efficient to guide focused attention specifically to the priority peaks reflecting candidates and targets, thereby effectively eliminating the processing of lures. In the same displays without lures (just candidates and the target), clump scanning of multiple candidates (and potentially the target) might be more efficient.

Nontargets can become candidates when their local feature contrast is high (i.e., if they are surrounded by dissimilar objects). This underscores the fact that efficiency of guidance is not determined by absolute target priority, but—as mentioned above—by *relative* target priority—that is, priority guidance is efficient only if the target priority is much higher than the priority of all (or most) other objects in the display. Lleras, Wang, Madison, and Buetti (2019) compared displays which contained only nontargets of medium similarity to displays containing also nontargets of low similarity to the target. If search efficiency were determined by pure target priority, the latter displays should produce more efficient search, because local target contrast is higher on average. Instead, Lleras et al. (2019) found that displays containing only medium-similarity nontargets produced more efficient search. This can be understood by considering that through the addition of low-similarity nontargets, the target is no longer the only item producing high local feature contrast. Now, low-similarity and high-similarity items produce relatively high local feature contrast as well, thus competing with the target for allocation of attention.

The notion of *relative target priority* is also useful to understand effects of distractor heterogeneity (Duncan & Humphreys, 1989), conjunction search (Treisman & Gelade,

1980), and search asymmetries (Treisman & Souther, 1985). These phenomena can be explained in terms of variation in the efficiency of priority guidance, largely in line with guided search (Wolfe, 1994, 2007).

With heterogeneous distractors and conjunction-search displays, the local feature contrast of the target is usually similar to that of nontargets, because many nontargets will generate (relatively) high local feature contrasts, too, and the local feature contrast of a conjunction target is necessarily lower compared with the respective feature-search displays (see also Wei, Yu, Müller, Pollmann, & Zhou, 2018; Zhaoping & May, 2007).[1] Consequently, with heterogeneous distractors and conjunction searches, the target is unlikely to produce a singular peak on the priority map, so that (similar to scenarios with low-saliency targets discussed above) nontargets often draw attention, and potential revisits become an issue.

As regards search asymmetries, these can be explained by variation in guidance due to low versus high local feature contrast of the target: For instance, a closed-circle target among broken-circle nontargets does not generate strong local feature contrast, because the shape feature "circularity" (or whatever shape feature the local contrast coded on the respective saliency map is based on) is present in all nontargets as well. A broken-circle target among closed-circle nontargets, in contrast, produces a strong local feature contrast, because "brokenness" is absent from all surrounding nontargets. Thus, various, seemingly disparate phenomena can be explained within our framework based on considerations of relative target priority.

This new proposal is admittedly not the most parsimonious visual-search theory ever proposed. In fact, both strategies could in principle explain the whole range of empirically observed search slopes on their own—for instance, by variation of guidance or window size, respectively. That is, a theory postulating both strategies in order to explain search slopes is nonparsimonious. Importantly, though, as detailed above, search slopes are nondiagnostic for differentiating models of visual search, and focusing on explaining search efficiency with as few assumptions as possible has, arguably, misguided model development in visual search to quite some degree. Also given that human cognition is complex, the theory's increase in complexity (as compared with, say, purely parallel theories) might be warranted. After all, we know quite well that the human brain contains redundancy (e.g., having two hemispheres with largely overlapping functions), and such redundancy might have evolutionary advantages in case one

system fails (e.g., due to injuries). Most importantly, we believe that our theory can account for all phenomena that previous theories were able to accommodate, some of which were discussed above and for each of which we will, of course, have to provide detailed proof in future work.

For the present purpose, we regard it to be of greater interest to consider conflicting results that, to date, have not been explained by any single existing theory. Thus, we now turn to a number of controversies in need of theoretical resolution—which, in our view, requires the existence of a search dichotomy similar to the one outlined above. The common denominator of our explanations of these phenomena—dimension versus feature specificity—might provide a fruitful new yardstick for distinguishing the two search modes, and we expect to develop additional yardsticks in future work. In what follows, we will focus on target-repetition effects—a form of history effects, and on distractor handling—a form of voluntary control. Together, history effects and voluntary control cover all (or, at least, the most intensively investigated) top-down influences on visual search (e.g., Gaspelin & Luck, 2018b), so that the two examples given here are representative of the whole set of phenomena to which the feature-specificity versus dimension-specificity distinction applies

Target-repetition effect

When the target changes unpredictably across trials, search can slow down under certain conditions (Egeth, 1977; Müller et al., 1995; Treisman, 1988). Closer analyses indicated that the loss in speed results mainly from trials on which the target changed with respect to the previous trial, while search is relatively fast when the target repeats (Müller et al., 1995). Crucially, Müller and colleagues found that the target-repetition effect is dimension specific: Response times were speeded independently of whether the previous target had exactly the same defining feature or just a feature within the same dimension (for a recent review, see Liesefeld, Liesefeld, Pollmann, & Müller, 2019; for a mathematical account, see Allenmark, Shi, & Müller, 2018). In stark contrast, another group of researchers observed a target repetition effect that was feature specific: Responses were speeded only if the previous target had the same feature, and they were slow when the previous target had a different feature from the same dimension. This feature-specific effect was termed *priming of pop-out* (PoP; Lamy, Zivony, & Yashar, 2011; Maljkovic & Nakayama, 1994). A crucial difference between the two conflicting lines of study is that only very few (and relatively widely spaced) nontargets are typically used in PoP studies. Indeed, the feature-specific PoP vanished when more (and densely spaced) nontargets were presented (Rangelov, Müller, & Zehetleitner, 2013; see also Krummenacher, Grubert, & Müller, 2010; Meeter & Olivers, 2006; Rangelov et al., 2011a, 2011b; Zehetleitner et al., 2012). Rangelov and

