Dear

|  |  |
| --- | --- |
| Editor comments |  |
| I take small issue with any conceptualization of human behavior being "random".  There is a big difference between "not guided by repeated configuration" (i.e. line 472), "random" and more likely: "being guided by some, unknown (and ineffective) to the experimenter, strategy adopted by the participant that could even vary from trial to trial. |  |
|  |  |
| Reviewer 1 comments |  |
| 1.      It is easy to see how the results are consistent with 'late' accounts. These propose that search is effectively random until the target is located, as which point contextual cuing reflects some response facilitation process. The results are also consistent with 'not-quite-so-late' accounts in which search is initially random, but once a familiar arrangement of distractors is identified near the target, search is guided towards the target. But what is less clear is what the accounts are that are inconsistent with the results? The main candidate appears to be a claim by Seitz et al., but this is described very briefly and no discussion or review is provided of the evidence on which Seitz et al. reached their conclusion. What's missing is a convincing argument against an early view which proposes that the initial processing of distractors far from the target facilitates eye movements towards the target, with this effect snowballing as search continues. This early facilitation may be overwhelmed by later facilitation and hence rather hard to detect, but ruling it out completely (as the authors' preferred account does) seems extreme. |  |
| 2.      There's an unacknowledged aspect of the procedure that I feel needs to be explicitly discussed.  The authors say (503-9) "The first of these is an inefficient search process where search fails to move towards the target in trials with more fixations. This is followed by a phase in which monotonic, positive increments are made toward the target position in the final 3 to 4 fixations... the effect of the endogenous central cue in the current study is to eliminate, or considerably reduce, the engagement with this first phase of the search process." The assumption is that by providing a valid cue to half of the display, a big chunk of the first stage is eliminated at minimal cost. But what is missing here is recognition that this will only be true on average. On any given trial, say with the cue pointing to the right, the participant might (if the cue had been absent) have initially search the right-hand side of the screen anyway. Or might have searched the top or the bottom. It is almost as if the authors assume that the part of the display that is discounted by the arrow (namely, the half to which it is not pointing) is where the participant would've looked. But this obviously can't be true. The authors need to incorporate this more realistic search behaviour in their discussion. The worry is that it is theoretically consequential. For example, on some proportion of trials the participant would anyway have initially searched on the side to which the arrow points, in which case the observed contextual cuing effect might in part be attributable to early search processes (contrary to the overall conclusion). |  |
|  |  |
| Reviewer 2 comments |  |
| my main issue with this work is that it is open to alternative interpretations. Specifically, the same set of (four) target positions was used for repeated and non-repeated displays. A given context was thus much less predictive for a given, individual, target position, which may have trained other than target-distractor mnemonic associations (cf. Geyer et al., 2024; PBR), such as where to expect critical search-for target positions in the display (i.e., probability cueing; Geng & Behrmann, 2005; AP&P) or learning of distractor-distractor associations (Beesley et al., 2015, JEP:LMC). Under this account, a contextual-facilitation effect wouldn't be a 'classical' contextual cueing effect (as described by Chun & Jiang, 1998; Cog Psy): instead, RT facilitation may be brought about by old distractor layouts serving as a better, i.e., more reliable, retrieval cue for repeated, i.e., learned, target positions. Applied to the present findings, this could explain the absence of a contextual facilitation effect in inconsistent trials of Experiment 1, that is, when invalid spatial cues misguide participants' attention to incorrect display regions and thus making RTs comparable between old- and new-context trials (as these two types of trials used the very same target positions). But this alternative account could also explain the presence of a contextual facilitation effect in consistent conditions (Experiment 1-3) by the joint, i.e., additive effects of absolute target position memories and distractor-distractor memories (the latter would provide the more effective retrieval cue for individual target positions). The far-reaching consequence would be that the present investigation tells us little about 'classical' contextual cueing. Instead, the result may be taken as evidence that the retrieval of other forms of contextual cueing, i.e., distractor-distractor learning, is time-dependent (and limited to the display quadrant containing the search-for target item). |  |
| As a minor comment: the authors interpret their results as support for a late, i.e., response-based, account of contextual cueing. But reaction times are typically indiscriminate as to the exact process that receives support by a specific (old/ new) display manipulation. And combining the contextual-cueing task with process-sensitive ERPs, it was found that contextual repetitions improve particularly late, albeit perceptual, stages of target-orientation discrimination decisions (e.g., Chen et al., 2022; Psychophysiology). Because of this, the authors may want to advocate a more nuanced view of the perceptual vs. response factors that may underlie contextual cueing. |  |
|  |  |
| Reviewer 3 comments |  |
| 1 - My biggest question from the current work concerns what happens during the early ineffective part of search. One view of what is happening here is that a bunch of display elements are being encoded as retrieval cues to connect the current display with potentially helpful representations in memory. The faster transition from ineffective to effective search for repeated displays reflects effective retrieval and the associated CC benefits coming online. What is especially interesting about the current results is that even a restricted set of display elements appears to be just as effective as the full set of display elements (e.g., similar cuing across arrow conditions in Experiment 2), so retrieval appears to remain efficient even with half the complement of retrieval cues (for the relevant search displays). The dampening of the CC benefit in Experiment 3 for the repeated-proximal displays suggests that there are some (small) costs of disrupting a proportion of the previously useful retrieval cues (i.e., randomizing distal distractor locations). To my mind, the work presented here does a bit more than speak against the monolithic search strategy advocated for by Seitz et al. (2023), and can further inform the quite complex combination of processes that appear to support CC. I think expanding the discussion along these lines could improve the impact of the manuscript, but I leave that determination to the discretion of the authors. |  |
| 2 - In Experiment 2, much of the analysis included data from all epochs (or pre- vs. post- removal of arrow cues). My intuition about what might happen when the cues were removed in the arrow condition would be that there would be a slight hit to performance because potentially, distractors on the opposite side of the display to the target would be encoded more weakly in memory, potentially reducing the effectiveness of retrieval once the arrow was removed. Based on the visualization of the data, I wouldn't expect a difference between arrow and no-arrow conditions in the early epochs of phase 2, but a more focused pairwise comparison of these conditions in Epoch 6 could help speak to how relevant the very distal distractors (i.e., those on the irrelevant side of the search display) were to memory retrieval and CC. If there is support for a difference, this might connect to the difference between repeated-proximal vs. fully repeated displays seen in Experiment 3. |  |
| 3 - On outlier exclusion, removing people based on the timeout criterion and then mean RT (even if this was the same person affected) seems to be double-dipping on criteria. Since the mean is pulled by timeouts, does the exclusion rate change if timeouts are filtered out prior to applying the mean RT criterion (or alternatively, if something less sensitive to trial-level outliers like median RT is used instead)? |  |
| 4 - Minor typo in Abstract (Line 18): "...presence of the endogenous [cue] did not impede..." |  |