Attentional capture: An ameliorable side-effect of searching for salient targets

Heinrich R. Liesefeld\textsuperscript{a,b}, Anna M. Liesefeld\textsuperscript{b} and Hermann J. Müllerb

\textsuperscript{a}Department of Psychology, University of Bremen, Bremen, Germany; \textsuperscript{b}Department Psychologie, Ludwig-Maximilians-Universität, München, Germany

ABSTRACT
This commentary highlights that some of the remaining discrepancies in the attentional-capture debate can be resolved by a simple assumption: observers do not use the priority map when this map is useless to solve the task. Rather, whenever search targets are known to be non-salient, observers resort to a previously postulated alternative search strategy for which (distractor) saliency signals are irrelevant. Equipped with this assumption, we trace thus-far unaccounted-for discrepancies between empirical studies on attentional capture back to specific design choices that affect relative target saliency (display density and non-target heterogeneity).

According to most theoretical positions in the debate, a salient distractor is needed for attentional capture to occur. This is because only a salient distractor would cause strong activation on the attention-guiding “priority map,” thereby summoning focal attention to its location. Based on theoretical considerations and an in-depth review of the literature, we recently added a second premise: observers must actually employ the priority map, that is: they must allow attention to be guided by saliency signals (Liesefeld et al., 2019, 2020). Obviously, rational observers would cease utilizing saliency signals if, in their current search environment, they experience priority map-based guidance to hardly ever yield a successful outcome: efficient attentional selection of the target. Pursuing such a strategy would be as futile as trying to peel an orange with a hammer – the priority map is a great tool in many situations, but not when the target is non-salient. Consequently, we argued, attentional capture would not occur when the target is known to be non-salient. Equipped with this premise and its consequences, we will go on to explain existing discrepancies in the literature and attempt to reconcile the Theeuwes versus Folk/Remington positions from a different perspective to that advocated by Gaspelin and Luck. These efforts are guided by a simple question: do observers expect the target to be salient? In other words, is the second necessary condition for capture to occur fulfilled?

Saliency is determined by local feature contrast: the same feature can be salient or non-salient depending on whether or not surrounding objects (non-targets) deviate in the respective feature dimension (Nothdurft, 2000). The more an object deviates from its immediate surround, the more salient it is and the more guidance it can provide. This results in, for example, a continuous relationship between local feature contrast of the target and search times (Liesefeld et al., 2016; Töllner et al., 2011). Targets are non-salient in displays with only a few non-targets (sparse displays), which produce insufficient local feature contrast (Rangelov et al., 2017). In displays with variable non-targets (heterogeneous displays), it is difficult to experimentally control target saliency (in particular, when the target is defined in the problematic shape dimension; Wolfe & Horowitz, 2017) and targets do typically not achieve a similarly high local feature contrast as in homogeneous displays. Furthermore, heterogeneous non-targets yield local feature contrast among each other and these salient non-targets therefore, generate peaks on the priority map that would...
compete with the target peak for attention allocations. Thus, display heterogeneity decreases the relative saliency of targets (see Constant & Liesefeld, 2021; Duncan & Humphreys, 1989). Accordingly, relying on saliency to find targets in highly heterogeneous or sparse displays is not expedient, because it is unlikely that the priority map will guide attention to a target of low (relative) saliency.

We have recently argued that observers possess an alternative mechanism to find non-salient targets that does not rely on saliency (Liesefeld & Müller, 2020): wilful deployment of spatial attention to one clump of search items after the other (clump scanning). This corresponds to the subjective experience of serial search, though processing within each clump is assumed to actually operate in parallel. Rather than letting attention be reflexively guided by the priority map, for reasons discussed in Liesefeld and Müller (2020), this wilful mechanism operates in a spatially systematic, though idiosyncratic manner: observers might deliberately search from left to right, or in a clockwise direction, for example. Processing within a clump involves a parallel matching to a target template and matching efficiency depends on featural similarity to a target (and possibly distractor) template, but not on local feature contrast (saliency) or overall priority. Indeed, others have argued that the priority map is not a stage of the visual stream, but “a control module sitting to one side of the main selective pathway” (Wolfe, 2007). Given this, it appears plausible that this module is bypassed if it provides no benefit. Thus, while distraction in searches for non-salient targets is an interesting research topic in its own right, it is – on this theoretical background – uninformative with regard to priority map mechanisms and so only of indirect relevance to the attentional-capture debate (see also Theeuwes’ position in Luck et al., 2021).

