

Contextual cueing of pop-out visual search: When context guides the deployment of attention

Thomas Geyer

Department Psychologie, Ludwig-Maximilians-Universität
München, Germany



Michael Zehetleitner

Department Psychologie, Ludwig-Maximilians-Universität
München, Germany



Hermann J. Müller

Department Psychologie, Ludwig-Maximilians-Universität
München, Germany, &
School of Psychology, Birkbeck College,
University of London, UK



Visual context information can guide attention in demanding (i.e., inefficient) search tasks. When participants are repeatedly presented with identically arranged ('repeated') displays, reaction times are faster relative to newly composed ('non-repeated') displays. The present article examines whether this 'contextual cueing' effect operates also in simple (i.e., efficient) search tasks and if so, whether there it influences target, rather than response, selection. The results were that singleton-feature targets were detected faster when the search items were presented in repeated, rather than non-repeated, arrangements. Importantly, repeated, relative to novel, displays also led to an increase in signal detection accuracy. Thus, contextual cueing can expedite the selection of pop-out targets, most likely by enhancing feature contrast signals at the overall-salience computation stage.

Keywords: contextual cueing, detection task, pop-out, attentional guidance

Citation: Geyer, T., Zehetleitner, M., & Müller, H. J. (2010). Contextual cueing of pop-out visual search: When context guides the deployment of attention. *Journal of Vision*, 10(5):20, 1–11, <http://journalofvision.org/content/10/5/20>, doi:10.1167/10.5.20.

Introduction

One of the key paradigms for investigating selective attention is visual search. In search studies, observers are presented with an array of stimuli, one of which is the task-relevant target and the others are irrelevant non-targets or distractors. The prime aim is to reveal the mechanisms that permit the target to be selected and responded to efficiently (Wolfe, 1998). One such mechanism is 'contextual cueing' (e.g., Chun & Jiang, 1998). Contextual cueing refers to the fact that when, over the course of an experiment, a target is repeatedly encountered at an invariant position within the same distractor arrangement ('context'), responses to the target are expedited relative to displays with non-repeated, random distractor arrangements. This suggests that the learnt distractor contexts guide focal attention to the target, even though observers are typically unable to explicitly recognize such predictive contexts. Spatial target-distractor associations learnt in a given visual context can persist across stimulus features changes (e.g., Chun & Jiang, 1998), task changes (e.g., Jiang & Song, 2005), or even modality changes (e.g., Nabeta, Ono, & Kawahara, 2003). Also, context effects can be observed for various types of display regularity, including spatial, temporal, featural, and semantic target-distractor contingencies (Chun,

2000; Goujon, Didierjean, & Marmèche, 2009; Ogawa, Watanabe, & Yagi, 2009). Although the neural ('memory') mechanisms underlying contextual cueing are not entirely clear at present, the hippocampus (e.g., Greene, Gross, Elsinger, & Rao, 2007) and, more generally, temporal lobe structures (i.e., medial temporal lobe, e.g., Manns & Squire, 2001; anterior temporal lobe, e.g., Chaumon, Hasboun, Baulac, Adam, & Tallon-Baudry, 2009) are likely to be implicated.

Response-based accounts of contextual cueing

While contextual cueing is recognized as an important adaptive control mechanism, there is a controversy as to the locus of the cueing effect: pre-attentive (perceptual) vs. post-selective (response-related). For example, Chun and Jiang (1998, Experiment 4) had their observers search for a target T among a variable number of distractor L's—a demanding search task requiring 'serial' shifts of focal attention. Examining the slopes of the RT/display size function—the prime measure of 'search efficiency' (Wolfe, 1998)—they found the search rates (per item) were reduced on trials with repeated, as compared to non-repeated, display arrangements (see also Tseng & Li, 2004, who

examined oculomotor as well as RT measures). This suggests that contextual cueing can enhance ‘early’ perceptual processes that provide the basis for attentional target selection, such as the computation of an overall-saliency map of the display, which is assumed to guide the allocation of focal attention in visual search (e.g., Wolfe, 1994).

However, other studies reported evidence for contextual cueing operating on ‘late’ processes (of response selection) following attentional target selection (e.g., Kunar, Flusberg, Horowitz, & Wolfe, 2007). Using stimuli similar to Chun and Jiang (1998), Kunar et al. (2007; Experiment 1) failed to find a reliable reduction in search rates for repeated relative to non-repeated displays. On the other hand, they observed contextual cueing even in much simpler, ‘parallel’-search tasks in which attention could be deployed efficiently towards the target location, making few demands on target selection. In their Experiment 3, observers had to detect a singleton color target, the only red letter among green letters, and respond to its identity (A vs. R); the distractor letters could all be either congruent (e.g., red A among green A’s) or incongruent (e.g., red A among green R’s) with the response to the target. Kunar et al. (2007) found contextual cueing only for congruent, but not for incongruent, trials. Because contextual cueing interacted with the response selection manipulation (and the effects of distractor congruence vs. incongruence are typically attributed to the response selection stage—see, e.g., Eriksen & Eriksen, 1974), Kunar et al. (2007) concluded that contextual cueing, rather than enhancing focal-attentional target selection, influences later processes of response selection. More precisely, they surmised that the presentation of a repeated display may lower the decision threshold required for initiating a response (e.g., Ratcliff & Smith, 2004), leading to benefits in RTs relative to non-repeated displays.

