Contextual cueing improves attentional guidance, even when guidance is supposedly optimal

Anthony M. Harris[1,2] & Roger W. Remington[2,3]

[1]Queensland Brain Institute, The University of Queensland

[2]School of Psychology, The University of Queensland

[3]Department of Psychology, University of Minnesota

Running head: Contextual cueing improves attentional guidance

Address for correspondence:

Anthony M. Harris

Queensland Brain Institute

The University of Queensland

St Lucia, QLD 4072

Australia

Email: anthmharris@gmail.com

Abstract

Visual search through previously encountered contexts typically produces reduced reaction times compared to search through novel contexts. This *contextual cueing* benefit is well established, but there is debate regarding its underlying mechanisms. Eye-tracking studies have consistently shown reduced number of fixations with repetition, supporting improvements in attentional guidance as the source of contextual cueing. However, contextual cueing benefits have been shown in conditions in which attentional guidance should already be optimal – namely, when attention is captured to the target location by an abrupt onset, or under pop-out conditions. These results have been used to argue for a response-related account of contextual cueing. Here, we combine eye tracking with response time to examine the mechanisms behind contextual cueing in spatially-cued and pop-out conditions. Three experiments find consistent response time benefits with repetition, which appear to be driven almost entirely by a reduction in number of fixations, supporting improved attentional guidance as the mechanism behind contextual cueing. No differences were observed in the time between fixating the target and responding – our proxy for response related processes. Furthermore, the correlation between contextual cueing magnitude and the reduction in number of fixations on repeated contexts approaches 1. These results argue strongly that attentional guidance is facilitated by familiar search contexts, even when guidance is near-optimal.

Keywords: Contextual Cueing, Eye-tracking, Visual Search, Attentional Guidance, Pop-out, Spatial Cueing

Significance Statement

When searching for a target amidst distractors, search is facilitated if one has previously encountered the particular arrangement of items in the display. Is this benefit due to repeated displays facilitating search itself (*attentional guidance*), or is it due to repeated displays allowing one to recognize the target faster, once it has been found (*response related processes*)? Here we used eye-tracking to assess attentional guidance and the time required for response related processes, while participants performed search tasks that have previously been suggested to remove the need for attentional guidance. The results showed that in all cases, when search was facilitated by display repetitions, this was associated with improved attentional guidance, even in situations where attentional guidance had previously been assumed to be optimal. These results suggest a single mechanism, attentional guidance improvements, underlies search facilitation by repeated distractor contexts.

Contextual Cueing Improves Attentional Guidance, Even Under Optimally Guided Conditions

When searching for targets in cluttered visual displays, search times decrease over trials due to practice at the task. However, if the task includes a small number of search displays that are repeated many times, search times for these displays improve beyond what is expected due to practice alone (Chun & Jiang, 1998). This effect has been termed *contextual cueing*. Despite almost twenty years of work examining the contextual cueing effect, we still do not know what cognitive processes are facilitated to bring this effect about (see Goujon, Didierjean, & Thorpe, 2015, for review). Two competing accounts of the contextual cueing benefit that have received some support attribute this benefit either to improved attentional guidance (e.g., Chun & Jiang, 1998; Peterson & Kramer, 2001), or to improved response related processing (e.g., Kunar, Flusberg, Horowitz, & Wolfe, 2007; Schankin & Schubö, 2009, 2010). Although these accounts are not mutually exclusive they represent independent contributions to the contextual cueing effect. At present we know little about how each contributes to the magnitude of observed contextual cueing benefits. Below we examine the evidence for each of these accounts. We then present three experiments that use eye tracking to examine what stages of search are facilitated by the presence of repeated search arrays.

The Evidence From Behavioral Studies

Initial studies examining the source of the contextual cueing effect suggested the effect was caused by improved attentional guidance on repeated displays (Chun & Jiang, 1998). Attentional guidance refers to the efficiency with which a target is located during visual search (Wolfe & Horowitz, 2004), typically assessed by manipulating the number of items in a search display and computing the slope of the resultant search function. This slope is an index of the time required to search each item in the display, whereas the intercept of the search function is interpreted as the time required for processes other than active search, such as perceptual processes, and the selection, preparation, and execution of a response. Thus, a manipulation that leads to improved attentional guidance is expected to produce shallower search slopes, while a manipulation that improves processes involved in selecting and/or executing a response is expected to influence the intercept of the search function, but not change the function's slope. Early support for attentional guidance as the source of contextual cueing came from Chun and Jiang (1998), who had participants search for a target among repeated and novel displays with set sizes of 8, 12, or 16 items. They found that repeated displays produced shallower search slopes than novel displays, and thus concluded that contextual cueing improved attentional guidance (for supporting evidence see Kunar, Flusberg, & Wolfe, 2008; Tseng & Li, 2004; Zhao, et al., 2012). Conflicting evidence emerged, however, when Kunar, Flusberg, Horowitz, and Wolfe (2007) ran a series of ten contextual cueing studies where they manipulated set size and observed significant slope differences in only one of these, raising questions about the role of attentional guidance in contextual cueing.

Indeed, Kunar et al. (2007) reasoned that if attentional guidance were the driving force behind contextual cueing, no contextual cueing effect should be produced under conditions in which attentional guidance is optimal[1]. To test for contextual cueing under optimally-guided conditions, Kunar et al. (2007) used 'pop-out' search displays, in which the target was always red, and was presented with a number of green distractors. These conditions typically produce flat search slopes in which the time required to find the target is independent of the number of distractors (hence the claim that attentional guidance is optimal). Using this pop-out contextual cueing paradigm, Kunar et al. (2007) showed that search times were still significantly faster on repeated displays than on novel displays, even though search efficiency was supposedly optimal. They interpreted this as evidence that attentional guidance has at best a small contribution to contextual cueing. In a separate pop-out search experiment, the authors manipulated the congruence of the identity of a red target and the identity of green distractors and demonstrated that the contextual cueing benefit was only present when the target and distractors were congruent. When the distractors signaled a response that was incongruent with the actual response required on a trial, the contextual cueing benefit was abolished. Based on this the authors suggested that response related processes, most probably response selection, were in fact the source of the majority of the contextual cueing effect. In subsequent work, Kunar, Flusberg, and Wolfe (2008) showed that given sufficiently slow search times, attentional guidance did emerge in contextual cueing, as evidenced by shallower search slopes in the repeated condition. However, they concluded that the slope reduction was too small, and the onset of guidance too late in the search, to account for the majority of contextual cueing effects in the literature.

Further evidence against the role of attentional guidance in contextual cueing comes from a study by Schankin and Schubö (2010), who examined the combined effects of display

[1] Note that our use of the term 'optimal' in this paper is used to reflect the position put forward by previous authors, and does not reflect a strong claim on our part that these conditions produce optimal guidance

repetition and spatial cueing. They augmented a standard contextual cueing paradigm by presenting a small circle flashed briefly at the location of the target just prior to the search display, designed to capture attention to the target location (Yantis & Jonides, 1984). They reasoned that the presence of a salient cue should produce optimal guidance of attention to the target location on these trials, eliminating all but response related effects. The results showed that contextual cueing was still observed when targets were pre-cued by an abrupt onset, even though attention should already have been captured to the target location. Furthermore, contextual cueing was no larger on uncued trials than on cued trials, leading the authors to conclude that under typical uncued conditions contextual cueing was likely not due to the guidance of attention. It should be noted, however, that Schankin and Schubö (2010) used only four target locations that were all close to the initial fixation point. This produced fast searches, in line with the times expected for efficient 'pop-out' search, even on trials in which the targets did not have their location pre-cued. Thus, the equal magnitude of the contextual cueing effects produced by cued and uncued trials may have resulted from a ceiling effect on the possible contextual cueing magnitude under these conditions, rather than an absence of attentional guidance benefits. Indeed, the contextual cueing effects found by Schankin and Schubö (2010) were less than 20ms, compared to a typical effect of 80–100ms or more. Thus, the behavioral results of Schankin & Schubö (2010) may represent a special case, not representative of those produced under more standard search conditions.