Which search mode (priority guidance vs. clump scanning) participants come to “tune into” in a given search scenario depends on which mode they experience as more efficient overall: displays with densely packed, homogenous non-targets, a salient target and many potential target locations must (after some experience with the task) induce priority guidance, but a closer look is necessary for many other displays employed in the literature. For example, non-target heterogeneity generally makes non-targets more salient and, as a consequence, reduces the relative saliency of the target, as discussed above. However, if non-targets are repeated within a display (which is often an accidental consequence of increasing set size, because the same limited set of non-target shapes is used; Theeuwes, 2004; Wang & Theeuwes, 2020), average target saliency increases and average non-target saliency decreases – so that, with sufficiently large displays, priority guidance gains the edge over clump scanning. This provides an attractive alternative explanation for a recent study by Wang and Theeuwes (2020). They found that the distractor used by Gaspelin et al. (2015) was suppressed in sparse displays but (sometimes) captured attention in dense displays. Wang and Theeuwes interpreted this as evidence that Gaspelin et al.’s finding of active distractor suppression is attributable to the fact that target and distractor were non-salient because of their use of sparse displays (see Theeuwes’ position in Luck et al., 2021, for details). While agreeing with most of Theeuwes’ argument, we disagree that the distractor in Gaspelin et al. was non-salient; in fact, it clearly was the most salient object in the respective displays, because it stood out in its defining (colour) dimension. From our stance, the most important consequence of increasing set size was an increase in relative target saliency due to the repetition of (a limited set of) non-target shapes, as explained above. Notably, Wang and Theeuwes manipulated set size between participants, so that the high-set-size group could rely on saliency to find the target and therefore likely employed priority guidance. By contrast, the low-set-size group knew that the target would be non-salient and so employed clump scanning. What Wang and Theeuwes demonstrated then is that salient distractors can interfere in priority guidance mode, but are more often suppressed (or rather initially ignored; see Kerzel & Burra, 2020) in clump scanning mode. Importantly, Wang and Theeuwes’ findings empirically confirm that, with the low-set-size conditions implemented by Gaspelin and Luck, the latter’s participants would not have used priority guidance either.

Differentiating tasks that induce clump scanning versus priority guidance helps to resolve many discrepancies in the literature. To give a further example: Gaspelin and Luck (2018) demonstrated that the specific distractor colour needs to be known in advance for attentional suppression to occur and that an unpredictable distractor would capture...
(overt) attention. Relatedly, changing the distractor colour after each block, Vatterott and Vecera (2012) found that the distractor initially produced strong interference and only later on, within a given block, ceased to interfere. Based on these findings, Won et al. (2019) had expected to find that varying the colour of their distractor would render it difficult to ignore. However, to their surprise and in direct contrast to Gaspelin and Luck, distractor interference did not differ between their mixed- and fixed-distractor-feature conditions, indicating that the specific distractor colour does not need to be known (or experienced) in advance. On the theoretical background provided by the present commentary, it is easy to resolve this discrepancy by noting that Gaspelin and Luck as well as Vatterott and Vecera employed heterogeneous non-targets, whereas the non-targets were homogeneous in Won et al. Thus, priority guidance was induced only under the conditions of Won et al. They found that observers can dimensionally down-weight colour saliency signals during the search for a shape target (second-order suppression), confirming our Dimension Weighting Account (DWA; Liesefeld & Müller, 2019). By contrast, distractor handling during the clump scanning induced in Gaspelin and Luck and Vatterott and Vecera appears to depend on the match to a distractor template (first-order suppression), which can apparently be generalized as a result of experience (Vatterott et al., 2018).