An alternative proposal to Kunar et al. (2007): Contextual cueing of target selection

However, although Kunar et al.’s (2007) findings suggest that contextual cueing modulates response selection, they do not necessarily exclude other sources of the cueing effect. It remains possible that in feature singleton search, there is a ‘race’ between bottom-up (i.e., stimulus) and top-down (i.e., memory) guidance signals for the allocation of attention. As the target is a highly salient, pop-out item, selection may be driven mainly by bottom-up guidance signals; that is, memory-based guidance signals (contextual cueing) could be ‘masked’ by the relatively fast computation of bottom-up saliency signals. Note that such a view has recently been suggested by Kunar and colleagues themselves (Kunar, Flusberg, & Wolfe, 2008), at least for inefficient (time-demanding) search tasks with targets defined by a conjunction of (shape) features.¹

Furthermore, despite its seemingly clear additive-factors logic, ‘compound’ search tasks (in which the search-critical

target attribute is different from the response-critical attribute), such as that employed by Kunar et al. (2007), encounter a general difficulty in dissociating (late) response-based effects from (early) search-based effects (see, e.g., Töllner, Gramann, Müller, Kiss, & Eimer, 2008). Examining stimulus- and, respectively, response-related components of the EEG, Töllner et al. (2008) showed that, in a compound-search task, certain stimulus–response (S–R) linkages are implicitly ‘favored’ by the post-selective processing system (in particular, a repeated stimulus is associated with the expectation of a repeated response, and a changed stimulus with a changed response), even if the stimuli and responses are in reality uncorrelated (as was the case in Kunar et al., 2007). Arguably, such linkages make it hard to unequivocally infer the locus of memory effects (here: contextual cueing) on the basis of RT data alone.

In summary, previous studies failed to demonstrate guidance of attention from contextual cueing in efficient, singleton feature search (though see Kunar et al., 2008, for positive findings at least for inefficient searches)—most likely because memory processes were too slow to influence target selection, which was therefore determined mainly by bottom-up (i.e., stimulus) factors. Also, because of (implicit) S–R linkages, compound-search tasks (such as that employed by Kunar et al., 2007) do not permit as clear-cut a dissociation of ‘early’ from ‘late’ effects as the compound-task logic suggests. Thus, it remains an open issue whether contextual cueing modulates ‘early’ perceptual coding processes leading up to target selection in efficient, pop-out search.

Arguably, this issue is best addressed by using a search task in which (i) the target is a feature singleton presented among homogeneous distractors (as in Experiment 3 of Kunar et al., 2007), and (ii) observers are simply required to discern the presence versus the absence of a target in the display (rather than to discriminate a response-critical attribute of the target, as has been typically the case in contextual-cueing studies), while (iii) being provided with some additional time to process the spatial layout of the search display. In simple detection of ‘pop-out’ stimuli, target-present and -absent decisions would involve similar processes, namely, to establish the presence versus the absence of an above-threshold (target) signal on the overall-saliency map (e.g., Müller, Heller, & Ziegler, 1995; Töllner, Zehetleitner, Krümmenacher, & Müller, *in press*; Wolfe, 1998). If, under these conditions, RTs were faster for repeated than for non-repeated displays, this effect is unlikely to be due to facilitated post-selective stimulus processing (e.g., Kunar et al., 2007). Rather, it would be attributable to retrieval of a learnt distractor layout (elicited by a repeated display) reducing the time required by the target to achieve saliency on the overall-saliency map and summon focal attention (e.g., Brady & Chun, 2007; Geyer, Shi, & Müller, *in press*).

Furthermore, if contextual cueing influences early saliency computations, this would predict an effect of contextual cueing on the accuracy, in terms of signal detection

measures (e.g., Green & Swets, 1966), of the target-present versus -absent discrimination under (i) conditions in which display exposure is time-limited without stress on response speed, or (ii) conditions with extreme stress on response speed, such as when responses have to be made within a preset, short time window starting with display onset. That is, benefits in signal detectability deriving from the presentation of repeated, as compared to non-repeated, displays would argue in favor of contextual cueing influencing early stages of perceptual coding, rather than late stages of response selection. This predicated on the assumption that RT measures index both perceptual and post-perceptual processing stages, whereas accuracy with brief displays or stress on response speed index only perceptual stages (e.g., Pashler, 1989; Santee & Egeth, 1982).

The present experiments were designed to test these predictions and examine for an early, perceptual source of contextual cueing. Experiment 1 employed a standard search RT task, that is, observers could view the search display as long as it took to make a speeded target-present/absent decision. Variably across trials, the target was either a color- or an orientation-defined singleton. That is, the task required cross-dimensional singleton feature search (see Found & Müller, 1996), where repetition (vs. change) of the target-defining across consecutive trials dimension has been shown to expedite pre-selective perceptual coding processes (e.g., Töllner et al., 2008; Töllner, Zehetleitner, Gramann, & Müller, *in press*). The aim was to ascertain whether a contextual cueing effect would also be evident under parallel-search conditions and whether it would interact with dimension repetition/change effects (for which an ‘early’ locus has already been demonstrated). By contrast, Experiment 2 was designed to measure detection accuracy (rather than RT) under two conditions: in

Experiment 2A, the search displays were shown only briefly, for an individually adjusted duration, and then masked prior to observers making a target-present/-absent decision; in Experiment 2B, stress on response speed was introduced by forcing observers to make a target-present/-absent response within an individually adjusted time window starting with the presentation of the search display. If contextual cueing effects could be demonstrated, in terms of RT and accuracy effects, in these simple-detection tasks, this would argue in favor of an early, perceptual locus of contextual cueing (without, of course, ruling out later sources contributing to the cueing effect).

Experiment 1

Method

Participants

13 unpracticed observers took part in Experiment 1 (11 female; mean age: 27.6 years). All observers reported normal or corrected-to-normal (color) vision. They were naïve as to the purpose of the study and gave informed consent prior to their participation. Observers received a course credit for their participation.

Stimuli and procedure

Figure 1 illustrates the sequence of display frames on a given trial. A trial started with the presentation, for 700 ms, of gray ‘placeholder’ squares marking the locations of the