---

[1] In typical conjunction-search displays, targets and nontargets produce roughly the same average local feature contrast in each dimension. For example, a red-horizontal-conjunction target will typically be surrounded by red and horizontal nontargets. Each red nontarget has the same chance as the red target of being surrounded mainly by blue nontargets and therefore achieve high saliency, or of being surrounded mainly by red nontargets and therefore achieve low saliency. The same holds for the orientation dimension.

colleagues argued, and later showed more directly (see Rangelov et al., 2017), that with a (for PoP studies typical) set size of three objects (one target and two nontargets), targets often are low in saliency due to the lack of local feature contrast—a consequence of the sparse layout (see Julesz & Bergen, 1983; Liesefeld et al., 2016; Nothdurft, 1993; Tsotsos et al., 1995)—and therefore actually fail to pop out (see also Becker, 2008).

On our account, low saliency makes the use of the priority map less efficient. Additionally, if only few objects are shown, these can be scanned at once, rendering clump scanning the more efficient strategy in typical PoP studies. On this interpretation, priming of pop-out is a speed-up of the template matching process rather than an increase in (attention-guiding) priority (see also Huang, Holcombe, & Pashler, 2004). That is, PoP search displays induce clump scanning, in which top-down influences are feature specific, while the (dense) displays used by Müller and colleagues induce priority guidance, in which top-down influences are dimensionally constrained (for an interesting exception, see Wolfe, Butcher, Lee, & Hyle, 2003).

Distractor suppression

The presence of a salient-but-irrelevant distractor, such as a red singleton during search for a form singleton, delays response times. This was interpreted in terms of involuntary attention allocation to the distractor preceding and therefore delaying the allocation of attention to the target (*attentional capture*; Theeuwes, 1991, 1992). However, recent electrophysiological evidence indicates that such a distractor does not typically capture attention, at least not if the target remains the same across trials (Burra & Kerzel, 2013; Gaspar & McDonald, 2014; Jannati, Gaspar, & McDonald, 2013). This would be predicted by the dimension-weighting account, because distractors in most attentional-capture studies are singletons in a dimension other than the target, and it would be reasonable to simply down-weight any signal from the respective distractor dimension. This down-weighting would not be possible, or it would be harmful, if the distractor stands out in the same dimension as the target (Liesefeld & Müller, 2019). Indeed, same-dimension distractors cause massive interference (an order of magnitude larger than different dimension distractors; Liesefeld, Liesefeld, & Müller, 2019; Sauter, Liesefeld, & Müller, 2019; Sauter, Liesefeld, Zehetleitner, & Müller, 2018) and clear electrophysiological signs of attentional capture (Liesefeld et al., 2017). We therefore concluded that distractors reliably capture attention when they are defined in the same dimension as the target—that is, distractor handling is dimensionally constrained (for a review, see Liesefeld & Müller, 2019, where we also explain in detail the special status of color, how dimensional relationship differs from mere similarity, and how imperfect preparatory weighting—for example, due to attentional resources being consumed by other aspects of the task—sometimes allow for attentional capture by different-dimension distractors).

Interestingly, another group of researchers came to the diametrically opposed conclusion—namely, that distractor handling is feature specific (*first order feature suppression*; Gaspelin & Luck, 2018a). They argued that if only the distractor's dimension is predictable, dimension weighting (*second order feature suppression*) but not feature weighting can be employed. They found, however, that knowing the distractor dimension was not sufficient to shield against distraction, but instead information on the specific distractor feature was needed—thus favoring the assumption that distractor down-weighting is necessarily feature specific. Importantly, Gaspelin and Luck (2018a) likely induced inefficient search (see also Gaspelin, Leonard, & Luck, 2015) among only a few objects. In particular, observers had to search for a specific target shape among various other nontarget shapes (i.e., high nontarget heterogeneity according to Duncan & Humphreys, 1989). We, in contrast, focused on tasks in which the target clearly stood out from the homogenous nontarget background (e.g., Liesefeld et al., 2017; Liesefeld et al., 2016; Liesefeld, Liesefeld, & Müller, 2019; Sauter et al., 2019; Sauter et al., 2018). Thus, in Gaspelin and Luck's task, using the priority map is likely inefficient, making observers resort to clump-scanning mode. As a consequence, dimensional-weight settings controlling the transfer from the dimension-specific saliency maps to the overall priority map would not play a role. Instead, in line with the present proposal, Gaspelin and Luck's observers might have performed clump scanning, using an efficient feature-specific distractor template that allows for fast identification and rejection of the distractor (e.g., Beck, Luck, & Hollingworth, 2018; Reeder, Olivers, & Pollmann, 2017; Woodman & Luck, 2007) or tuning the target template in a way that optimized target processing in the presence of the salient distractor (Geng et al., 2017; Navalpakkam & Itti, 2007).

Relations to other theories

The theoretical framework of visual search proposed here does not only draw from feature-integration theory and the dimension-weighting account as discussed above, but unifies various existing accounts (e.g., Duncan & Humphreys, 1989, 1992; Humphreys & Müller, 1993; Müller, Humphreys, & Donnelly, 1994; Pashler, 1987a; Theeuwes, 1994; Wolfe, 1994, 2007; Wolfe et al., 1989). Although a comparison with each existing account in adequate depth is beyond the scope of the present paper, we will briefly address and integrate a few selected thoughts that we find particularly intriguing in the context of current debates. Others would likely prioritize other

aspects, and the following considerations are by no means meant to be exhaustive.