Studies seeking to examine the mechanisms underlying contextual cueing using signal detection theory have also provided mixed results. Evidence has appeared supporting both the attentional guidance account (Geyer, Zehetleitner, & Müller, 2010) and the response related account (Schankin, Hagemann, & Schubö, 2011). In sum, whether behavioral contextual cueing results support improvements to attentional guidance or response related processes seems to depend on the particular paradigm and measure used. This may be a problem inherent to inferring cognitive processes from behavioral manipulations alone, as these methods tend to be rather indirect. Other studies have attempted to use more direct measures of processes related to attentional guidance and response related processing to get at these questions.

The Evidence From Eye Tracking

Improved search efficiency would predict fewer fixations required to find the target in repeated compared to novel displays. In contrast, improved response processing would not affect search processes, and, thus, not reduce the number of fixations. Instead, improved efficiency in response selection, programming, or execution, predicts faster responses once the target has been fixated. Eye-tracking studies of contextual cueing provide ample evidence that repeated displays reduce the number of fixations needed to find the target, evidence of improved attentional guidance (Brockmole & Henderson, 2006; Geringswald, Baumgartner, & Pollmann, 2012; Geringswald, Herbik, Hoffmann, & Pollmann, 2013; Geringswald & Pollmann, 2015; Manelis & Reder, 2012; Manginelli & Pollmann, 2009; Peterson & Kramer, 2001; Tseng & Li, 2004; Zang, Jia, Müller, & Shi, 2015; Zhou et al., 2012). Moreover, repeated displays have a greater proportion of trials on which the first fixation lands on the target (Peterson & Kramer, 2001). In contrast, eye-tracking studies have generally found little (Zhou et al., 2012) or no (Manelis & Reder, 2012; Tseng & Li, 2004) shortening of the time from target fixation to button press on repeated displays, suggesting the role of response related processes in producing the contextual cueing effect is minimal at best.

The Evidence From Cognitive Neuroscience

Directing spatial attention to an object on the left or right side of a visual scene has been associated with enhanced amplitude and/or reduced latency of the N2pc component of the evoked potential contralateral to the attended side (see Luck, 2011, for review). N2pc amplitude contralateral to targets have been found to be greater on repeated displays than on novel displays, evidence in favor of improved guidance (Johnson, Woodman, Braun, & Luck, 2007; Kasper, Grafton, Eckstein, & Giesbrecht, 2015; Olsen, Chun, & Allison, 2001; Schankin & Schubö, 2009, 2010; but see Schankin, Hagemann, & Schubö, 2011, for a counterexample).

ERP examinations of response related processes typically focus on a component called the Lateralized Readiness Potential (LRP). There are two forms of the LRP that are thought to reflect different response related processes. The stimulus-locked LRP component is typically associated with processes occurring before response preparation and execution (Mordkoff & Gianaros, 2000), and in the absence of latency differences in earlier ERP components is often taken as an index of ease of response selection. In contrast, the response-locked LRP is commonly taken as an index of the rate or ease of response preparation and execution.

The evidence for contextual cueing affecting LRPs is somewhat mixed. Schankin and Schubö (2009) found that reaction time differences between repeated and non-repeated displays correlated with differences in response-locked LRP onset for the two display types, with no effects observed for the stimulus-locked LRP. However, a subsequent study by the same authors (Schankin & Schubö, 2010) found the opposite result: no effect of contextual cueing on the response-locked LRP, and a non-significant trend towards an earlier stimulus-locked LRP. Thus, based on evidence from ERPs it is unclear whether contextual cueing is associated with response related processes, and, if it is, what specific process it is related to (response selection, preparation, or execution).

Functional-MRI studies of contextual cueing have consistently found repeated displays to be associated with a deactivation of medial temporal lobe structures, particularly the hippocampus (Geyer, Baumgartner, Müller, & Pollman, 2012; Giesbrecht, Sy, & Guerin, 2013; Goldfarb, Chun, & Phelps, 2016; Greene, Gross, Elsinger, & Rao, 2007; Kasper, Grafton, Eckstein, & Giesbrecht, 2015; Manelis & Reder, 2012). This result is not, in itself, support for either the attentional guidance or response related accounts of contextual cueing (although, studies have implicated the hippocampus in the guidance of eye movements by the contents of long-term memory; Hannula & Ranganath, 2009; Summerfield, Lepsien, Fitelman, Mesulam, & Nobre, 2006). However, a study employing combined fMRI and EEG found that the magnitude of signal change in the medial temporal lobe was correlated with both the magnitude of the behavioral effect, and N2pc magnitude (Kasper, Grafton, Eckstein, & Giesbrecht, 2015). Furthermore, display repetition has been associated with heightened activation in brain regions involved in attentional control, such as the inferior parietal lobe and the superior temporal gyrus, among others (Giesbrecht, Sy, & Guerin, 2013; Goldfarb, Chun, & Phelps, 2016; Greene, Gross, Elsinger, & Rao, 2007; Kasper, Grafton, Eckstein, & Giesbrecht, 2015; Manelis & Reder, 2012; Manginelli, Baumgartner, & Pollmann, 2013).

The Present Study

In summary, response time, gaze metrics, EEG, and neuroimaging measures have produced an inconsistent pattern of results, and it remains unresolved whether attentional guidance, response facilitation, or an as yet unspecified combination of the two, underlies the contextual cueing effect.

The present study takes a closer look at pop-out and spatial cueing search paradigms that have provided behavioral evidence against attentional guidance and in favor of response facilitation. Specifically, we examined the pattern of eye fixations to determine if reduced search times are associated with reductions in the number of fixations, the time from target fixation to response, or both. In Experiments 1, 2a, and 2b, we employed eye tracking with a spatially cued contextual cueing paradigm similar to that of Schankin and Schubö (2010). In

Experiment 3 we examined eye fixations in a pop-out contextual cueing task similar to that of Kunar et al. (2007), with gaze contingent targets to limit peripheral target processing and provide a strict control of response related processing time. If these paradigms produce contextual cueing effects that are not driven by attentional guidance, we would expect to see contextual cueing in the reaction time results of these experiments that is not reflected in a reduction in the number of fixations required to find the target on repeated displays. Rather, if response related processes are the driving force behind the contextual cueing effect, the reaction time benefit for repeated displays should be mirrored by a corresponding reduction in the time between fixating the target and emitting a response.

Experiment 1

In Experiment 1 we sought to replicate the effect that repeated displays produce faster responses, even when the target location is pre-cued by a salient onset cue that should capture attention (Schankin & Schubö, 2010). To ensure that the contextual cueing we are looking at reflects the typically observed contextual cueing effect, and is not interfered with in some way by the presence of the spatial cues, we first established the contextual learning across four uncued epochs of search before introducing the spatial cues across the final four epochs. Schankin & Schubö (2010) used only four fixed target locations, a departure from the majority of contextual cueing studies. It is possible that the small number of locations leads to learning that is not representative of more standard paradigms. We allowed the targets to be located randomly in the displays (see below for details), as is typical in contextual cueing experiments. If contextual cueing is driven by processes other than attentional guidance, we should observe faster responses to repeated displays than to novel displays, even when the target location is pre-cued by a spatial cue (*validly cued trials*), and this effect should not be reflected in a difference in the number of fixations required to find the target. Finally, if facilitation of response related processes contributes to the contextual cueing effect, facilitation of responses should be reflected in a shortening of the time between fixating the target and emitting a response.

Methods

Participants

Sixteen participants took part in the current experiment (12 females, M = 22.38 years, SD = 2.55 years). This sample size was selected *a priori* as typical for contextual cueing experiments. Participation was voluntary and participants were compensated for their time at a rate of $10 per hour. All reported normal vision – participants with glasses or contact lenses were excluded as these can interfere with the eye tracking. This study was approved by the University of Queensland Human Research Ethics Committee.

Apparatus and Stimuli

Stimuli were displayed on a 19-inch CRT color monitor with a resolution of 1024x768 pixels and a refresh rate of 85 Hz, controlled by a computer running Windows XP. Eye movements were tracked using a video-based infrared eye-tracking system (Eyelink 1000, SR Research, Ontario, Canada) with a spatial resolution of 0.01° of visual angle and a sampling rate of 500 Hz. Participants had their head supported by the eye tracker's chin rest and forehead support and viewed the screen from a distance of 60 cm. For registration of manual responses, a standard USB keyboard was used. Event scheduling and response time measurement were controlled by Matlab, using the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007).