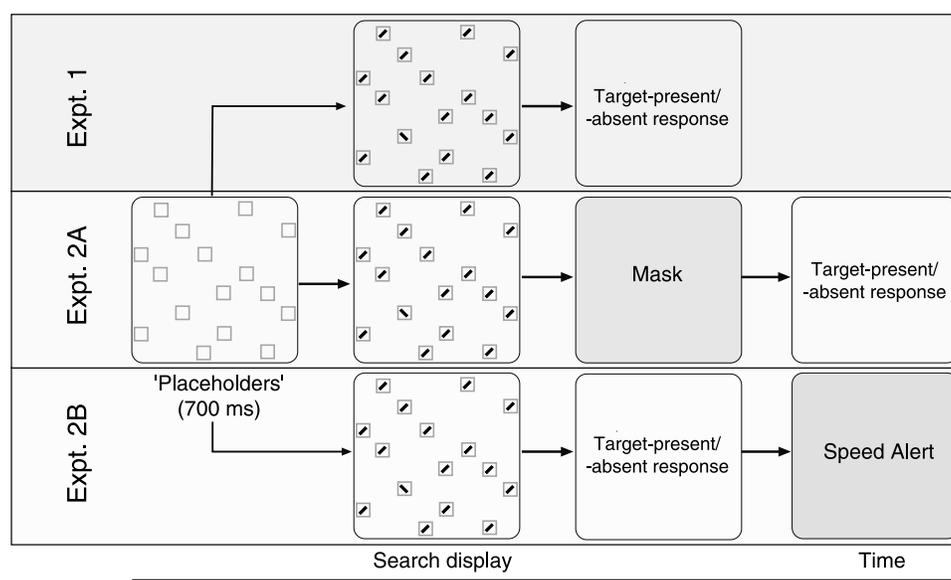


Figure 1. Illustration of the methods used in Experiments 1 and 2 (A, B). Explanations are provided in the text. In the figure, the target is the left-tilted orientation-singleton.

subsequent search items (size: $0.80^\circ \times 0.80^\circ$ at a viewing distance of 57 cm; brightness: 9.1 cd/m^2 ; see Ogawa & Kumada, 2008, for a similar procedure). The purpose of this placeholder display was to provide observers with additional ‘preview’ time to process the spatial layout of the search display, without being able to actually search for the target. At the end of the preview period, the search items were presented within the placeholder squares. The distractors were green and (relative to the vertical) 45° right-tilted bars ($0.35^\circ \times 0.35^\circ$; 9.1 cd/m^2). The singleton-feature target, if present, was either color-defined (a red colored bar oriented 45° to the right; $0.35^\circ \times 0.35^\circ$; 9.1 cd/m^2) or orientation-defined (a green colored bar oriented 45° to the left; $0.35^\circ \times 0.35^\circ$; 9.1 cd/m^2). The target-defining feature was determined randomly across trials. All stimuli were presented against a white background (31.9 cd/m^2). Observers’ task was to press one key on the computer keyboard when a target was present and another key when it was absent.

Experiment 1 consisted of two sessions, with a total of 1,536 trials. Observers had unlimited time for their response. The dependent variable was RT. The independent variables were display type (repeated vs. non-repeated), target (present vs. absent), and display size (9 vs. 16 items; constant across an experimental session; counterbalanced across observers). The latter manipulation was intended to demonstrate that the present task permitted ‘efficient’ search (cf. Wolfe, 1998).

There were eight randomly arranged target-distractor layouts, generated at the beginning of the experiment, which were repeated on randomly selected trials throughout the experiment. Note that, on target-present trials, the target appeared always at the same location in any such (repeated) configuration. The same repeated arrangements that were displayed on target-present trials were also used on target-absent trials, however, with the target replaced by a distractor. This was done in order to prevent observers learning to associate a certain (repeated) arrangement with a certain response (such as “this arrangement means target present”). Furthermore, the same (eight) repeated arrangements that were used for singleton color targets were also used for singleton-orientation targets. This means that singleton color and singleton orientation targets were equally likely to appear in any of the eight repeated arrangements. Presenting singleton color and singleton orientation targets among the same configurations appeared acceptable as prior studies had already shown contextual cueing to operate independently of the (featural) identity of the target (e.g., Chun & Jiang, 1998). However, whether or not contextual associations can be learned (and subsequently retrieved) under conditions of variable dimension-to-context mappings remains an open issue.

Non-repeated target-distractor arrangements were generated on-line on a given experimental trial. In half the trials, a repeated arrangement was presented, and a non-repeated arrangement in the other half. To equate target location repetition effects between repeated and non-

repeated displays, the target appeared equally often at each of 16 possible locations throughout the experiment: eight locations were used in repeated and the other eight in non-repeated configurations. In Experiment 1, the 9 and 16 search items appeared on an underlying matrix of 6×6 and 8×8 stimulus locations, respectively, to equate stimulus density across these display size conditions.

After the experiment, observers performed a recognition test, which was designed to examine whether they would be able to explicitly discern repeated from non-repeated displays. The recognition test consisted of 64 trials: 32 with predictive displays and 32 with non-predictive displays (generated online), presented in randomized order. Given this, the chance rate for recognizing a repeated display was 50%.

Results and discussion

RT performance

Data analysis was performed using R (R Development Core Team, 2007). For each experimental condition (display type \times target \times display size), RTs outside the range of ± 2.5 standard deviations from mean RTs were discarded as ‘outliers’ (overall, 2.3% of all trials). Trials on which a response error occurred were also excluded from the analysis (1.7% of trials). Separate repeated-measures ANOVAs of the miss and false alarm rates, each with display type and display size as factors, failed to reveal any significant effects [all F ’s < 1].

Mean RTs were examined by a 2 (target-present vs. -absent) \times 2 (repeated vs. non-repeated display) \times 2 (9 vs. 16 display items) repeated-measures ANOVA. There were significant main effects of target [$F(1,12) = 26.87, p < .01$] and display type [$F(1,12) = 18.84, p < .01$]. No other effects were significant. Importantly: there were no significant effects involving display size, that is, increasing the display size from 9 to 16 items did not influence the RTs (9- vs. 16-item displays: target-present: 561 vs. 546 ms; target-absent: 649 vs. 663 ms; see also Figure 2), indicating efficient, spatially parallel search.

Target presence was determined faster than establishing target absence [main effect of target: 553 vs. 656 ms], and RTs were overall faster for repeated than for non-repeated displays [main effect of display type: 599 vs. 611 ms; see also Figure 2]. Although small, the reliable RT benefit for repeated over non-repeated displays suggests that contextual cueing operates in simple feature singleton detection tasks.