Feature-integration theory

As in FIT, we assume that observers can choose between two search modes. In contrast to FIT, rather than conceiving of these modes as purely parallel or purely serial, we consider both to include serial and parallel aspects. Furthermore, instead of assuming that search becomes serial because of the need to integrate free-floating features, we assume that both modes can in principle be employed in all search tasks, and we suggest a variety of factors that determine which mode is most efficient, most notably display density, set size, and the composition of target and nontarget features. The actual choice, though, depends only partly on these task characteristics; it is also influenced by previous experience and, likely, some element of voluntary control. More research is needed to determine the relative influence of these factors on the "choice" of search mode.

Further search dichotomies

Wolfe (1998) used the unimodality of the distribution of search slopes across various tasks to argue that all search slopes stem from a single distribution, reflecting the same basic processes and differing only in efficiency. Reanalyzing this extensive data set, Haslam, Porter, and Rothschild (2001), by contrast, found indication that the slopes rather stem from two distinct but overlapping distributions, potentially reflecting two qualitatively different types of search. Interestingly, they speculated that "two distinct search processes might exist, both operating on continua of some sort" (p. 746)—which would be in line with our assumption of two search modes which can both vary in efficiency.[2] Haslam et al., however, refrained from speculating what these two distinct processes might be.

Bacon and Egeth (1994; see also, e.g., Leber & Egeth, 2006) proposed that people might choose to search for a specific feature or simply for any singleton (*feature-search* vs. *singleton-detection mode*). They observed that with homogenous nontargets, a salient distractor interfered with search. With heterogeneous nontargets, by contrast, the same distractor did not cause any interference. They reasoned that homogenous nontargets incentivize a singleton-detection mode, whereas heterogeneous nontargets force observers into a feature-search mode.

---

[2] It is of interest to note that we conceived of this dichotomy independently from and for different reasons than Haslam et al. (2001). Their work was brought to our attention only during the second round of reviews for the present manuscript.

Singleton detection might be an instance of priority guidance with all weights set equally. However, equal weights would not produce the pattern observed by Liesefeld, Liesefeld, and Müller (2019; see also Liesefeld et al., 2017; Sauter et al., 2019; Sauter et al., 2018): A physically identical distractor much more salient than the target produces either weak or strong interference, depending on whether the observer searches for a target defined in a different or the same dimension. Furthermore, with distractors of comparable saliency unpredictably intermixed across trials, only the same-dimension distractor produced strong interference. Both findings are indicative of unequal weighting of the various dimensions in the presence of distractor interference, arguing against both singleton-detection and feature-search modes. Consistent with this, Zehetleitner, Goschy, and Müller (2012) found that a different-dimension distractor caused significant interference even when observers were operating in feature-search mode, provided the distractor was absent in the preceding mode-induction phase—pointing to a crucial role of distractor practice for minimizing interference (see also Müller, Krummenacher, Geyer, & Zehetleitner, 2009). Thus, the interference caused by different-dimension distractors—which has been taken, by others, as indication of attentional capture (Theeuwes, 1991, 1992) or singleton-detection mode (Bacon & Egeth, 1994)—is likely only the residual remaining after incomplete down-weighting of the distractor dimension.

Two characteristics of the typical task used to induce feature-search mode would make it suited to induce clump scanning: First, the use of distractor heterogeneity, which is known to produce inefficient search (Duncan & Humphreys, 1989). Second, studies inducing feature-search mode typically use relatively small set sizes (five vs. nine items). Indeed, although search slopes are typically not significant in tasks used to induce feature search, the slopes in Leber and Egeth (2006), for example, were 3.5 ms/item and 3.25 ms/item for distractor-present and distractor-absent displays, respectively. Slopes of that size can be shown to significantly differ from zero if the tests have sufficient power (Liesefeld et al., 2016). Thus, even if the task was not very inefficient, it likely did not produce pop out. Instead it might have induced relatively efficient clump scanning, which our theoretical framework allows for. During clump scanning, a different-dimension distractor does not produce any costs because it bears no similarity to the target template (e.g., Duncan & Humphreys, 1989). In sum, feature-search mode would roughly correspond to our clump scanning and singleton-detection mode to our priority guidance (though the two theoretical dichotomies differ in specific details).

Another dichotomy was suggested by Watson et al. (2010; see also Smilek et al., 2006): Participants either focus on passive processing with stable fixation (*seeing*) or active exploration including eye movements (*looking*). These different strategies can be induced in the same search tasks by simply

instructing participants accordingly. Passive seeing likely corresponds to priority guidance, whereas active looking would correspond to clump scanning. Accordingly, the search modes proposed here can be regarded as mechanistic specifications of Watson et al.'s (2010) seeing versus looking modes.

Guided search

Priority guidance is basically a version of the broader class of models termed "guided search" (Wolfe, 1994, 2007; Wolfe et al., 1989), with a dimension-weighting flavor (Found & Müller, 1996; Liesefeld, Liesefeld, & Müller, 2019; Liesefeld, Liesefeld, Pollmann, & Müller, 2019; Liesefeld & Müller, 2019; Müller et al., 1995). One might argue against the proposed dichotomy by maintaining that what we describe as clump-scanning mode is just a search with zero guidance, where each object (or clump of objects) has to be processed until the target is found. There are two main reasons why, despite this possibility, we favor the notion of a search dichotomy: (a) Without the dichotomy, we cannot explain the conflicting findings regarding dimension versus feature specificity considered above; and (b) there is strong evidence that physically identical displays are treated differently depending on the search mode induced in a preceding training phase (Leber & Egeth, 2006; Zehetleitner et al., 2012).