Despite many similarities, our notion of priority guidance differs from Theeuwes’ position in a further crucial respect that renders it compatible with the position of Folk and Remington (Luck et al., 2021): Even if conditions are “ideal” for capture to occur (salient target together with a more salient distractor), we believe that capture can be avoided: in particular, saliency signals can be modulated, so that even if attention is guided by the priority map, salient distractors do not necessarily capture attention (Liesefeld et al., 2019). Technically, distractor saliency and searching for saliency signals are the two necessary, though not sufficient, conditions for capture to occur. As detailed in Liesefeld and Müller (2020), the notion of “dimensional weighting” sets priority guidance also apart from the singleton-detection mode introduced by Bacon and Egeth (1994) and explains a highly reliable finding incompatible with either feature-search or singleton-detection mode (as well as most other theoretical stances in the attentional-capture debate, including Theeuwes’): unpredictably intermixed distractors of the same saliency produce large or, respectively, little interference depending on the distractors’ dimensional relationship to the current search target (Liesefeld & Müller, 2019; Liesefeld & Müller, 2021). Importantly, dimensional modulation of saliency signals is never perfect. In a computational model (Liesefeld & Müller, 2021), we have implemented dimension weighting by multiplying saliency signals from the target dimension by a certain weight parameter. In this model, saliency signals from non-target dimensions (including the distractor dimension) would become irrelevant only if this target dimension weight parameter became infinite (which is never the case). Thus, residual interference caused by different-dimension distractors during priority guidance is not a sign of pure bottom-up processing (cf. Theeuwes’ position in Luck et al., 2021), but simply a sign of imperfect (graded) top-down dimensional weighting. How strong this weighting typically is in studies claiming evidence for lack of top-down control, becomes apparent only in comparison with scenarios in which weighting is impossible and interference is dramatically higher because the distractor is salient in the same dimension as the target (Liesefeld et al., 2017, 2019). The graded nature and dimensional constraints of saliency weighting constitute strong theoretical specifications of top-down control during priority guidance; dimensional-weight settings as well as the general search mode (priority guidance vs. clump scanning) can be considered (high-level) aspects of the control state, as envisioned by Folk and Remington (see their position in Luck et al., 2021).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the German Research Foundation (DFG) under grant MU 773/16-2 (to H.J.M.) and grant LI 2868/3-1 (to H.R.L.).

References

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. Perception & Psychophysics, 55(5), 485–496. https://doi.org/10.3758/BF03205306
Constant, M., & Liesefeld, H. R. (2021). Massive effects of saliency on information processing in visual working memory. *Psychological Science*, 32(5), 682–691. https://doi.org/10.1177/0956797620975785

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

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychological Science*, 26(11), 1740–1750. https://doi.org/10.1177/0956797615597913

Liesefeld, H. R., & Müller, H. J. (2019). Distractor handling via Constant, M., & Liesefeld, H. R. (2021). Massive e

Liesefeld, H. R., & Müller, H. J. (2020). A theoretical attempt to lies between two hierarchical levels of priority computation revealed by spatial statistical distractor learning. *Journal of Experimental Psychology. General*, 150(4), 710–728. https://doi.org/10.1037/xge0000970

Liesefeld, H. R., Liesefeld, A. M., & Müller, H. J. (2019). Distractor-interference reduction is dimensionally constrained. *Visual Cognition*, 27(3–4), 247–259. https://doi.org/10.1080/13506285.2018.1561568

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(6), 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

Liesefeld, H. R., & Müller, H. J. (2020). A theoretical attempt to revive the serial/parallel-search dichotomy. *Attention, Perception & Psychophysics*, 82(1), 228–245. https://doi.org/10.3758/s13414-019-01819-z

Liesefeld, H. R., & Müller, H. J. (2021). Modulations of saliency signals at two hierarchical levels of priority computation revealed by spatial statistical distractor learning. *Journal of Experimental Psychology. General*, 150(4), 710–728. https://doi.org/10.1037/xge0000970

Luck, S. J., Gaspeolin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress toward resolving the attentional capture debate. *Visual Cognition*, 29(1), 1–21. https://doi.org/10.1080/13506285.2020.1848994

Nothdurft, H. C. (2000). Salience from feature contrast: Variations with texture density. *Vision Research*, 40(23), 3181–3200. https://doi.org/10.1016/s0042-6989(00)00168-1

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(5), 651–671. https://doi.org/10.1037/xge0000284

Theeuwes, J. (2004). Top-down search strategies cannot over-ride attentional capture. *Psychonomic Bulletin & Review*, 11(1), 65–70. https://doi.org/10.3758/bf03206462

Töllner, T., Zehetleitner, M., Gramann, K., & Müller, H. J. (2011). Stimulus saliency modulates pre-attentive processing speed in human visual cortex. *PloS One*, 6(1), e16276. https://doi.org/10.1371/journal.pone.0016276

Vatterott, D. B., Mozer, M. C., & Vecera, S. P. (2018). Rejecting salient distractors: Generalization from experience. *Attention, Perception, & Psychophysics*, 80(2), 485–499. https://doi.org/10.3758/s13423-012-0280-4

Wang, B., & Theeuwes, J. (2020). Salience determines attentional orienting in visual selection. *Journal of Experimental Psychology. Human Perception and Performance*, 46(10), 1051–1057. https://doi.org/10.1037/xhp0000796

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). Oxford University Press.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, 1(3), 1–8. https://doi.org/10.1038/s41562-017-0058

Won, B.-Y., Kosoyan, M., & Geng, J. J. (2019). Evidence for second-order singleton suppression based on probabilistic expectations. *Journal of Experimental Psychology. Human Perception and Performance*, 45(1), 125–138. https://doi.org/10.1037/xhp0000594