19th July 2024

Dear Dr Shomstein,

We are very grateful for the reviews of our submission PP-ORIG-24-110, and for an opportunity to address these points in a revision. We have detailed below all of the points of feedback made in the reviews, as well as our direct response to each point, including any substantial text added to the manuscript to address these points. We also made a number of small phrasing edits to the manuscript.

We look forward to receiving your response to this revision.

Yours sincerely,

Tom Beesley – on behalf of all authors

Editor

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.

We agree that "ineffective" is a better description of this process. We have amended all places in which we referred to this phase of search as "random".

Reviewer 1

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.

We have now substantially expanded the discussion of the Seitz et al. study in the introduction:

"In contrast to the localised facilitation account, it's possible that contextual cuing involves learning of a procedural template that guides eye-movements in a consistent manner following experience with the task. A recent study by Seitz et al. has provided evidence to support this claim. Participants eye-movements were monitored for repeated and randomly arranged configurations, and similarity metrics were computed to identify the consistency of scan-paths over time. Several findings point towards the establishment of a more general type of procedural learning in CC. Firstly, it was found that scan-path similarity increased over the course of training, but that the similarity of scan-paths was higher in repeated compared to random

configurations. Secondly, scan-path similarity was higher in the initial half of the search trial compared to the second half. These data suggest that, in contrast to the earlier characterisations of the initial search process as "inefficient", this early phase may be an important part of the behavioural response in CC, potentially involving the development of a generic scanning behaviour.

Importantly, Seitz et al. suggest that CC is best characterised as involving the acquisition of this generic procedural scanning response, and a configuration-specific facilitation. These behaviours occur in the early and late period of oculomotor guidance, respectively. The question remains as to how critical the early activation of procedural knowledge is to the development of CC. The current article examines this by significantly interrupting the search process with an endogenous cue for attention. In all experiments participants complete a contextual cuing visual search task but are also presented with an arrow that signals the side of the screen on which the target will appear. Thus, this cue disrupts the natural search process, considerably reducing the operation of the generic scanning response in the early phase of search. The experiments therefore examine whether this initial part of the search process is inconsequential for the observation of the CC effect, or whether the development and maintenance of the generic scanning behaviour contributes substantially to the CC effect."

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).

The reviewer makes the valid point that the initial search process is variable and that on some trials (on average half) the initial movement of gaze would be towards the side of the screen with the target. While search would likely terminate at an earlier point on these trials, it is not necessarily the case that search would proceed in the same manner as it does on cued trials. This is because on non-cued trials the participant has not ruled out the target being on the opposite side of the screen, while on cued trials it is always the case that the target would not appear on this opposite side of the screen. In any case, we do not dispute that uncued search has variable consequences, and therefore on some trials where initial movements are towards the target, search times would be shorter. The point is that it is this final part of the search that is crucial. One might call this "early" on a particular trial, but what is more relevant here is the

characteristic of this behaviour as “effective” search, and its relationship to an initial period of ineffective search. We originally noted in our General Discussion: “Taken together, the results here point towards the possibility of three components to the behaviour in CC: an early ineffective search, followed by enhanced localisation and increased perceptual discrimination of the target, driven by the distractors closest to the target.”

To hopefully clarify our position, we’ve added to the discussion of Experiment 1:

“It should be acknowledged that the variable search behaviour that participants would exhibit during the early part of the search process would naturally lead them to search the area around the target on many trials. As such, on trials without the endogenous cue, the termination of the inefficient phase of search will occur earlier on some trials compared to others. The cuing of attention by the valid arrow cue ensures this termination happens on every cued trial, eliminating the inefficient phase of search.”

And to the General Discussion:

“Search behaviour under CC conditions is necessarily variable, however, and each time a configuration is encountered, the pattern of eye-movements will inevitably be driven by a range of factors that lead to variation in the scan paths taken. What is clear is that it is the final few fixations and saccades that are crucial to the search behaviour that facilitates CC, and this crucial period will follow a variable length of ineffective search.”

Reviewer 2

1. 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).

Response: We thank the reviewer for this comment. As we understand it, the reviewer is suggesting that instead of distractor-target associations, the CC effect in our tasks may result from probability cuing and distractor-distractor learning.

Regarding probability cuing, it's hard to see how this would lead to a CC effect in this task: target positions were as common in repeated configurations as random configurations (an arrangement that is held in almost all CC studies). RT facilitation could not be due to some target locations being more frequent than others.