Also note that the top-down mechanism envisaged by guided search differs from the one proposed here: While dimension weighting operates on dimension-specific local-feature-contrast signals, in guided search specific feature channels are up-weighted (i.e., feature-weighting with some imprecision) prior to the transformation into saliency values. This is likely done in a way that optimizes the signal-to-noise ratio (i.e., relative target saliency; Geng et al., 2017; Navalpakkam & Itti, 2007). It is, of course, conceivable that both top-down influences exist concurrently. In any case, more research is needed to determine whether we can maintain the plain-vanilla version of priority guidance or whether we have to add such a presalience feature-specific mechanism. Researchers embarking on this mission should keep in mind that it remains unclear exactly what constitutes a dimension in terms of saliency computations (for instance, color likely consists of multiple dimensions; see Liesefeld, Liesefeld, Pollmann, & Müller, 2019), and that observers likely also have mechanisms of spatial suppression at their disposal (Ferrante et al., 2018; Klein, 1988; Goschy, Bakos, Müller, & Zehetleitner, 2014; Klein, Schmidt, & Müller, 1998; Sauter et al., 2019; Sauter et al., 2018; Wang & Klein, 2010; Wang & Theeuwes, 2018; Watson & Humphreys, 1997).

Attentional-window theories

From the perspective of attentional-window theories (Humphreys & Müller, 1993; Pashler, 1987a; Theeuwes, 1994), one might argue against the proposed dichotomy by maintaining that what we interpret as priority guidance really is a relatively large window, affording rather efficient search. A similar argument might be made by theories assuming a functional viewing field (Hulleman & Olivers, 2017). Again, the dependence of the search strategy on the search mode adopted during a preceding training phase and the findings regarding dimension versus feature specificity reviewed above render these theories unlikely to provide a general account of all kinds of visual search. Put another way, rather than viewing attentional-window/template-matching theories and guided-search theories as competitors, we considers them to simply characterize different strategies that are preferred in different task situations. This underscores the point that in the extant literature, the label *visual search* has been applied to what constitute, from a cognitive-processing perspective, actually quite dissimilar behaviors. Accordingly, future research ought to distinguish clearly between tasks prone to induce priority guidance or, respectively, clump scanning to avoid theoretical stalemates and foster scientific progress.

Summary of the proposed dichotomy

To conclude, the available evidence leads us to believe that there are two main strategies or processing modes employed in visual-search tasks: One that can be used when the target is singled out easily from the stimulus array by a combination of bottom-up and top-down factors (priority guidance), and one that is used when all items need to be inspected systematically (clump scanning). Using the example from the introduction, people will use priority guidance if they expect the target to stand out, such as the single (or the few) red-cased book(s) on our bookshelf full of black-cased books, and they will carefully scrutinize each element via clump scanning when searching for the proverbial needle lying in plain view on top of the haystack. In between these extremes, strategy choice might be a matter of individual preference and habituation.

Depending on the type of search a particular researcher is focusing on, it would seem self-evident that the respectively other camp is fundamentally mistaken: If one thinks of the book with the red binder on the bookshelf, it would appear clear that search is best described as massively parallel; if one thinks of the needle in the haystack, it appears clear that search is painfully serial. Thus, acknowledging that there simply are two fundamentally different search modes that are employed in different situations will help us arrive at a theoretical reconciliation between the opposing camps, and researchers can expressly decide and clearly communicate whether they study priority-guided or clump-scanning search, thereby avoiding much of the confusion.

We propose that the "dimension versus feature specificity" of effect patterns can be used as an empirical yardstick to distinguish between the two modes of search. At first sight, this criterion might appear to make dimension weighting (and our new theory) unfalsifiable: Whenever there is compelling evidence against dimension weighting, we can simply claim that (involuntarily) the task induced a clump-scanning mode. But this would grossly overstate the flexibility of the theory: We have detailed various factors that make one or the other strategy more efficient. Thus, evidence against dimension weighting in a task that renders priority guidance highly efficient (e.g., dense, homogenous arrays of nontargets) would seriously challenge our theory. Additionally, we aim to develop further criteria to differentiate between the two modes in future work.

Obviously, the assumption of two rather flexible search modes renders our idea considerably less parsimonious compared with its one-mode competitors. Parsimony is usually preferred in science, but parsimonious theories can, of course, be inaccurate models for reality (e.g., the debunked assumption that planetary orbits are circular). Simple theories might often be able to explain data patterns in one study or a selected set of data patterns, but then fail to account for other data patterns or for more diverse sets of patterns when forced to keep the exact-same set of (central and auxiliary) assumptions. Future work on the present proposal must identify more examples such as the target-repetition and distractor-suppression effects reviewed above and directly show within the same appropriately designed studies that one-mode models fail to account for the complete data pattern.

Also consider that the brain has not evolved to perform the artificial tasks employed in typical visual-search studies. Rather, when confronted with such tasks, it likely has to exploit mechanisms that have evolved for different, real-life reasons. There is certainly a huge range of fundamentally different mechanisms to pick from, and it is not too surprising that (at least) two of these cognitive mechanisms are useful for solving some (or most) laboratory search tasks.

Given that the dichotomy advocated here differs greatly from Treisman's original proposal, the reader might wonder whether the title is appropriate. As detailed above, clump scanning is serial in that the attentional window needs to shift several times, but it is also parallel because multiple items are processed in parallel. Conversely, priority guidance is parallel in that all objects are initially processed in parallel to extract priority values, but it is also serial because attention is subsequently allocated to each hotspot on the priority map in a serial manner. This might be taken to indicate that the parallel-versus-serial contrast is not well suited to describe the search dichotomy. Alternatively, one can think of the dichotomy from the standpoint of dialectical monism: Although priority guidance has a serial aspect, it is best described (and intuitively felt) as relatively passive and parallel. Although clump scanning contains a parallel aspect, it is best described (and intuitively felt) as relatively active and serial. Theorists from a Western background might feel uncomfortable with this way of thinking, though the notion of such complementary, interconnected, and interdependent forces ("yin and yang"') is well-accepted among (ancient) East-Asian philosophers.