All stimuli were white (RGB: 255, 255, 255), and appeared on a grey background (RGB: 160, 160, 160). Participants first saw a fixation screen, consisting of a small fixation cross (0.4° x 0.4°), located centrally. On cued trials, this was followed by the appearance of a ring (the cue; 1.5° diameter), either at the location of the upcoming target (*valid trials*) or at the location of an upcoming distractor (*invalid trials*). Invalid cues always cued an item in a different quadrant of the display to the target. On uncued trials the cue period was replaced with an empty grey screen. The cue period was followed by an inter-stimulus interval consisting of an empty grey screen. Finally, the search display was presented (**Figure 1**).

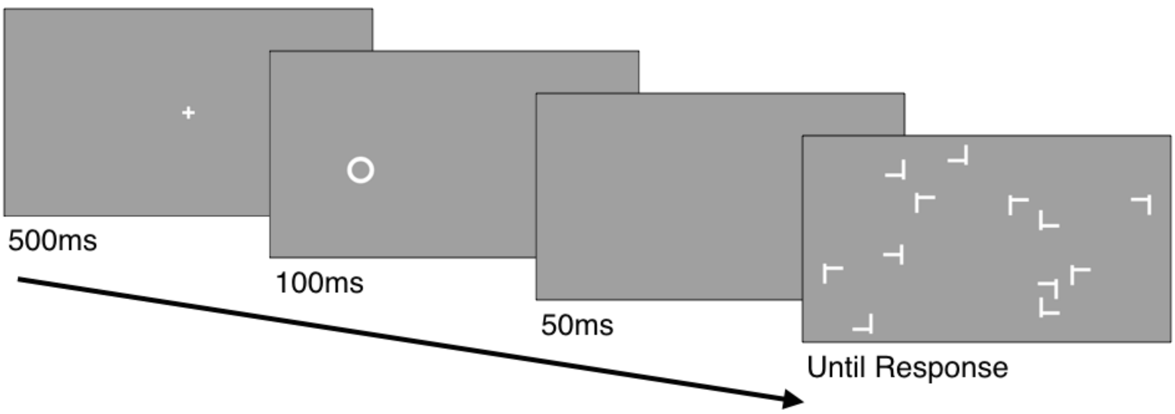

Figure 1: The spatially cued contextual cueing paradigm. Participants search for the sideways T among Ls. Prior to onset of the search display, participants may be presented with an onset cue either at the location of the target (*validly cued trials*), or at the location of a distractor item (*invalidly cued trials*). On *uncued trials* the cue display was replaced with a blank screen.

The search display consisted of 12 white stimuli (1.5° x 1.5°; **Figure 1**). Eleven of these were L shapes, with the horizontal bar offset slightly (1/6th of the length of the vertical bar) from the end of the vertical bar. These were rotated either 0°, 90°, 180°, or 270°. The target was a T shape of the same dimensions as the distractor Ls, rotated either 90° or 270°. Stimuli were randomly placed into the cells of an invisible 11x11 grid such that three stimuli appeared in each quadrant of the search display, with the constraint that stimuli could not appear within the central nine cells, or in any cells of the central column or row of the matrix. Furthermore, targets (but not distractors) were prevented from appearing in the three cells in each corner of the display. Targets appeared equally often in each quadrant of the search display in both repeated and novel conditions. Stimuli were randomly jittered up to 0.5° both vertically and horizontally to reduce collinearity and grouping among the stimuli. Throughout the task, incorrect responses elicited a 350ms, 1000Hz tone.

Procedure and Design

Participants were calibrated with the eye-tracker's standard 13-point calibration. They were then instructed that they were to search for the sideways T, pressing the 'z' key with their left index finger if the stem of the T pointed leftward, and pressing the '/' key with their right index finger if the stem of the T pointed rightward. They were asked to search as quickly and accurately as possible. Each trial began with a fixation control, such that the trial only began once the participant had been fixating within 1° of the central fixation cross for 500ms within a two second window. If two seconds passed without the participant having fixated for 500ms, the participant was calibrated anew and the trial was begun again with the fixation control. After this fixation control the cue display (or a blank screen on uncued trials) was presented for 100ms, followed by the inter-stimulus interval for 50ms. Finally, the search display was presented until the participant elicited a response. After this response the next trial began again with the fixation control.

Participants completed 16 practice trials of the task, followed by 960 trials of the experiment proper. These trials were divided into 60 blocks of 16 trials each. Each block was made up of eight *repeated displays* and eight *novel displays*. The eight repeated displays were generated at the start of the experiment, controlled such that there were two repeated displays with targets in each of the four quadrants of the search display. Each time a repeated display was presented it was identical to all other presentations of that same display, except for the orientation of the target, which was randomly determined on each trial. Target location, distractor locations, and distractor orientations were maintained across repetitions of the same display. The eight novel displays had their target locations determined at the start of the experiment such that there were two novel displays with targets in each of the four search quadrants, and these target locations were fixed throughout the experiment. All other aspects of novel displays were randomly determined each time a novel display was presented (target orientation and distractor locations and orientations). The 16 target locations used throughout the experiment were controlled to be unique, so that no two displays within a block shared a target location.

Participants first completed 20 blocks of the search task with no cues. This was to ensure that the presence of the spatial cues could not change the mechanisms involved in the learning of the repeated displays. At the end of the first 20 blocks, participants were informed of the upcoming spatial cues and were instructed to attend to them and use them to improve their search. Cues were presented at the target location on 50% of cued trials, and at a distractor location on 50% of cued trials, and participants were informed of this. The 40 cued blocks were divided into 20 two-block pairs. In each pair, each of the displays (repeated displays and novel displays with consistent target locations) was associated with one valid cue and one invalid cue. The order of these was shuffled so that half the displays received a valid cue in the first block of a pair and an invalid cue in the second block, and for the other half of displays this order was reversed. For our analysis each of these block pairs was considered to be one block containing one valid and one invalid cue for each display. The resulting 20 uncued blocks and 20 cued blocks were further collapsed to produce five-block *epochs* (four epochs uncued, four epochs cued) to increase statistical power. A self-paced break was presented at the end of every five uncued blocks, and every five cued block pairs.

Eye Tracking Measures

We analyzed two eye-tracking measures: the total number of fixations required to locate the target, our proxy for attentional guidance (Peterson & Kramer, 2001), and the duration between when the eyes first land on a target, and when that target is responded to, our proxy for the duration of response related processes (Tseng & Li, 2004). In these analyses only fixations longer than 50ms were included, and successive fixations on the same search item were treated as a single fixation.

Results

Analyses for the current experiments were performed using a mixture of R (R Core Team, 2015) and JASP (JASP Team, 2016). Greenhouse Geisser corrected *p*-values are reported where appropriate. All *t*-tests report two-tailed significance. As we have a strong directional hypothesis that contextual cueing should produce a benefit in repeated displays relative to novel displays, it could be argued that use of one-tailed tests is appropriate. We note and discuss cases where use of a one-tailed test would change the result.

Trials were rejected from analysis if their reaction time was below 200ms or above 5000ms. This resulted in an average loss of 4.26% (SD = 2.54%) of trials per participant. Use of a more liberal time criterion (8000ms) led to exclusion of < 1% of the data and did not change the results in any meaningful way. We report results using the stricter criterion to reduce contamination from lapses of attention or other extraneous factors. One participant was excluded from analysis for having an error rate more than 2.5 standard deviations above the group mean of 1.51% (SD = 1.57%). Incorrect trials were excluded from all reaction time and eye-tracking analyses.

Reaction Times

A 4 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on reaction time (RT) data from the uncued trials (**Figure 2a**), revealed a significant main effect of epoch, $F(3,42) = 9.88$, $p < .001$, $\eta^2 = .41$, demonstrating that participants' performance improved throughout the training period. A significant main effect of display repetition also emerged, $F(1,14) = 5.20$, $p = .039$, $\eta^2 = .27$, such that participants responded significantly faster to repeated displays (M = 2135ms) compared to novel displays (M = 2280ms). The interaction between epoch and display repetition was not significant, $F < 1$. This is a common finding due to contextual cueing emerging within the first epoch of the experiment (e.g., Chun & Jiang, 1998; Peterson & Kramer, 2001). As reaction times were faster for repeated displays than for novel displays, we have demonstrated the presence of the typical contextual cueing effect prior to the introduction of the spatial cues.