One additional analysis was conducted to examine how the cueing effect developed over time (see Figure 2). For this, RTs to repeated and non-repeated target-distractor arrangements in 9- and 16-item displays were re-analyzed as a function of experimental block (for each display size, there were 12 blocks, each comprising of 64 trials) and entered in a 2 (repeated vs. non-repeated display) \times 12 (block 1 vs. block 2 vs. ... block 12) \times 2 (9 vs. 16 items)

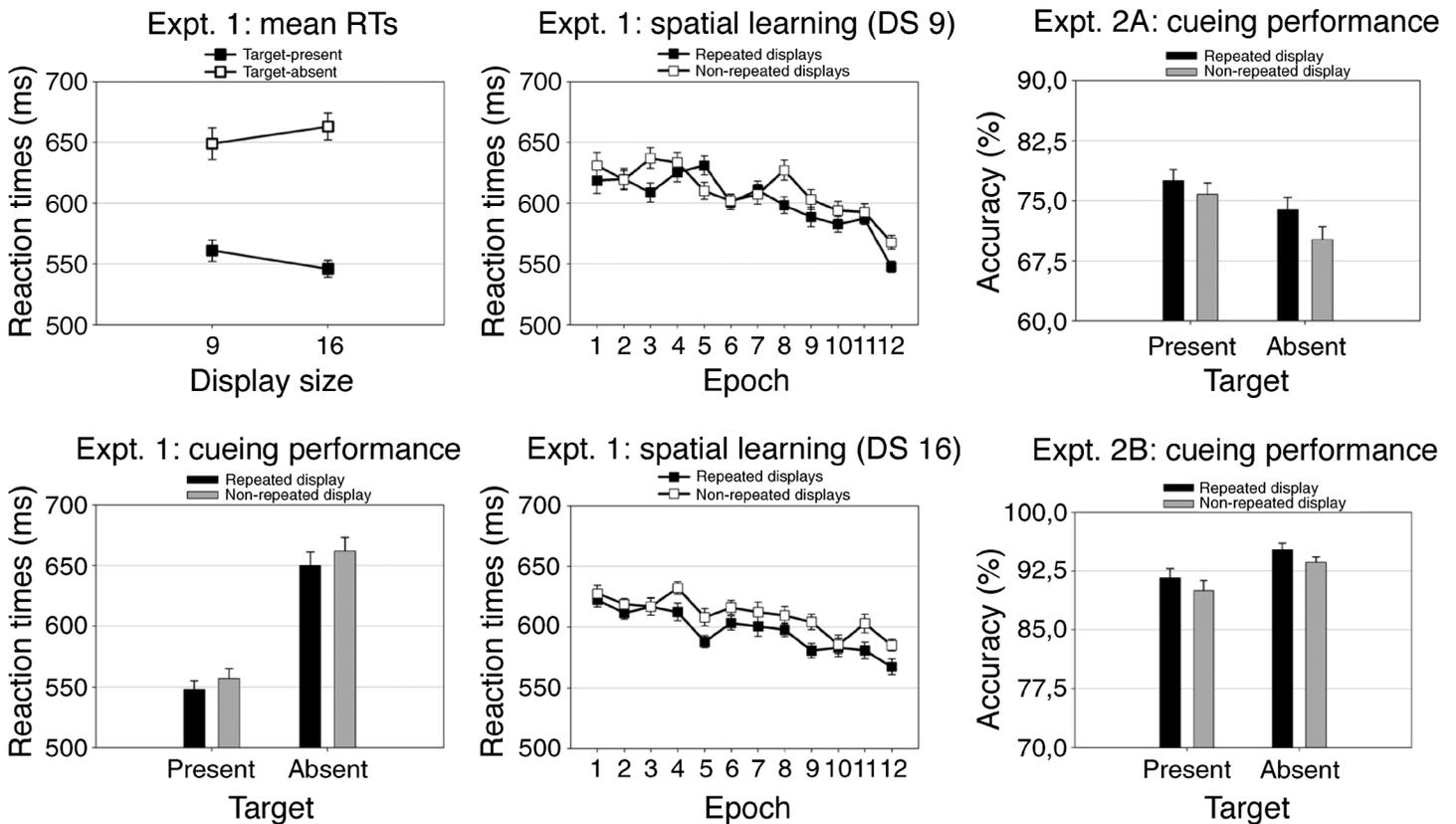


Figure 2. Results of Experiments 1 (left & middle panels) and 2 (right panel). Left panel: mean RTs (ms) for target-present and -absent trials as a function of display size (9- vs. 16-item displays; top), and mean RTs (ms) for repeated and non-repeated displays as a function of target (present vs. absent; bottom). Middle panel: mean RTs (ms) for repeated and non-repeated 9-item displays (top) and 16-item displays (bottom) as a function of block (1 through 12). The data are collapsed across target-present and -absent trials. Right panel: mean correct responses (%) for repeated and non-repeated displays as a function of target (present vs. absent) in Experiment 2A (top) and Experiment 2B (i.e., ‘speed task’; bottom). In all panels, the small black vertical bars represent the mean standard error.

ANOVA. This analysis (only) revealed main effects of block [$F(11,132) = 5.74, p < .01$] and display type [$F(1,12) = 21.04, p < .01$]. The main effect of block was owing to a significant improvement in response speed as the experiment progressed (block 1: 624 ms; block 12: 567 ms). The main effect of display type was due to RTs being faster for repeated than for non-repeated displays (see above). Although the block \times display type was non-significant [$F < 1$], the main effect of display type confirmed that observers detected targets more rapidly in repeated than non-repeated configurations. Finer-grained analyses showed that there were no significant RT differences between repeated and non-repeated 9-item displays in blocks 1 and 2 (both p 's $> .10$) and 16-item displays in blocks 1 through 3 (all p 's $> .10$).

Recognition performance

Observers' ability to explicitly recognize repeated displays was assessed by calculating the hit rate (correct classification of repeated displays as ‘repeated’, i.e., seen before) and false alarm rate (incorrect classification of non-

repeated displays as ‘repeated’) on recognition trials. More hits than false alarms would indicate that observers could tell apart repeated from non-repeated displays. However, the mean hit rate was comparable to the mean false alarm rate [.43 vs. .42; one-tailed $t(1,12) = .02, p = .49$].