Open practices statement This is a theoretical article, so the questions regarding open practice do not apply.

References

Allenmark, F., Shi, Z., & Müller, H. J. (2018). Inter-trial effects in visual search: Factorial comparison of Bayesian updating models. *PLOS Computational Biology, 14.* e1006328. https://doi.org/10.1371/journal.pcbi.1006328

Allenmark, F., Wang, B., Liesefeld, H. R., Shi, Z., & Müller, H. J. (2019). Probability cueing of singletondistractor regions in visual search: the locus of spatial distractor suppression is determined by color swapping. Visual Cognition. Manuscript in press.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, 55, 485–496. https://doi.org/10.3758/BF03205306

Beck, V. M., Luck, S. J., & Hollingworth, A. (2018). Whatever you do, don't look at the . . . : Evaluating guidance by an exclusionary attentional template. *Journal of Experimental Psychology: Human Perception and Performance, 44,* 645–662. https://doi.org/10.1037/xhp0000485

Becker, S. I. (2008). Can intertrial effects of features and dimensions be explained by a single theory?. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1417–1440. https://doi.org/10.1037/a001138

Buetti, S., Cronin, D. A., Madison, A. M., Wang, Z., & Lleras, A. (2016). Towards a better understanding of parallel visual processing in human vision: Evidence for exhaustive analysis of visual information. *Journal of Experimental Psychology: General*, 145, 672–707. https://doi.org/10.1037/xge0000163

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

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, 112, 291–328. https://doi.org/10.1037/0033-295X.112.2.291

Burra, N., & Kerzel, D. (2013). Attentional capture during visual search is attenuated by target predictability: Evidence from the N2pc, Pd, and topographic segmentation. *Psychophysiology*, 50, 422–430. https://doi.org/10.1111/psyp.12019

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, 80, 2918–2940.

Duncan, J., & Humphreys, G. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 578–588. https://doi.org/10.1037/0096-1523.18.2.578

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458. https://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*, 313–315. https://doi.org/10.1038/369313a0

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics*, *62*, 425–451. https://doi.org/10.3758/BF03212096

Egeth, H. (1977). Attention and preattention. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 277–320). New York, NY: Academic Press.

Engel, F. L. (1977). Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Research 17,* 95–108. https://doi.org/10.1016/0042-6989(77)90207-3.

Eriksen, C. W., & Hoffman, J. E. (1972). Some characteristics of selective attention in visual perception determined by vocal reaction time. *Perception & Psychophysics*, *11*, 169–171. https://doi.org/10.3758/BF03210367

Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology*, *79*, 1–16. https://doi.org/10.1037/h0026873

Ferrante, O., Patacca, A., Di Caro, V., Della Libera, C., Santandrea, E., & Chelazzi, L. (2018). Altering spatial priority maps via statistical learning of target selection and distractor filtering. *Cortex*, *10*, 267–95. https://doi.org/10.1016/j.cortex.2017.09.027

Found, A., & Müller, H. J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a 'dimension-weighting' account. *Perception & Psychophysics*, *58*, 88–101. https://doi.org/10.3758/BF03205479

Gardner, G. T. (1973). Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology*, *4*, 130–155. https://doi.org/10.1016/0010-0285(73)90009-1

Gaspar, J. M., & McDonald, J. J. (2014). Suppression of salient objects prevents distraction in visual search. *The Journal of Neuroscience*, *34*, 5658–5666. https://doi.org/10.1523/JNEUROSCI.4161-13.2014

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychological Science*, *26*, 1740–1750. https://doi.org/10.1177/0956797615597913

Gaspelin, N., & Luck, S. J. (2018a). Distinguishing among potential mechanisms of singleton suppression. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 626–644. https://doi.org/10.1037/xhp0000484

Gaspelin, N., & Luck, S. J. (2018b). "Top-down" does not mean "voluntary". *Journal of Cognition, 1*, 25. https://doi.org/10.5334/joc.28

Geisler, W. S., & Chou, K. L. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review, 102*, 356–378. https://doi.org/10.1037/0033-295X.102.2.356

Geng, J. J., DiQuattro, N. E., & Helm, J. (2017). Distractor probability changes the shape of the attentional template. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1993–2007. https://doi.org/10.1037/xhp0000430

Goschy, H., Bakos, S., Müller, H. J., & Zehetleitner, M. (2014). Probability cueing of distractor locations: Both intertrial facilitation and statistical learning mediate interference reduction. *Frontiers in Psychology, 5,* 1195. https://doi.org/10.3389/fpsyg.2014.01195

Grubert, A., & Eimer, M. (2016). The speed of serial attention shifts in visual search: Evidence from the N2pc component. *Journal of Cognitive Neuroscience*, *28*, 319–332. https://doi.org/10.1162/jocn_a_00898

Haslam, N., Porter, M., & Rothschild, L. (2001). Visual search: Efficiency continuum or distinct processes? *Psychonomic Bulletin & Review*, 8, 742–746. https://doi.org/10.3758/BF03196212

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

Hoffman, J. E. (1979). A two-stage model of visual search. *Perception & Psychophysics*, *25*, 319–327. https://doi.org/10.3758/BF03198811

Huang, L., Holcombe, A. O., & Pashler, H. (2004). Repetition priming in visual search: Episodic retrieval, not feature priming. *Memory & Cognition*, *32*, 12–20. https://doi.org/10.3758/BF03195816