The other mechanism suggested is distractor-distractor learning. The cited paper by Geyer et al. (2024) is very interesting in terms of understanding how the overlap of target positions can affect the flexibility of the learned representations in contextual cuing. In that paper, the authors suggest that shared targets may lead the learning in CC to shift from distractor-target associations towards distractor-distractor associations. It is this possible mechanism that the reviewer notes could be playing a role in the current experiments. We have two points which we hope will provide an adequate response to the suggestion.

1. The shift from distractor-target to distractor-distractor learning is a proposed mechanism for the relearning effect observed in Geyer et al. (2024) but it certainly has not been proven that this is what is driving greater flexibility in the representations. That is, to our knowledge there has not been a demonstration that a procedure in which targets are shared amongst repeated and random configurations leads to these types of representations. Therefore, in the case of the current procedure, common targets across repeated and random configurations may change how flexible the representations are in CC, but we cannot be sure that distractor-target learning ("classical CC") is not playing a dominant role here.

2. We have conducted further analysis that compares the effects observed in the current study with known effect sizes for CC driven only by (what we might suppose are) distractor-distractor associations. Vadillo et al. (2021) performed a meta-analysis of the CC effect that emerges when distractor configurations are not predictive of a particular target position. They found that the resulting CC effect size is d = 0.60, 95% CI [0.42, 0.78]. We calculated Cohen's d for the initial phase of the three experiments in the current paper, which produced an effect size of d = 0.80, 95% CI [0.40, 1.19]. There is overlap here in CIs (our Cis being much greater due to being from 3 studies, compared to a much larger set of studies in Vadillo et al.'s meta analysis), but it is worth noting that our effect is towards the upper bound of Vadillo et al.'s CI. It is also worth noting that our current training procedures in phase 1 were shorter (20 presentations) than those used in the studies detailed by Vadillo et al. (2021) (30 presentations). We also computed the effect size in Phase 2 of Experiment 2 in the current study, when CC is well established and the endogenous cue has been removed, and this produced an effect size of d = 1.25, 95% CI [.78, 1.70]. This is clearly more in line with the reported effect size of typical CC studies (d = 1; Vadillo et al., 2016) than those with targets that are non-predictive (d = .60; Vadillo et al., 2021). We are therefore confident that the CC effect we are observing derives primarily from distractor-target associations ("classical CC") rather than distractor-distractor associations.

2. 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.

We thank the reviewer for this suggested amendment. In the General Discussion we now clarify this point and cite this paper.

Reviewer 3

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.

We thank the reviewer for this comment. We've added the following paragraph to the General Discussion:

"Our data are consistent with previous theoretical (Brady & Chun, 2007) and empirical (Olson and Chun, 2002) work that has highlighted the influence of distractor configurations localised to the target. Experiment 2 in particular demonstrates that acquisition of effective representations is equivalent if search is limited to one half of the display from the outset. In Experiment 3 the CC effect was observed only when fully repeated and proximal-repeated configurations were presented. Interestingly the CC effect was substantially weaker in the case of configurations with only proximal-repeated distractors. This must reflect a generalisation decrement between the stored representation and the available cues for the target. Our manipulation of the influence of repeated distractors was based on disrupting the repeating configurations on a quadrant basis: those inside the quadrant retained their positions, while those outside were randomised. This somewhat crude manipulation will not perfectly capture the impact of all distractors: it is likely that the influence of distractors at increasing distances from the target will have a gradually reducing influence on driving a CC effect."

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.

The reviewer makes an interesting point about the influence of the unprocessed repeated distractors once the arrow is removed in epoch 6. However, as they suggest, this isn't borne out in the data. If these unprocessed distractors led to the suggested reduction in effective retrieval, then we'd see weaker contextual cuing in the arrow condition compared to the no-arrow condition. In fact, numerically the opposite is true, with the arrow condition showing a larger difference between repeated and random configurations. We conducted the phase 2 Bayesian analysis on just the data from epoch 6, which found that the best fitting model contained only the factor of trial type (BF = 38). The addition of the interaction term to the model provided a weaker fit to the data compared to the model without this term, though this fell just short of substantial evidence for the absence of an interaction effect, BF = 0.36. Given the observed numerical trend in the presented figure is in the opposite direction to that predicted by this account, we have decided not to include this analysis in the manuscript.

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)?

Timeouts were indeed removed before any RT analysis was conducted. We have added this detail in the exclusion criteria for mean RT.

4 - Minor typo in Abstract (Line 18): "...presence of the endogenous [cue] did not impede..."

Thank you. Fixed.