One additional analysis was conducted to further explore the relationship between explicit recognition performance and contextual cueing. In this analysis, the sensitivity measure d' [i.e., $z(\text{hit}) - z(\text{false alarm})$] was calculated for each observer and, on the basis of a d' median split, observers were assigned to an ‘aware group’ (individual-observer $d' > \text{median}$; mean d' for ‘aware’ observers: .42) or, respectively, a ‘non-aware group’ (individual-observer $d' < \text{median}$; mean d' for ‘non-aware’ observers: $-.36$). Of note, mean d' in Experiment 1 was zero ($d' = .00$). Importantly, the contextual cueing performance did not differ significantly between ‘aware’ and ‘unaware’ observers: 12 vs. 10 ms [one-tailed $t(1,6) = .35, p = .37$]. In sum, observers did not produce significantly more hits than false alarms. In addition, even ‘unaware’ observers exhibited contextual cueing, and the size of the cueing effect was uninfluenced by whether or not an observer did

show ‘awareness’ of repeated displays. This pattern may be taken as evidence that contextual cueing in feature singleton detection search is not systematically influenced by explicit awareness of display layout.

Experiment 2

Experiment 2 was a close replication of Experiment 1, except for the differences set out below.

Participants

23 new observers participated in Experiment 2 (9 female; mean age: 29.1 years). They were paid at a rate of €8.00 (US\$12.00) per session.

Stimuli and procedure

Experiment 2 comprised of two conditions (see also Figure 1). In Experiment 2A (10 observers performing one session of 768 trials), the search displays were followed by a mask display prior to observers’ target-present/absent response. The masking stimuli consisted of ten randomly oriented (0° – 360°) and colored (gray, red, and green) lines (sized 0.35° – 0.80°) presented at each possible stimulus location. The stimulus onset asynchrony (SOA) between the search displays and masks was determined individually, in a pre-experimental session, using an adaptive staircase procedure that aimed at an accuracy level of 75% correct responses (averaged across target-present/absent trials). The mean SOA established in this way was 90 ms. Observers were instructed to respond as accurately as possible, without stress on response speed. In Experiment 2B (with 13 different participants performing one session of 768 trials), observers performed the search task under two different response instructions: speed and accuracy, respectively. In the ‘speed task’, observers were instructed to respond as fast as possible. To reinforce this, a warning message (German words ‘zu langsam!!!’, meaning ‘too slow!!!’) was presented if the RT on a given trial was above the 75th percentile of the individual RTs, established in a pre-experimental session. Across all participants, the mean 75th percentile was 845 ms. In the ‘accuracy task’, observers were instructed to respond as accurately as possible. To encourage this, the display screen remained blank for 8 sec following an error (target miss or false-alarm) response. Note that for both the speed and accuracy tasks, the search items remained visible until observers executed their target-present/absent response. The speed and accuracy tasks were run in separate experimental blocks, presented in randomized order across the experiment. Observers were informed about the

respective task at the beginning of each block. In Experiment 2 (A, B), the dependent variable was accuracy, and the independent variables were display type and target (display size was held constant at 16 items, because Experiment 1 had revealed no differences in contextual cueing between 9- and 16-item displays—see above). The 16 items were presented on a grid of 8×8 possible locations. In both Experiments 2A and 2B, the search task was followed by a recognition test of a total of 64 trials (32 trials repeated displays, 32 trials non-repeated displays; randomized order).

Results and discussion

Accuracy performance

If contextual cueing facilitates ‘early’ attentional processes, then signal detection accuracy, assumed to reflect the operation of pre-attentive sensory coding stages, should be higher for repeated than for non-repeated displays. This was explored by subjecting the error rates (target misses and false alarms) in Experiments 2A and 2B (i.e., speed task) to separate 2 (target-present vs. -absent) \times 2 (repeated vs. non-repeated displays) ANOVAs.² Both ANOVAs revealed main effects of display type [Experiment 2A: $F(1,9) = 16.36$, $p < .01$; Experiment 2B: $F(1,12) = 9.57$, $p < .01$]. In addition, for Experiment 2B, the main effect of target was significant [$F(1,12) = 7.32$, $p < .05$], with miss rates being higher than false alarm rates (9.2 vs. 5.6%). Concerning the main effects of display type, when the search displays were presented only briefly and then masked (Experiment 2A), response accuracy was higher for repeated than for non-repeated displays (target-present: 77.5% vs. 75.8%; target-absent: 73.9% vs. 70.2% of correct responses). Similarly, when observers were forced to respond as fast as possible (Experiment 2B), accuracy was larger for repeated than for non-repeated displays (target-present: 91.6% vs. 90.0%; target-absent: 95.2% vs. 93.6% of correct responses). This pattern of results is important because it suggests that contextual cueing modulates ‘early’ attentional processes involved in target selection (as indexed by detection accuracy). Additional ANOVAs on response speed (with target and display type as factors) confirmed this assumption: detection RTs were significantly faster for repeated than for non-repeated displays [Experiment 2A: 652 vs. 667 ms; $F(1,9) = 5.11$, $p < .05$; Experiment 2B: 531 vs. 541 ms; $F(1,12) = 8.76$, $p < .01$] (in addition to main effects of target).³

To further explore participants’ performance, the signal detection sensitivity measure d' and the response criterion β (see, e.g., Macmillan & Creelman, 1991) were computed and subjected to two additional ANOVAs, each with display type (repeated vs. non-repeated) as factor. Although the above analyses of the target miss and false-alarm rates provide evidence that contextual cueing increases

perceptual accuracy (i.e., hit rates were increased and false-alarm rates reduced for repeated as compared to non-repeated displays), it remains possible that contextual cueing also influences decision-making factors, such as lowering the threshold required for initiating a detection response, which would be indicated by a more ‘liberal’ (i.e., smaller) response criterion β for repeated than for non-repeated displays.