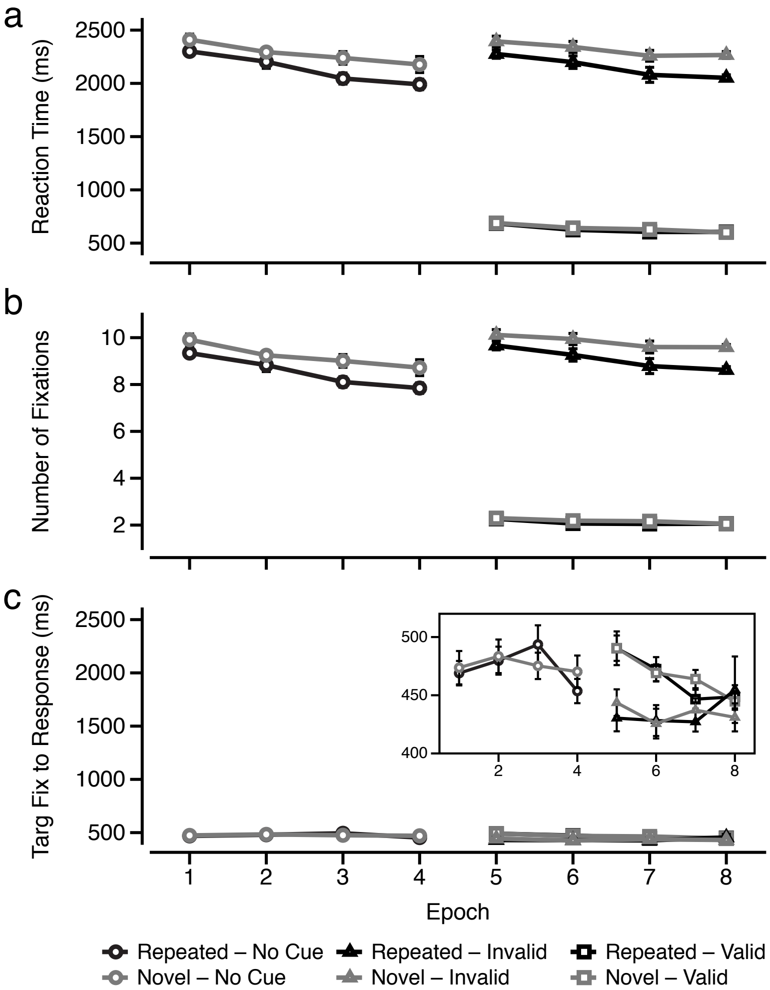

Figure 2: Experiment 1 results. **a** Reaction time data. **b** Number of fixations prior to response. **c** Time from target fixation to response, plotted on the same y-scale as the reaction time data for a valid comparison of effect magnitudes. Inset in c is the data scaled to show any differences. Error bars are within-participants SEM (Cousineau, 2005; Morey, 2008), but are generally too small to be seen.

A 4 (epoch) x 2 (display repetition: repeated, novel) x 2 (cue validity: valid, invalid) repeated measures ANOVA run on RT data from the second half of the experiment, during which spatial cues were present, revealed a significant main effect of epoch, $F(3,42) = 12.02$, $p < .001$, $\eta^2 = .46$. There was also a significant main effect of display repetition, $F(1,14) = 10.08$, $p = .007$, $\eta^2 = .42$, reflecting faster responses for repeated displays ($M = 1391$ms) compared to novel displays ($M = 1478$ms), and a significant main effect of validity, $F(1,14) = 821.98$, $p < .001$, $\eta^2 = .98$, reflecting faster responses to validly cued displays ($M = 635$ms) than to invalidly cued displays ($M = 2234$ms). The interaction between epoch and display repetition was non-significant, $F < 1$, and the interaction between epoch and cue validity fell just shy of significance, $F(3,42) = 2.80$, $p = .051$, $\eta^2 = .17$. Critically, however, the interaction between display repetition and cue validity was significant, $F(1,14) = 25.25$, $p < .001$, $\eta^2 =$

.64. There was no significant three-way interaction between epoch, display repetition, and cue validity, $F < 1$.

We followed up the interaction of display repetition and cue validity with linear contrasts, comparing repeated and novel displays for each level of cue validity. This revealed that responses were significantly faster on invalidly cued repeated displays (M = 2152ms) than on invalidly cued novel displays (M = 2315ms), $t(21.96) = 5.22$, $p < .001$, $d = 1.35$. However, there was no significant RT difference between validly cued repeated displays (M = 630ms) and validly cued novel displays (M = 640ms), $t < 1$. Furthermore, comparing the magnitude of contextual cueing between validly cued and invalidly cued conditions (novel RT – repeated RT, collapsed across epochs) revealed contextual cueing magnitude to be significantly larger for the invalidly cued condition, $t(14) = 4.94$, $p < .001$, $d = 1.28$. Thus, contextual cueing was observed following invalid spatial cues, however, no evidence of contextual cueing was observed following valid spatial cues.

Eye Tracking

Number of fixations. The pattern of results for the number of fixations per trial (**Figure 2b**) closely matched the RT results. A 4 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the number of fixations per trial in the uncued period revealed a significant main effect of epoch, $F(3,42) = 13.04$, $p < .001$, $\eta^2 = .48$, and a significant main effect of display repetition, $F(1,14) = 5.98$, $p = .028$, $\eta^2 = .30$, such that repeated displays (M = 8.53) were completed with significantly fewer fixations than novel displays (M = 9.22). There was no significant interaction between epoch and display repetition, $F < 1$.

A 4 (epoch) x 2 (display repetition: repeated, novel) x 2 (cue validity: valid, invalid) repeated measures ANOVA on the number of fixations made in each trial of the cued period of the experiment produced a significant main effect of epoch, $F(3,42) = 9.59$, $p < .001$, $\eta^2 = .41$, a significant main effect of display repetition, $F(1,14) = 10.16$, $p = .007$, $\eta^2 = .42$, and a significant main effect of validity, $F(1,14) = 772.12$, $p < .001$, $\eta^2 = .98$. There was no significant interaction between epoch and display repetition, $F < 1$. The interaction between epoch and cue validity was significant, $F(3,42) = 4.09$, $p = .012$, $\eta^2 = .23$, as was the interaction between display repetition and cue validity, $F(1,14) = 23.53$, $p < .001$, $\eta^2 = .63$. There was no significant three-way interaction between epoch, display repetition, and cue validity, $F(3,42) = 1.05$, $p = .380$, $\eta^2 = .07$.

As in the RT data, we followed up the significant interaction of display repetition and cue validity with linear comparisons between repeated and novel displays for each level of cue validity. These revealed that there were significantly more fixations required to find the target on invalidly cued novel displays ($M = 9.81$) than on invalidly cued repeated displays ($M = 9.08$), $t(21.77) = 5.13$, $p < .001$, $d = 1.33$, however, there was no difference between the number of fixations required to find the target on validly cued novel displays ($M = 2.18$) and validly cued repeated displays ($M = 2.11$), $t < 1$. Comparing the magnitude of contextual cueing between validly cued and invalidly cued conditions (number of fixations on novel trials – number of fixations on repeated trials, collapsed across epochs) revealed contextual cueing magnitude to be significantly larger for the invalidly cued condition, $t(14) = 4.74$, $p < .001$, $d = 1.22$.

Time from target fixation to button press. In the uncued period, the 4 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the time between fixating the target and emitting a response (**Figure 2c**), showed no significant main effects or interactions, *p*s > .3. In the cued period, the 4 (epoch) x 2 (display repetition: repeated, novel) x 2 (cue validity: valid, invalid) repeated measures ANOVA on the time between fixating the target and emitting a response, revealed a main effect of validity, $F(1,14) = 7.53$, $p = .016$, $\eta^2 = .35$, and an interaction between epoch and validity, $F(3,42) = 4.25$, $p = .010$, $\eta^2 = .23$. None of the other main effects or interactions were significant, *p*s > .15, including all effects involving display repetition. As such, we find no evidence that contextual cueing influences response processes in this experiment.