Hulleman, J., & Olivers, C. N. L. (2017). The impending demise of the item in visual search. *The Behavioral and Brain Sciences*, *40*, e132. https://doi.org/10.1017/S0140525X15002794

Humphreys, G. W., & Müller, H. J. (1993). SEarch via Recursive Rejection (SERR): A connectionist model of visual search. *Cognitive Psychology, 25*, 43–110. https://doi.org/10.1006/cogp.1993.1002

Irons, J. L., & Leber, A. B. (2018). Characterizing individual variation in the strategic use of attentional control. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 1637–1654. https://doi.org/10.1037/xhp0000560

Ishibashi, K., Kita, S., & Wolfe, J. M. (2012). The effects of local prevalence and explicit expectations on search termination times. *Attention, Perception, & Psychophysics*, *74*, 115–123. https://doi.org/10.3758/s13414-011-0225-4

Itti, L. (2017). Until the demise of the functional field of view. *Behavioral and Brain Sciences, 40*, e140. https://doi.org/10.1017/S0140525X1600010

Jannati, A., Gaspar, J. M., & McDonald, J. J. (2013). Tracking target and distractor processing in fixed-feature visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1713–1730. https://doi.org/10.1037/a0032251

Jenkins, M., Grubert, A., & Eimer, M. (2018). The speed of voluntary and priority-driven shifts of visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 27–37. https://doi.org/10.1037/xhp0000438

Julesz, B., & Bergen, J. R. (1983). Textons, the fundamental elements in preattentive vision and perception of textures. *Bell System Technical Journal*, *62*, 1619–1645. https://doi.org/10.1002/j.1538-7305.1983.tb03502.x

Kinchla, R. A. (1974). Detecting target elements in multielement arrays: A confusability model. *Perception & Psychophysics*, *15*, 149–158. https://doi.org/10.3758/BF03205843

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

Klein, R. M., Schmidt, W. C., & Müller, H. J. (1998). Disinhibition of return: Unnecessary and unlikely. *Perception & Psychophysics*, *60*, 862–872. https://doi.org/10.3758/BF03206069

Krummenacher, J., Grubert, A., & Müller, H. J. (2010). Inter-trial and redundant-signals effects in visual search and discrimination tasks: Separable pre-attentive and post-selective effects. *Vision Research*, *50*, 1382–1395. https://doi.org/10.1016/j.visres.2010.04.006

Lamy, D., Zivony, A., & Yashar, A. (2011). The role of search difficulty in intertrial feature priming. *Vision Research*, *51*, 2099–2109. https://doi.org/10.1016/j.visres.2011.07.010

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*, 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review, 13*, 132–138. https://doi.org/10.3758/BF03193824

Li, Z. (1999). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 10530–10535.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6*, 9–16. https://doi.org/10.1016/S1364-6613(00)01817-9

Liesefeld, H. R., Liesefeld, A. M., & Müller, H. J. (2019). Distractor-interference reduction is dimensionally constrained. *Visual Cognition*. Advance online publication. https://doi.org/10.1080/13506285.2018.1561568

Liesefeld, H. R., Liesefeld, A. M., Pollmann, S., & Müller, H. J. (2019). Biasing allocations of attention via selective weighting of saliency signals: Behavioral and neuroimaging evidence for the dimension-weighting account. In T. Hodgson (Ed.), *Current topics in behavioral neurosciences: Processes of visuo-spatial attention and working memory.* Basel, Switzerland: Springer.

Liesefeld, H. R., Liesefeld, A. M., Töllner, T., & Müller, H. J. (2017). Attentional capture in visual search: Capture and post-capture dynamics revealed by EEG. *NeuroImage, 156*, 166–173. https://doi.org/10.1016/j.neuroimage.2017.05.016

Liesefeld, H. R., Moran, R., Usher, M., Müller, H. J., & Zehetleitner, M. (2016). Search efficiency as a function of target saliency: The transition from inefficient to efficient search and beyond. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 821–836. https://doi.org/10.1037/xhp0000156

Liesefeld, H. R., & Müller, H. J. (2019). Distractor handling via dimension weighting. *Current Opinion in Psychology, 29*, 160–167. https://doi.org/10.1016/j.copsyc.2019.03.003

Lleras, A., Wang, Z., Madison, A., & Buetti, S. (2019). Predicting search performance in heterogeneous scenes: Quantifying the impact of homogeneity effects in efficient search. *Collabra: Psychology, 5*, 2. https://doi.org/10.1525/collabra.151

Luck, S. J., & Ford, M. A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 825–830.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22*, 657–672. https://doi.org/10.3758/BF03209251

McLeod, P., Driver, J., & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature, 332*, 154–155. https://doi.org/10.1038/332154a0

Meeter, M., & Olivers, C. N. L. (2006). Intertrial priming stemming from ambiguity: A new account of priming in visual search. *Visual Cognition, 13*, 202–222. https://doi.org/10.1080/13506280500277488

Moore, C. M., & Wolfe, J. M. (2001). Getting beyond the serial/parallel debate in visual search: A hybrid approach. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints on human information processing* (pp. 178–198). Oxford, UK: Oxford University Press.

Moran, R., Liesefeld, H. R., Usher, M., & Müller, H. J. (2017). An appeal against the item's death sentence: Accounting for diagnostic data patterns with an item-based model of visual search. *Behavioral and Brain Sciences, 40*, e148. https://doi.org/10.1017/S0140525X16000182

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 & Review, 23*, 1300–1315. https://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), 24. https://doi.org/10.1167/13.8.24

Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics, 57*, 1–17. https://doi.org/10.3758/BF03211845

Müller, H. J., Humphreys, G. W., & Donnelly, N. (1994). SEarch via Recursive Rejection (SERR): Visual search for single and dual form-conjunction targets. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 235–258. https://doi.org/10.1037/0096-1523.20.2.235

Müller, H. J., Krummenacher, K., Geyer, T., & Zehetleitner, M. (2009). Attentional capture by salient color singleton distractors is modulated by top-down dimensional set. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1–16. https://doi.org/10.1037/0096-1523.35.1.1.