Unsurprisingly (given the hit and false-alarm rate results), d' prime was significantly higher for repeated than for non-repeated displays [Experiment 2A: 1.61 vs. 1.36; $F(1,9) = 15.95$, $p < .01$, $MSE = .02$; Experiment 2B: 3.36 vs. 3.05; $F(1,12) = 7.09$, $p < .05$, $MSE = .08$]. By contrast, the response criterion β was uninfluenced by the display type manipulation [β for ‘old’ and ‘new’ displays: Experiment 2A: 1.2 vs. 1.1 [$F(1,9) = .42$, $p = .53$, $MSE = .01$]; and Experiment 2B: 2.0 vs. 1.5 [$F(1,12) = 2.35$, $p = .15$, $MSE = .92$].] If anything, responses to ‘old’ displays tended to be more conservative than responses to ‘new’ displays (indicated by numerically higher β values in Experiments 2A and 2B for ‘old’ displays). This pattern suggests that contextual cueing in visual pop-out (detection) search enhances ‘early’ processes of target selection, rather than ‘late’ processes of response selection.

Summary and conclusions

Early locus of contextual cueing

Experiments 1, 2A, and 2B showed that when observers were required to discern the presence versus the absence of a singleton feature target, search performance was enhanced, in terms of both response speed (Experiment 1) and accuracy (Experiments 2A and 2B), when the search items were presented in identical (repeated), rather than random (non-repeated), arrangements. This is in line with the idea that contextual cueing facilitates focal-attentional target selection (e.g., Chun & Jiang, 1998; see also Peterson & Kramer, 2001, and Brockmole & Henderson, 2006, who showed that finding the target required fewer eye movements—i.e., overt attention shifts—in ‘old’ relative to ‘new’ displays), while not ruling out that later, response selection processes may also be modulated by the cueing (in particular, in demanding, ‘serial’ search tasks; e.g., Kunar et al., 2007). One way how contextual cueing may work is that contextual memory reinforces target saliency coding at some stage in the computation of the overall-saliency map (see Brady & Chun, 2007, and Geyer et al., *in press*, for similar conclusions). This account is illustrated in Figure 3. It assumes that target selection operates from an overall-saliency (or master) map of the display, with the relative activation of the master map units determining the allocation of attention (e.g., Koch & Ullman, 1985). The master map units integrate, in parallel,

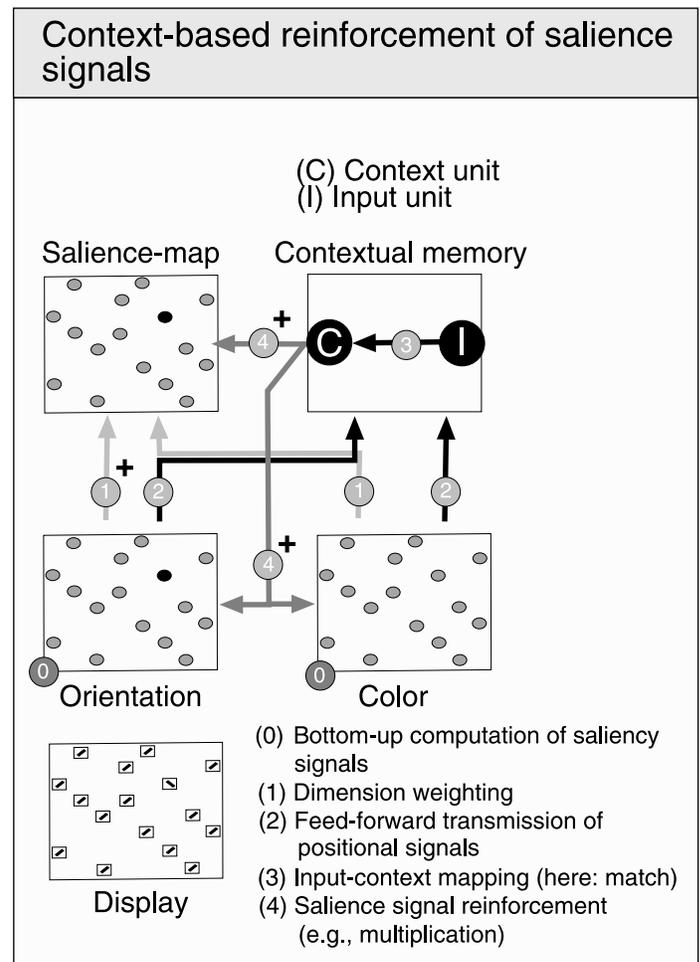


Figure 3. A psychological model of target selection in feature-singleton accommodating memory-based search guidance (i.e., contextual cueing). Explanations are provided in the text.

the output of dimension-specific feature contrast signal computations via spatiotopic connections between feature analyzers and the master map (indexed by ‘0’ in the model). Prior to information integration at the level of the master map, the dimension-specific feature contrast signals are weighted dependent on the intertrial ‘history’: in cross-dimension search, dimensional weights are dynamically increased for target-defining dimensions, and decreased for non-defining dimensions (e.g., Müller et al., 1995; Zehetleitner & Müller, *in preparation*; ‘1’ in the model). An important assumption of the model is the idea of spatiotopic connections also between feature analyzers and the (hippocampal) memory underlying contextual cueing. Following the bottom-up computation of saliency signals, active feature analyzers send, in parallel, a pattern of position signals to the ‘input units’ of the contextual memory (‘2’ in the model). If this input matches some stored representation in the memory’s ‘context units’ (i.e., if observers encounter a repeated display; ‘3’ in Figure 3), the target location specified by the activated context memory (signaled by an ‘output unit’) will receive an

amplification signal, either via the feature analyzers (enhancing feature contrast coding) or directly to the overall-saliency map ('4' in the model). As a result, a memory-based guidance signal (contextual cueing) may reinforce bottom-up saliency computations, expediting and enhancing target pop-out.

This argument presupposes that contextual cueing, that is, context-based reinforcement of salience signals, occurs relatively late after the presentation of the search display. Consistent with this, examining intracranial event-related potentials, Olson, Chun, and Allison (2001) found presentation of repeated (vs. non-repeated) displays to be associated with activity modulations in visual brain areas (e.g., V1, V2) that were observed only around 200 ms post search display onset. Based on this, it is suggested that contextual cueing acts as a 'late' feedback mechanism influencing perceptual processing.