Discussion

The results here clearly support improved attentional guidance on repeated displays as the source of the contextual cueing effect. The effect of display repetition on reaction times was present only in conditions in which guidance was not already optimal – uncued trials, and invalidly cued trials. Furthermore, the eye-tracking data show repeated displays reduce the number of fixations required to find a target, but have no effect on the time between fixating the target and emitting a response. All of this evidence converges on the conclusion that contextual cueing improved attentional guidance in this experiment.

This result is at odds with Schankin and Schubö (2010), who observed contextual cueing on validly cued trials. One difference between Schankin and Schubö (2010) and the current experiment that may account for these different results, is that search in this experiment was quite slow, with reaction times above two seconds in all of the conditions that showed contextual cueing. Previous evidence arguing against the involvement of attentional guidance improvements in contextual cueing has come from papers with relatively fast search times, well under one second on average. Indeed, Kunar, Flusberg, and Wolfe (2008) have previously argued that contextual cueing does improve attentional guidance when search times are slow, but that in typical search tasks that are completed in under a second, attentional guidance does not have time to be engaged. They argue that another process, probably response related (Kunar et al. 2007), underlies the contextual cueing effect in faster searches.

Another difference between Schankin and Schubö (2010) and our Experiment 1, that may account for the different pattern of results, is that Schankin and Schubö employed only four target locations across all of their repeated and novel search displays, while we allowed targets to appear almost anywhere in the search display, and never in the same location for two different repeated displays. Schankin and Schubö's use of overlapping target locations with different displays may have reduced the guidance component of contextual cueing by speeding search, as noted above, but it may also have affected the acquisition or expression of contextual cueing in some other way. For instance, their procedure associated a given target location with multiple distractor contexts, potentially interfering with guidance by repeated distractors. Experiments 2a and 2b were performed to test these possibilities.

Experiment 2

As noted above, attentional guidance may have driven contextual cueing in Experiment 1 because search was slow, or because each target location was associated with only one display layout (in the repeated displays condition). To test these possibilities, we ran two new experiments, identical to Experiment 1 except that the targets were restricted to only four locations. In Experiment 2a, these were relatively distant from fixation, while in Experiment 2b the targets were situated quite close to fixation. Thus, if repetition of target locations across multiple displays interferes with attentional guidance in contextual cueing, both of these experiments should show no reduction in the number of fixations needed to find the target. However, if attentional guidance is related to search speed, then we may see a dissociation between these experiments, with fewer fixations on repeated trials in the slower search of Experiment 2a, but not in the faster search of Experiment 2b.

Method

Participants

Sixteen participants took part in each of Experiment 2a (9 females, *M* = 23.94 years, *SD* = 6.97 years) and Experiment 2b (11 females, *M* = 21.94 years, *SD* = 5.95 years). These sample sizes were selected *a priori* as being typical for contextual cueing experiments. Participation was voluntary and participants were compensated for their time at a rate of $10 per hour. All reported normal vision. One participant was excluded from Experiment 2a due to a technical error that resulted in no eye-tracking data being recorded for that participant.

Stimuli and Procedure

All aspects of the methods of Experiments 2a and 2b were the same as Experiment 1 except that the target locations were fixed to four locations at the corners of an imaginary square around fixation. Thus, the target locations for different displays overlapped such that each of the four target locations was associated with two repeated displays and two novel displays per block. In Experiment 2a the target locations were relatively distant from the starting position of the search (10.5°). In Experiment 2b they were relatively close to the starting position (5°).

Results - Experiment 2a

Trials were rejected from analysis if their reaction time was below 200ms or above 5000ms. This resulted in an average loss of 3.56% (*SD* = 2.50%) of trials per participant. Use of a more liberal time criterion did not change the results in any meaningful way. Trials with incorrect responses resulted in a further loss of 1.14% (*SD* = 0.91%) of all trials.

Reaction Times

A 4 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the reaction time data from the first, uncued, half of the experiment (**Figure 3a**) revealed a significant main effect of epoch, $F(3,42) = 12.97$, $p < .001$, $\eta^2 = .48$. There was no significant main effect of display repetition, $F(1,14) = 2.65$, $p = .126$, $\eta^2 = .16$, and no significant interaction between epoch and display repetition, $F(3,42) = 1.84$, $p = .155$, $\eta^2 = .12$. However, a planned comparison of the difference between repeated and novel trials within the fourth epoch revealed that contextual cueing did emerge by the end of the uncued period, $t(14) = 2.17$, $p = .048$, $d = 0.56$.

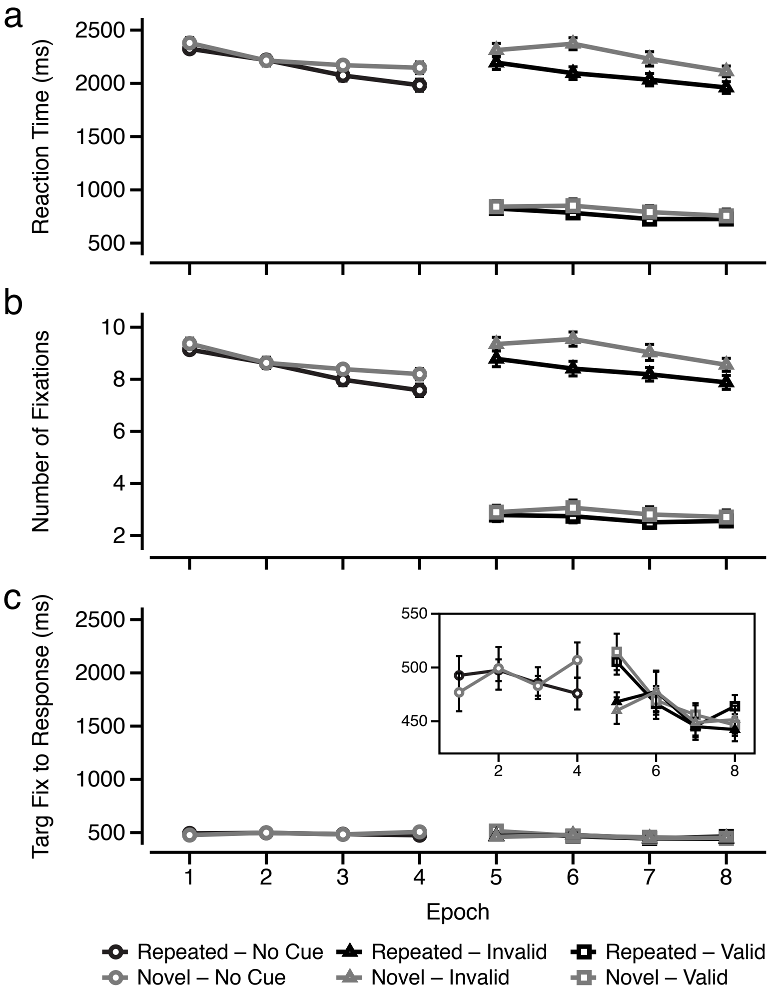

Figure 3: Experiment 2a results. **a** Reaction times. **b** Number of fixations prior to response. **c** Time from target fixation to response, plotted on the same y-scale as the reaction time data for a valid comparison of effect magnitudes. Inset in c is the data scaled to show any differences. Error bars are within-participants SEM (Cousineau, 2005; Morey, 2008), but are often too small to be seen.

A 4 (epoch) x 2 (display repetition: repeated, novel) x 2 (cue validity: valid, invalid) repeated measures ANOVA run on the RT data from the second half of Experiment 2a, during which spatial cues were present, revealed a significant main effect of epoch, $F(3,42) = 12.37$, $p < .001$, $\eta^2 = .47$. There was also a significant main effect of display repetition, $F(1,14) = 27.32$, $p < .001$, $\eta^2 = .66$, and a significant main effect of validity, $F(1,14) = 203.54$, $p < .001$, $\eta^2 = .94$. All two-way interactions were significant: epoch by display repetition, $F(3,42) = 4.43$, $p = .009$, $\eta^2 = .24$; epoch by cue validity, $F(3,42) = 6.68$, $p < .001$, $\eta^2 = .32$; display repetition by cue validity, $F(1,14) = 21.75$, $p < .001$, $\eta^2 = .61$. The three-way interaction of epoch, display repetition, and cue validity was non-significant, $F < 1$.