Müller, H. J., Liesefeld, H. R., Moran, R., & Usher, M. (2017). Parallel attentive processing and pre-attentive guidance. *Behavioral and Brain Sciences, 40*, e149. https://doi.org/10.1017/S0140525X16000194

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America, A, Optics, Image & Science, 7*, 1209–1217. https://doi.org/10.1364/JOSAA.7.001209

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320*, 264–265.

Narbutas, V., Lin, Y.-S., Kristan, M., & Heinke, D. (2017). Serial versus parallel search: A model comparison approach based on reaction time distributions. *Visual Cognition, 25*, 306–325. https://doi.org/10.1080/13506285.2017.1352055

Navalpakkam, V., & Itti, L. (2007). Search goal tunes visual features optimally. *Neuron, 53*, 605–617.

Newton, I. (1846). *The mathematical principles of natural philosophy (Andrew Motte, Trans.).* New York, NY: Daniel Adee.

Nothdurft, H. (1993). The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision Research, 33*, 1937–1958. https://doi.org/10.1016/0042-6989(93)90020-W

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: when it guides attention and when it does not. *Trends In Cognitive Sciences, 15*, 327–334. https://doi.org/10.1016/j.tics.2011.05.004

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science, 4*, 118–123. https://doi.org/10.1111/1467-8721.ep10772534

Palmer, J., & McLean, J. (1995, August). *Imperfect, unlimited-capacity, parallel search yields large set-size effects.* Paper presented at the annual meeting of the Society for Mathematical Psychology, Irvine, CA. Retrieved from http://faculty.washington.edu/jpalmer/files/palmer_mclean1995talk.pdf

Pashler, H. (1987a). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics, 41*, 191–201. https://doi.org/10.3758/BF03208218

Pashler, H. (1987b). Target-distractor discriminability in visual search. *Perception & Psychophysics, 41*, 285–302. https://doi.org/10.3758/BF03208228

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160–174. https://doi.org/10.1037/0096-3445.109.2.160

Prinz, W. (1986). Continuous selection. *Psychological Research, 48*, 231–238. https://doi.org/10.1007/BF00309087

Rangelov, D., Müller, H. J., & Zehetleitner, M. (2011a). Independent dimension-weighting mechanisms for visual selection and stimulus identification. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1369–1382. https://doi.org/10.1037/a0024265

Rangelov, D., Müller, H. J., & Zehetleitner, M. (2011b). Dimension-specific intertrial priming effects are task-specific: Evidence for multiple weighting systems. *Journal of Experimental Psychology:*

Human Perception and Performance, *37*, 100–114. https://doi.org/10.1037/a0020364

Rangelov, D., Müller, H. J., & Zehetleitner, M. (2013). Visual search for feature singletons: Multiple mechanisms produce sequence effects in visual search. *Journal of Vision*, *13*(3), 22. https://doi.org/10.1167/13.3.22

Rangelov, D., Müller, H. J., & Zehetleitner, M. (2017). Failure to pop out: Feature singletons do not capture attention under low signal-to-noise ratio conditions. *Journal of Experimental Psychology: General*, *146*, 651–671. https://doi.org/10.1037/xge0000284

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 849–860. https://doi.org/10.1037/0096-1523.18.3.849

Reeder, R. R., Olivers, C. N. L., & Pollmann, S. (2017). Cortical evidence for negative search templates. *Visual Cognition*, *25*, 278–290. https://doi.org/10.1080/13506285.2017.1339755

Roper, Z. J. J., Cosman, J. D., & Vecera, S. P. (2013). Perceptual load corresponds with factors known to influence visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1340–1351.

Rosenholtz, R., Huang, J., & Ehinger, K. A. (2012). Rethinking the role of top-down attention in vision: Effects attributable to a lossy representation in peripheral vision. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00013

Rosenholtz, R., Huang, J., Raj, A., Balas, B. J., & Ilie, L. (2012). A summary statistic representation in peripheral vision explains visual search. *Journal of Vision*, *12*(4). https://doi.org/10.1167/12.4.14

Sauter, M., Liesefeld, H. R., & Müller, H. J. (2019). Learning to suppress salient distractors in the target dimension: Region-based inhibition is persistent and transfers to distractors in a non-target dimension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0000691

Sauter, M., Liesefeld, H. R., Zehetleitner, M., & Müller, H. J. (2018). Region-based shielding of visual search from salient distractors: Target detection is impaired with same- but not different-dimension distractors. *Attention, Perception, & Psychophysics*, *80*, 622–642. https://doi.org/10.3758/s13414-017-1477-4

Smilek, D., Enns, J. T., Eastwood, J. D., & Merikle, P. M. (2006). Relax! Cognitive strategy influences visual search. *Visual Cognition*, *14*, 543–564. https://doi.org/10.1080/13506280500193487

Snodgrass, J. G., & Townsend, J. T. (1980). Comparing parallel and serial models: Theory and implementation. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 330–354. https://doi.org/10.1037/0096-1523.6.2.330

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences*, 12, 342–348. https://doi.org/10.1016/j.tics.2008.05.007

Steinman, S. B. (1987). Serial and parallel search in pattern vision? *Perception*, *16*, 389–398. https://doi.org/10.1068/p160389