Reinforcement of salience signals at the level of the master map

Chun and Jiang (1998) surmised that the positional memory underlying contextual cueing is identity-independent (see also Jiang & Song, 2005, and Nabet et al., 2003). Evidence for this was their finding (in Experiment 2) that contextual associations acquired in the first half of the experiment continued to expedite RT performance even when the distractors changed their defining features (shape), but maintained their locations, in the second half.⁴ That is, contextual cueing expedites a processing stage that codes *where* a critical difference is in the visual display, but not *what* this difference is. This is consistent with signaling on the overall-saliency map: an active saliency map unit only signals that there is a feature difference around the corresponding display location, not what the difference is (dimensionally or featurally). This makes the master map a target candidate for the effect of contextual cueing.

The present findings bear on this issue in at least two respects. First, Experiments 1 and 2 (A, B) revealed evidence of contextual cueing when the very same repeated layouts were used for singleton color and orientation targets (which occurred in random order across trials). Recall that the same target-distractor arrangements were used for both types of targets throughout the course of the present experiments (there was no change between the first and second experimental halves, as in Chun & Jiang's, 1998, Experiment 2). This thus suggests that not only the expression (Chun & Jiang, 1998), but also the learning of spatial target-distractor arrangements is (dimension and feature) identity-independent.

In line with this assumption are the results of a control experiment which, contrary to the present Experiment 1, used four different arrangements for singleton color and, respectively, orientation targets (fixed dimension-context

mapping). Yet the contextual cueing benefit was of the same magnitude as that observed with the variable dimension-context mapping in the present Experiment 1: 14 vs. 12 ms [data collapsed across target-present and -absent trials; $p > .10$]. If the memory underlying contextual cueing were contingent on the dimension of the target, and expressing its influence at a dimension- (or feature-) specific level of coding, one would have expected the cueing effect to be more pronounced in the fixed than in the variable dimension-context-mapping condition (e.g., presenting color and orientation targets within the same contexts could have made learning less robust). Thus, the finding of a comparable cueing effect between fixed and variable dimension-context mapping conditions adds to the existing evidence that contextual cueing acts on identity-independent representations, such as the overall-saliency map (cf. Wolfe, 1994).

Second, to examine for the locus of the cueing effect more directly, we re-analyzed the data (of Experiments 1, 2A, and 2B) dependent on whether observers encountered a repeated or a non-repeated display and, importantly, whether the dimension of the singleton target was the same or different relative to the previous trial (e.g., Found & Müller, 1996). Typically, search performance is improved following repetitions, as compared to changes, of the target-defining dimension—an effect which has been shown to facilitate electrophysiological markers of focal-attentional target selection (specifically, the so-called N2pc component; e.g., Töllner et al., 2008; Töllner, Zehetleitner, Gramann et al., *in press*). Given this, and following Sternberg's (1969) additive-factors logic, if the cueing effect aids processes at the focal-attentional stage, then it should interact with the dimension repetition manipulation. For example, it is possible that following a change of the target's defining dimension (e.g., a color target followed by an orientation target), salience computations are less efficient (i.e., temporally extended), as a result of which there would be more time for the cueing effects to 'kick in' and aid performance. If so, contextual cueing should have a larger effect with different- relative to same-dimension targets. This was examined, for the target-present trials, by separate 2 (repeated arrangement, non-repeated arrangement) \times 2 (same dimension, different dimension) repeated-measures ANOVAs which revealed significant main effects of stimulus arrangement [all F 's > 4] and stimulus dimension [all F 's > 37]. Importantly, the stimulus arrangement \times dimension interaction was also significant, at least for Experiments 2A and 2B [both F 's > 4]. This means that when observers encountered a repeated relative to a non-repeated display, the benefit in performance accuracy was larger for different- compared to same-dimension targets (2.0 vs. 1.3%, data combined across Experiments 2A and 2B). Although this interaction was only marginal for the RT performance in Experiment 1 [$2 > F > 1$], there was a tendency for the cueing effect being larger for different- than for same-dimension targets

(13 vs. 6 ms). Thus, contextual cueing was modulated by the dimensional manipulation, suggesting both contextual cueing and dimensional weighting effects exert their influences at the overall-salience computation stage.

Contextual cueing in pop-out search

Across all experiments, contextual cueing was observed even though the search to be performed was highly efficient. This finding is consistent with recent evidence that contextual cueing is not necessarily dependent on ‘serial’ deployments of attention (e.g., Ogawa & Kumada, 2008), as long as observers are provided with ‘sufficient’ time (here: 700 ms) to encode and learn the display arrangements. Nevertheless, contextual cueing has been shown more consistently with ‘serial’ than with ‘parallel’ tasks (e.g., Jiang & Song, 2005). The reason for this may be that all search tasks, whether ‘serial’ or ‘parallel’, involve some component of contextual cueing, but cueing effects are masked in efficient-search tasks by fast, bottom-up driven target individuation processes (compared to which context-based guidance signals are relatively slow). Alternatively, or in addition, contextual learning may be impeded under pop-out search conditions because the display items (which are normally removed upon the response) are not available for the (minimum) time required for contextual learning to take place. Future research is required to distinguish between these alternatives.

In conclusion, presentation of a singleton feature target in a repeated display yielded significant RT gains (Experiment 1), demonstrating the operation of contextual cueing in a simple (‘efficient’) pop-out detection task. Furthermore, there were gains in detection accuracy with briefly presented search displays (Experiment 2A) and task performance under strict response deadline conditions (Experiment 2B). This pattern of effects argues that contextual cueing facilitates perceptual processes associated with focal-attentional target selection.

Acknowledgments

This research was supported by Deutsche Forschungsgemeinschaft (DFG) Research Group (FOR 480), Research Project (GE 1889/1-1), and CoTeSys Excellence Cluster (142) grants. Correspondence concerning this article should be addressed to: Thomas Geyer, Ludwig Maximilian University Munich, Department of Psychology, Leopoldstraße 13, 80802 München, Germany (email: geyer@lmu.de).

Commercial relationships: none.
Corresponding author: Thomas Geyer.
Email: geyer@lmu.de.

Address: Department Psychologie, Lehrstuhl für Allgemeine und Experimentelle Psychologie, Ludwig-Maximilians-Universität München, Leopoldstraße 13, 80802 München, Germany.