As we are not concerned with the change in contextual cueing across epochs, we chose to only follow up the interaction between display repetition and cue validity. We compared repeated with novel displays for valid and invalid cues separately using linear comparisons. Responses were significantly faster to invalidly cued repeated displays (M = 2072ms) than to invalidly cued novel displays (M = 2257ms), $t(24.73) = 6.95$, $p < .001$, $d = 1.79$. However, there was no significant difference between response times to validly cued repeated displays (M = 765ms) and validly cued novel displays (M = 810ms), $t(24.73) = 1.68$, $p = .105$, d = 0.43. It could be argued that use of a one-tailed test is more appropriate here, given our clear directional hypothesis. This would have produced a result of $p = .0525$. This may indicate the presence of a weak effect of display repetition following valid spatial cues. Comparing the magnitude of contextual cueing between validly cued and invalidly cued conditions (novel RT – repeated RT, collapsed across epochs) revealed contextual cueing magnitude to be significantly larger for the invalidly cued condition, $t(14) = 4.68$, $p < .001$, $d = 1.21$. Thus, contextual cueing was primarily observed on invalidly cued trials.

Eye Tracking

Number of fixations. The pattern of results for the number of fixations per trial closely matched the RT results (**Figure 3b**). A 4 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the number of fixations per trial in the uncued period revealed a significant main effect of epoch, $F(3,42) = 18.02$, $p < .001$, $\eta^2 = .56$. There was no significant main effect of display repetition, $F(1,14) = 2.03$, $p = .176$, $\eta^2 = .13$, and no significant interaction, $F(3,42) = 1.60$, $p = .204$, $\eta^2 = .10$.

A 4 (epoch) x 2 (display repetition: repeated, novel) x 2 (cue validity: valid, invalid) repeated measures ANOVA on the number of fixations from each trial of the cued period produced a significant main effect of epoch, $F(3,42) = 12.22$, $p < .001$, $\eta^2 = .47$, a significant main effect of display repetition, $F(1,14) = 21.54$, $p < .001$, $\eta^2 = .61$, and a significant main effect of validity, $F(1,14) = 162.34$, $p < .001$, $\eta^2 = .92$. The interaction of epoch and display repetition was significant, $F(3,42) = 3.84$, $p = .016$, $\eta^2 = .22$, as was the interaction of epoch and cue validity, $F(3,42) = 8.21$, $p < .001$, $\eta^2 = .37$, and the interaction of display repetition and cue validity, $F(1,14) = 19.22$, $p < .001$, $\eta^2 = .58$. The three-way interaction of epoch, display repetition, and cue validity was non-significant, $F < 1$.

The interaction between display repetition and cue validity was followed up with linear comparisons comparing repeated with novel displays for valid trials, and for invalid trials.

These revealed that significantly fewer fixations were required to find the target on repeated invalidly cued trials (M = 8.32) than on novel invalidly cued trials (M = 9.12), $t(22.96) = 6.24$, $p < .001$, $d = 1.61$, however, there was no difference in the number of fixations between repeated (M = 2.65) and novel (M = 2.87) validly cued trials, $t(22.96) = 1.72$, $p = .099$, d = 0.44. Use of a one-tailed test here gave a significant difference between repeated and novel validly cued trials, $p = .0495$. Comparing the magnitude of contextual cueing between validly cued and invalidly cued conditions (number of fixations on novel trials – number of fixations on repeated trials, collapsed across epochs) revealed contextual cueing magnitude for the number of fixations per trial to be significantly larger for the invalidly cued condition, $t(14)$ = 4.38, $p < .001$, $d = 1.13$.

Time from target fixation to button press. A 4 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the time between fixating the target and responding in the uncued period of Experiment 2a (**Figure 3c**) revealed no significant main effects or interaction, all ps > .3. In the cued period, the only significant effects from the 4 (epoch) x 2 (display repetition: repeated, novel) x 2 (cue validity: valid, invalid) repeated measures ANOVA were the main effect of epoch, $F(3,42) = 5.12$, $p = .004$, $\eta^2 = .27$, and the interaction between epoch and validity, $F(3,42) = 4.03$, $p = .013$, $\eta^2 = .22$. All other main effects and interactions, including all effects involving display repetition, were non-significant, ps > .2. Thus, we find no evidence of display repetitions influencing response processes in this experiment.

Results - Experiment 2b

Trials were rejected from analysis if their reaction time was below 200ms or above 5000ms. This resulted in an average loss of 0.27% (SD = 0.29%) of trials per participant. Exclusion of trials containing incorrect responses resulted in a loss of a further 2.04% of trials (SD = 1.31%).

Reaction Times

A 4 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the reaction time data from the uncued portion of Experiment 2b (**Figure 4a**) revealed a significant main effect of epoch, $F(3,45) = 10.78$, $p < .001$, $\eta^2 = .42$. The main effect of display repetition was non-significant, $F < 1$, as was the interaction of epoch and display repetition, $F(3,45) = 1.31$, $p = .284$, $\eta^2 = .08$. A planned comparison showed no evidence of a difference between repeated and novel displays in the final epoch of the uncued period, $t < 1$, suggesting contextual cueing had not emerged by this time.

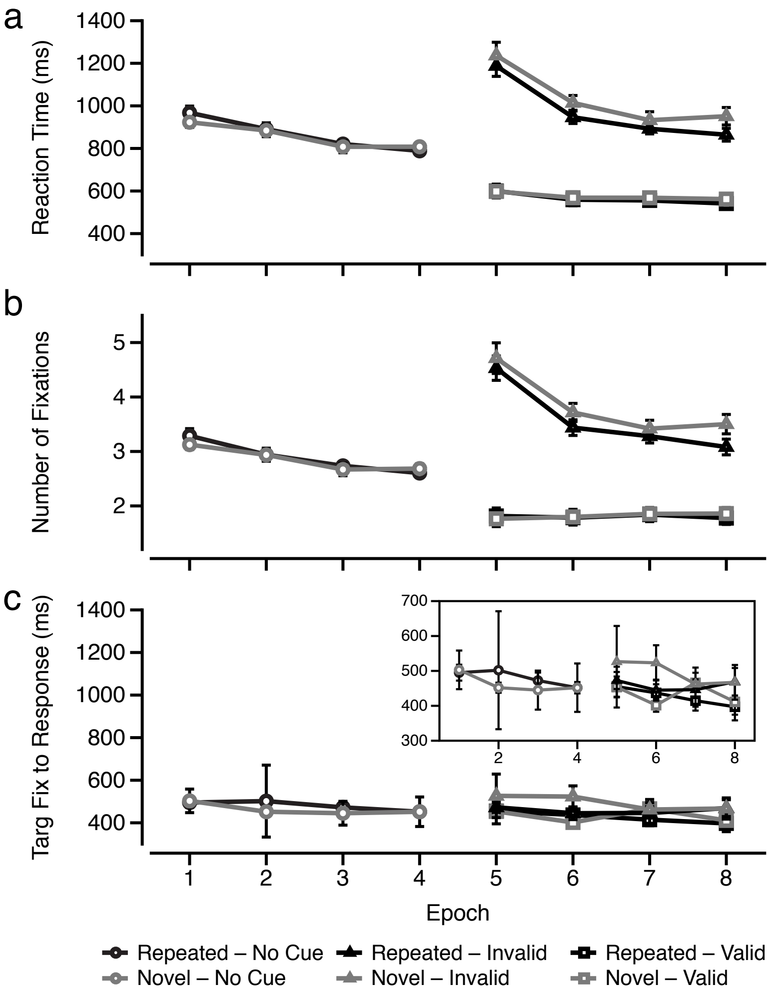

Figure 4: Experiment 2b results. **a** Reaction time data. **b** Number of fixations prior to response. **c** Time from target fixation to response, plotted on the same y-scale as the reaction time data for a valid comparison of effect magnitudes. Inset in c is the data scaled to show any differences. Error bars are within-participants SEM (Cousineau, 2005; Morey, 2008), but are generally too small to be seen.