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, *50*, 184–193. https://doi.org/10.3758/BF03212219

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*, 599–606. https://doi.org/10.3758/BF03211656

Theeuwes, J. (1994). Endogenous and exogenous control of visual selection. *Perception*, *23*, 429–440. https://doi.org/10.1068/p230429

Theeuwes, J. (1996). Parallel search for a conjunction of color and orientation: The effect of spatial proximity. *Acta Psychologica*, *94*, 291–307. https://doi.org/10.1016/S0001-6918(96)00003-0

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

Theeuwes, J., & Kooi, F. L. (1994). Parallel search for a conjunction of contrast polarity and shape. *Vision Research*, *34*, 3013–3016. https://doi.org/10.1016/0042-6989(94)90274-7

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

Töllner, T., Rangelov, D., & Müller, H. J. (2012). How the speed of motor-response decisions, but not focal-attentional selection, differs as a function of task set and target prevalence. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *109*, e1990–e1999. https://doi.org/10.1073/pnas.1206382109

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

Townsend, J. T. (1972). Some results concerning the indentifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, *25*, 168–199. https://doi.org/10.1111/j.2044-8317.1972.tb00490.x

Townsend, J. T. (1976). Serial and within-stage independent parallel model equivalence on the minimum completion time. *Journal of Mathematical Psychology*, *14*, 219–238. https://doi.org/10.1016/0022-2496(76)90003-1

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*, 46–54. https://doi.org/10.1111/j.1467-9280.1990.tb00067.x

Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *40*, 201–237. https://doi.org/10.1080/02724988843000104

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*, 15–48. https://doi.org/10.1037/0033-295X.95.1.15

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 459–478. https://doi.org/10.1037/0096-1523.16.3.459

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, *114*, 285–310. https://doi.org/10.1037/0096-3445.114.3.285

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

Tsotsos, J. K. (1990). Analyzing vision at the complexity level. *Behavioral and Brain Sciences*, *13*, 423–469. https://doi.org/10.1017/S0140525X00079577

Tsotsos, J. K., Culhane, S. M., Wai, W. Y. K., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*, *78*, 507–545. https://doi.org/10.1016/0004-3702(95)00025-9

VanRullen, R., Carlson, T., & Cavanagh, P. (2007). The blinking spotlight of attention. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *104*, 19204–19209. https://doi.org/10.1073/pnas.0707316104

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, *31*, 523–535. https://doi.org/10.1016/S0896-6273(01)00392-0

von Mühlenen, A., & Müller, H. J. (2000). Perceptual integration of motion and form information: Evidence of parallel-continuous processing. *Perception & Psychophysics*, *62*, 517–531. https://doi.org/10.3758/BF03212103

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

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, *19*, 1395–1407. https://doi.org/10.1016/j.neunet.2006.10.001

Wang, Z., & Klein, R. M. (2010). Searching for inhibition of return in visual search: A review. *Vision Research*, *50*, 220–228. https://doi.org/10.1016/j.visres.2009.11.013

Wang, B., & Theeuwes, J. (2018). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception & Performance*, 44, 13–17. https://doi.org/10.1037/xhp0000472

Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology*, *30*, 79–109. https://doi.org/10.1006/cogp.1996.0003

Ward, R., Duncan, J., & Shapiro, K. (1997). Effects of similarity, difficulty, and nontarget presentation on the time course of visual attention. *Perception & Psychophysics*, 59, 593–600. https://doi.org/10.3758/BF03211867-120

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

Watson, D. G., & Humphreys, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review*, *104*, 90–122. https://doi.org/10.1037/0033-295X.104.1.90

Watson, M. R., Brennan, A. A., Kingstone, A., & Enns, J. T. (2010). Looking versus seeing: Strategies alter eye movements during visual search. *Psychonomic Bulletin & Review*, *17*, 543–549. https://doi.org/10.3758/PBR.17.4.543

Wei, P., Yu, H., Müller, H. J., Pollmann, S., & Zhou, X. (2018). Differential brain mechanisms for processing distracting information in task-relevant and -irrelevant dimensions in visual search. *Human Brain Mapping, 40*, 110–124. https://doi.org/10.1002/hbm.24358

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

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

Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York, NY: Oxford University Press.

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental*

Psychology: Human Perception and Performance, 29, 483–502. https://doi.org/10.1037/0096-1523.29.2.483

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 419–433. https://doi.org/10.1037/0096-1523.15.3.419

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature*, *435*, 439–440. https://doi.org/10.1038/435439a

Wolfe, J. M., Klempen, N. L., & Shulman, E. P. (1999). Which end is up? Two representations of orientation in visual search. *Vision Research*, *39*, 2075–2086. https://doi.org/10.1016/S0042-6989(98)00260-0

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

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

Woodman, G. F., & Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 363–377. https://doi.org/10.1037/0096-1523.33.2.363

Wyble, B., Callahan-Flintoft, C., Chen, H., Marinov, T., Sarkar, A., & Bowman, H. (2018). Understanding visual attention with RAGNAROC: A reflexive attention gradient through neural attractor competition. *bioRxiv*, 406124. https://doi.org/10.1101/406124

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*, 72–75. https://doi.org/10.1038/23936

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

Zehetleitner, M., & Müller, H. J. (2010). Salience from the decision perspective: You know where it is before you know it is there. *Journal of Vision, 10*(14), 35. https://doi.org/10.1167/10.14.35

Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences*, *1339*, 154–164. https://doi.org/10.1111/nyas.12606

Zhaoping, L., & May, K. A. (2007). Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex. *PLOS Computational Biology, 3*, e62. https://doi.org/10.1371/journal.pcbi.0030062