Footnotes

¹The critical finding was that contextual cueing gains (in terms of the slopes of functions relating RT to display size) were larger when participants were provided with more time to search the displays (e.g., when the ‘T’ target and the various, orthogonally oriented ‘L’ distractors were presented on a background that consisted of multiple line segments with vertical and horizontal orientations relative to a background that contained no lines).

²In Experiment 2B, RTs and accuracy were influenced by task instructions. First, mean RTs were faster in the speed than in the accuracy task [535 vs. 567 ms; 32-ms effect; one-tailed $t(1,12) = 3.18, p < .01$]. Second, response errors were less frequent in the accuracy than in the speed-task [2.9 vs. 7.6%, 4.7% effect; one-tailed $t(1,12) = 4.66, p < .01$].

³Concerning the relationship between explicit recognition and contextual cueing, Experiment 2 (A, B) exhibited a similar pattern of results to Experiment 1: the mean hit rates did not differ from the mean false-alarm rates [Experiment 2A: .42 vs. .40; one-tailed $t(1,9) = .16, p = .44$; Experiment 2B: .46 vs. .42; one-tailed $t(1,12) = .65, p = .26$]; the sensitivity measure d' was close to zero (mean d' in Experiments 2A and 2B: .01 and .10, respectively); and the contextual-cueing effects did not differ significantly between ‘aware’ and ‘unaware’ observers: Experiment 2A, 2.1% vs. 3.3% [one-tailed $t(1,4) = .80, p = .23$]; Experiment 2B, 1.1% vs. 2.2% [one-tailed $t(1,6) = 1.07, p = .16$].

⁴Note that the RT benefits arising from repeated displays were of the same size in the epochs immediately preceding and immediately following the change of distractor shapes.

References

- Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cueing. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 798–815. [PubMed]
- Brockmole, J. R., & Henderson, J. M. (2006). Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *59*, 1177–1187. [PubMed]

- Chaumon, M., Hasboun, D., Baulac, M., Adam, C., & Tallon-Baudry, C. (2009). Unconscious contextual memory affects early responses in the anterior temporal lobe. *Brain Research, 1285*, 77–87. [PubMed]
- Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*, 170–178.
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28–71. [PubMed]
- Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 224–234. [PubMed]
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143–149.
- 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. [PubMed]
- Geyer, T., Shi, Z., & Müller, H. J. (in press). Contextual cueing in multi-conjunction search is dependent on color- and configuration-based intertrial contingencies. *Journal of Experimental Psychology: Human Perception and Performance*.
- Goujon, A., Didierjean, A., & Marmèche, E. (2009). Semantic contextual cuing and visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 50–71.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, England: Wiley.
- Greene, A. J., Gross, W. L., Elsinger, C. L., & Rao, S. M. (2007). Hippocampal differentiation without recognition: An fMRI analysis of the contextual cueing task. *Learning and Memory, 14*, 548–553. [PubMed]
- Jiang, Y., & Song, J.-H. (2005). Spatial context learning in visual search and change detection. *Perception & Psychophysics, 67*, 1128–1139. [PubMed] [Article]
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4*, 219–227. [PubMed]
- Kunar, M. A., Flusberg, S., Horowitz, T. S., & Wolfe, J. M. (2007). Does contextual cuing guide the deployment of attention? *Journal of Experimental Psychology: Human Perception and Performance, 33*, 816–828. [PubMed]
- Kunar, M. A., Flusberg, & Wolfe, J. M. (2008). Time to guide: Evidence for delayed attentional guidance in contextual cueing. *Visual Cognition, 16*, 804–825. [PubMed] [Article]
- Macmillan, N. A., & Creelman, C. D. (1991). *Signal detection theory: A user's guide*. Cambridge: Cambridge University Press.
- Manns, J. R., & Squire, L. R. (2001). Perceptual learning, awareness, and the hippocampus. *Hippocampus, 11*, 776–782. [PubMed]
- Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics, 57*, 1–17. [PubMed]
- Nabeta, T., Ono, F., & Kawahara, J. (2003). Transfer of spatial context from visual to haptic search. *Perception, 32*, 1351–1358. [PubMed]
- Ogawa, H., & Kumada, T. (2008). The encoding process of nonconfigural information in contextual cuing. *Perception & Psychophysics, 70*, 329–336. [Article]
- Ogawa, H., Watanabe, K., & Yagi, A. (2009). Contextual cueing in multiple object tracking. *Visual Cognition, 17*, 1244–1258.
- Olson, I. R., Chun, M. M., & Allison, T. (2001). Contextual guidance of attention: Human intracranial event-related potential evidence for feedback modulation in anatomically early temporally late stages of visual processing. *Brain, 124*, 1417–1425. [PubMed]
- Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology, 21*, 469–514.
- Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics, 63*, 1239–1249. [PubMed] [Article]
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*, 333–367. [PubMed] [Article]
- R Development Core Team (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Santee, J. L., & Egeth, H. E. (1982). Do reaction time and accuracy measure the same aspects of letter recognition? *Journal of Experimental Psychology: Human Perception and Performance, 8*, 489–501. [PubMed]
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315.
- Töllner, T., Gramann, K., Müller, H. J., Kiss, M., & Eimer, M. (2008). Electrophysiological markers of visual dimension changes and response changes. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 531–542. [PubMed] [Article]

- Töllner, T., Zehetleitner, M., Gramann, K., & Müller, H. J. (in press). Top-down weighting of visual dimensions: Behavioral and electrophysiological evidence. *Vision Research*.
- Töllner, T., Zehetleitner, M., Krummenacher, J., & Müller, H. J. (in press). Perceptual basis of redundancy gains in visual pop-out search. *Journal of Cognitive Neuroscience*.
- Tseng, Y. C., & Li, C. S. R. (2004). Oculomotor correlates of context-guided learning in visual search. *Perception & Psychophysics*, *66*, 1363–1378. [[PubMed](#)]
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238.
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–73). Hove: Psychology Press.
- Zehetleitner, M., & Müller, H. J. (in preparation). A formal model of the dimension weighting account.