A 4 (epoch) x 2 (display repetition: repeated, novel) x 2 (cue validity: valid, invalid) repeated measures ANOVA on reaction times from the cued period of Experiment 2b revealed a significant main effect of epoch, $F(3,45) = 23.47$, $p < .001$, $\eta^2 = .61$, a significant main effect of display repetition, $F(1,15) = 7.76$, $p = .014$, $\eta^2 = .34$, and a significant main effect of validity, $F(1,15) = 81.66$ $p < .001$, $\eta^2 = .85$. Significant interactions between epoch and validity, $F(3,45) = 15.71$, $p < .001$, $\eta^2 = .51$, and between display repetition and validity, $F(1,15) = 6.72$, $p = .020$, $\eta^2 = .31$, were also revealed. The interaction of display repetition and epoch, and the three-way interaction, were both non-significant, *p*s > .3.

As in the previous experiments, we chose to follow up the interaction between display repetition and validity by computing linear comparisons comparing repeated with novel displays for valid and invalid cues separately. Responses were significantly faster to invalidly cued repeated displays (*M* = 973ms) than to invalidly cued novel displays (*M* = 1034ms), $t(28.14) = 3.79$, $p < .001$, $d = 0.95$. However, there was no significant difference between response times to validly cued repeated displays (*M* = 564ms) and validly cued novel displays (*M* = 574ms), $t < 1$. Comparing the magnitude of contextual cueing between validly cued and invalidly cued conditions (novel RT – repeated RT, collapsed across epochs) revealed contextual cueing magnitude to be significantly larger for the invalidly cued condition, $t(15) = 2.59$, $p = .021$, $d = 0.65$. Thus, contextual cueing was observed only on invalidly cued trials in Experiment 2b.

Eye Tracking

Number of fixations. Once again, the pattern of results for the number of fixations per trial closely matched the RT results (**Figure 4b**). A 4 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the number of fixations per trial in the uncued period revealed a significant main effect of epoch, $F(3,45) = 11.91$, $p < .001$, $\eta^2 = .44$. There was no significant main effect of display repetition and no significant interaction, *F*s < 1

A 4 (epoch) x 2 (display repetition: repeated, novel) x 2 (cue validity: valid, invalid) repeated measures ANOVA on the number of fixations from each trial of the cued period produced a significant main effect of epoch, $F(3,45) = 18.11$, $p < .001$, $\eta^2 = .55$, a significant main effect of display repetition, $F(1,15) = 7.26$, $p = .017$, $\eta^2 = .33$, and a significant main effect of validity, $F(1,15) = 62.98$, $p < .001$, $\eta^2 = .81$. The interaction of epoch and display repetition was non-significant, $F(3,45) = 1.96$, $p = .134$, $\eta^2 = .12$. However, the interaction of epoch and cue validity was significant, $F(3,45) = 22.87$, $p < .001$, $\eta^2 = .60$, as was the interaction of display repetition and cue validity, $F(1,15) = 7.99$, $p = .013$, $\eta^2 = .35$. The three-way interaction of epoch, display repetition, and cue validity was non-significant, $F < 1$.

We followed up the interaction between display repetition and cue validity with linear comparisons, comparing repeated with novel displays for invalid trials, and for valid trials. These revealed that significantly fewer fixations were required to find the target on invalidly cued repeated displays (*M* = 3.58) than on invalidly cued novel displays (*M* = 3.84), $t(29.23) = 3.88$, $p < .001$, $d = 0.97$, however, there was no difference in the number of fixations between repeated (*M* = 1.80) and novel (*M* = 1.82) validly cued trials, $t < 1$. Comparing the magnitude of contextual cueing between validly cued and invalidly cued conditions (number of fixations on novel trials – number of fixations on repeated trials, collapsed across epochs) revealed contextual cueing magnitude for the number of fixations per trial to be significantly larger for the invalidly cued condition, $t(15) = 2.80$, $p = .014$, $d = 0.70$.

While the difference in fixations between invalidly cued repeated and novel displays is quite small, on average only about one quarter of a fixation per display (or one fixation roughly every four trials), this effect is both statistically reliable and consistent with the smaller reaction time benefit observed in this experiment. Indeed, the average combined saccade and fixation duration in this experiment was 249ms, which gives an expected RT benefit of (249ms/fixation * 0.26 fixations) 64.74ms on invalidly cued repeated trials, which aligns closely with the observed RT benefit of 61ms.

Time from target fixation to button press. There were four participants who never directly fixated the target in at least one condition of this experiment, presumably because, with the targets this close to fixation, these participants were able to discriminate the targets in the near-periphery. As such, these participants were excluded from this analysis. A 4 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the data from the remaining twelve participants (**Figure 4c**), in the uncued period, revealed no significant main effects or interaction, all *p*s > .2. Likewise, a 4 (epoch) x 2 (display repetition: repeated, novel) x 2 (validity: valid, invalid) repeated measures ANOVA on the data from the cued period revealed no significant main effects or interactions, all *p*s > .1. Thus, once again, we find no evidence of display repetitions influencing response processes.

Discussion

The results of Experiment 2 replicate the main results of Experiment 1. All conditions in which contextual cueing was observed produced differences in the number of fixations required to find the target between repeated and novel displays. No differences were observed between repeated and novel displays in the time between fixating the target and emitting a response, in any condition. In both of these experiments contextual cueing was observed on invalidly cued trials, but there was little to no evidence of contextual cueing following valid cues, contrary to the predictions of a response related account of contextual cueing. Furthermore, when weak evidence of contextual cueing was found for validly cued trials in Experiment 2a, this was again accompanied by the same effect in the number of fixations required to find the target, suggesting any contextual cueing present on validly cued trials was also driven by improved attentional guidance. Thus, both the eye-tracking and reaction time data from Experiments 2a and 2b suggest attentional guidance as the driving force behind the contextual cueing effect.

One interesting aspect of the results that bears commenting on is the slow emergence of contextual cueing as search variability reduced. In Experiment 2a, targets were presented at one of only four locations, and reaction time benefits did not emerge until the final epoch of the training period. In this experiment, search was still quite slow (~2 seconds per trial in the training period). In Experiment 2b, targets were presented at only four locations, and were placed considerably closer to fixation. This had the effect of considerably speeding search (<1 second per trial in the training period), and further slowing the emergence of contextual cueing, which was not apparent until the cued period of the experiment. Note that with only four locations, each target location is associated with many more, novel displays than would be the case with Experiment 1, or the typical contextual cueing experiment. These results are consistent with our earlier conjecture that associating repeated and novel distractor contexts with the same target locations weakens learning. This effect may be similar to previous reports that multiple associations between target locations and distractor contexts can interfere with contextual cueing (Kunar & Wolfe, 2011; Zellin, Conci, von Mühlenen, & Müller, 2011; Zellin, von Mühlenen, Müller, & Conci, 2013). ``

It is important to note that the slow emergence of contextual cueing in Experiments 2a and 2b does not provide an alternative explanation for why little to no contextual cueing was observed in the validly cued condition, for two reasons. First, contextual cueing was observed in the validly cued condition of Schankin and Schübo (2010) despite the absence of pre-established learning in their paradigm. Second, no contextual cueing was observed in the valid condition of Experiment 1, despite the preexisting learning established in the uncued period. Thus, the lack of contextual cueing on valid trials cannot be attributed to the absence of a preexisting effect. Furthermore, it cannot be argued that the limited contextual cueing on validly cued trials is because participants simply hadn't learned the displays, because the same displays produced strong contextual cueing in the invalidly cued condition. Thus, the difference in contextual cueing between the validly cued and invalidly cued conditions seems to be due to how participants' attention was guided by the spatial cues in the two conditions.

It might be argued that as we observed significant guidance-related contextual cueing in Experiment 2b, search must still have been too slow to produce contextual cueing without attentional guidance (Kunar, Flusberg, & Wolfe, 2008). This is an implausible explanation, however, as search times on invalidly cued trials were generally quite fast, averaging roughly 1 second per trial, which is faster than the average responses to conditions that produced no evidence of guidance in Kunar, Flusberg, and Wolfe (2008; Experiment 2, Standard condition). Thus, if attentional guidance improvements only emerged in slower searches, we would not expect them to be shown here. However, as there were a number of differences between our Experiment 2b and Kunar, Flusberg, and Wolfe (2008), we cannot be sure that time-related guidance improvements would emerge across the same timeline in these experiments. Experiment 3 was conducted to provide a stricter test of the claim that attentional guidance improvements related to contextual cueing only emerge under slower searches.

Experiment 3

As noted in the introduction, one of the arguments put forward against the involvement of attentional guidance in the contextual cueing effect is that contextual cueing is observed when guidance is supposedly optimal (i.e., following valid spatial cues, or in pop-out search). In the previous experiments we found little to no evidence of contextual cueing following valid spatial cues (but see combined analysis below). This supports an attentional guidance account of contextual cueing, but limits the conclusions we can draw about contextual cueing under conditions of optimal guidance. Contextual cueing under pop-out conditions, however, has been shown several times (Geyer, Zehetleitner, & Müller, 2010; Kunar, Flusberg, Horowitz, & Wolfe, 2007), providing evidence that learning occurs even under conditions of near-optimal search.

Experiment 3 further explores the effect of display repetition in near-optimal search conditions by having participants complete a pop-out contextual cueing task while we tracked their gaze. Furthermore, to provide a more complete delineation of the search into guidance related and response related periods, we used gaze-contingent targets that were masked until participants fixated them. This allowed us precise knowledge of how long participants had to execute response related processing, as target identity could not be extracted while the target was in the periphery. If contextual cueing effects involve improvements to processes other than attentional guidance, particularly when attentional guidance is already optimal, then this experiment should show evidence of such.

Methods

Participants

Given the small and highly variable contextual cueing effect observed in the pop-out experiment of Kunar et al. (2007, their Experiment 2b), we decided *a priori* to increase our statistical power by doubling the number of participants tested for this experiment, relative to the previous experiments. As such, thirty-two participants took part in Experiment 3 (26 females, M = 22.41 years, SD = 2.76 years). Participation was voluntary and participants were compensated for their time at a rate of $10 per hour. All reported normal vision.

Stimuli and Procedure

Experiment 3 was patterned after Experiment 1, but modified as specified below. In Experiment 3, spatial clues were not employed, so trials contained only the fixation display that offset after 500ms of fixation (see Experiment 1 Methods for detailed description), followed by the search display that was present until a response was made. Participants completed 960 trials, divided into 10 epochs of 96 trials each. All distractor items were green H shapes (RGB: 0, 255, 0; dimensions identical to the T stimuli from Experiment 1). Due to the ease of responding without fixating the target in experiment 2b, we used gaze contingent targets to ensure all participants fixated the target prior to responding. These were red H stimuli (RGB: 255, 0, 0; dimensions identical to the distractor items). When fixated within 2° from the center of the target, one vertical bar of the target H offset, transforming the target item to a sideways T, to which participants responded in the same way as in the previous experiments. The offset bar was randomly determined as the left or right bar, and was not predicted by the search display. By using gaze contingent targets, we are able to know precisely how long a participant was able to extract target related information prior to emitting a response, as there was no possibility of peripheral targets being identified prior to being fixated. Thus, we provide a strict test of the claim that response related processes are the root source of contextual cueing (Kunar et al. 2007), and that attentional guidance only comes into play in slower searches (Kunar, Flusberg, & Wolfe, 2008).

Results

Trials were rejected from analysis if their reaction time was below 200ms or above 2000ms. The earlier limit on exclusion time relative to the previous experiments was selected *a priori* to be more appropriate to the faster search expected in this experiment. This resulted in an average loss of 0.47% (*SD* = 0.79%) of trials per participant. Trials with incorrect responses were excluded from analysis. This resulted in the loss of a further 1.61% of all trials (*SD* = 1.20%). One participant was excluded from analysis for having an error rate more than 2.5 standard deviations higher than the group mean. Two further participants were excluded from analysis for having average reaction times more than 2.5 standard deviations slower than the group mean. Reanalysis with these participants included did not qualitatively alter the results.

Reaction Times

A 10 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on participants' reaction times (**Figure 5a**) revealed a significant main effect of epoch, $F(9,252) = 5.10$, $p = .001$, $\eta^2 = .15$, and a significant main effect of display repetition, $F(1,28) = 10.64$, $p = .003$, $\eta^2 = .28$, such that responses to repeated displays (M = 708ms) were significantly faster than to novel displays (M = 722ms). There was no significant interaction between epoch and display repetition, $F < 1$. It is worth noting that the small 14ms contextual cueing effect observed here is consistent with the magnitude of contextual cueing observed in past pop-out contextual cueing studies (Geyer, Zehetleitner, & Müller, 2010; Kunar, Flusberg, Horowitz, & Wolfe, 2007).

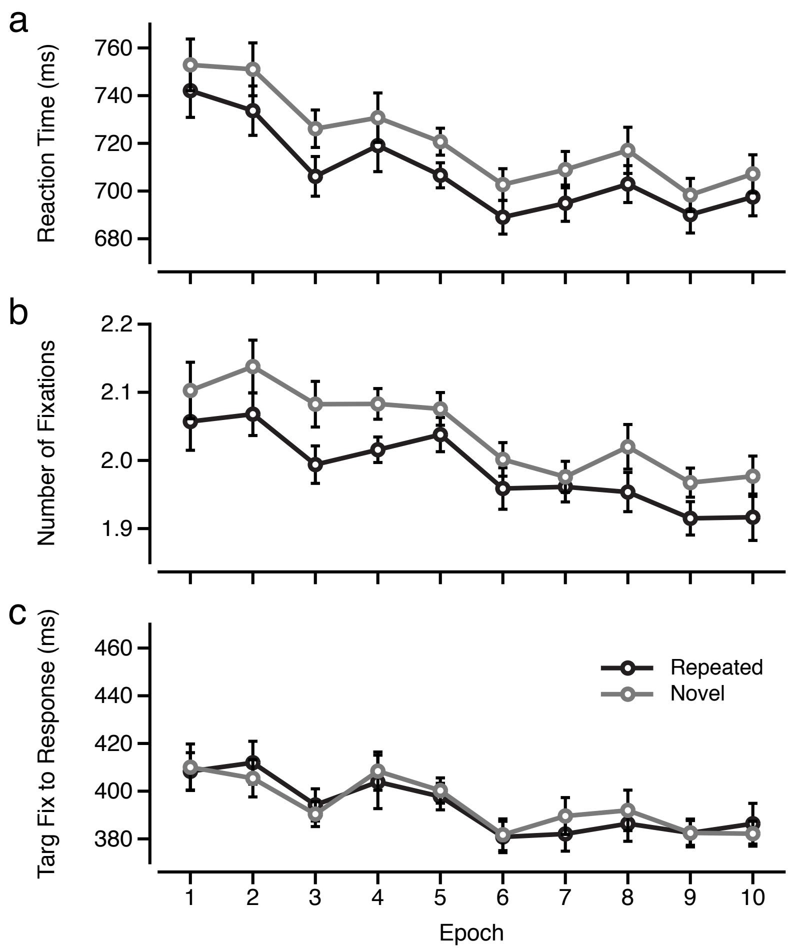

Figure 5: Experiment 3 results. **a** Reaction times. **b** Number of fixations prior to response. **c** Time from target fixation to response. Error bars are within-participants SEM (Cousineau, 2005; Morey, 2008).

Eye Tracking

Number of fixations. A 10 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the number of fixations required to find the target (**Figure 5b**) revealed a significant main effect of epoch, $F(9,252) = 4.92$, $p < .001$, $\eta^2 = .15$, and a significant main effect of display repetition, $F(1,28) = 10.31$, $p = .003$, $\eta^2 = .27$, such that on repeated displays ($M = 1.99$) fewer fixations were required to find the target than on novel displays ($M = 2.04$). This tiny but significant difference between repeated and novel displays (0.05 fixations per trial) may seem surprising. However, given that saccades and fixations had an average combined duration of 236ms in this experiment, a difference of one saccade every twenty trials is consistent with the average contextual cueing effect of 14ms observed in this experiment (236ms/fixation * .05 fixations = 11.82ms). The interaction between epoch and display repetition was non-significant, $F < 1$.

Time from target fixation to button press. A 10 (epoch) x 2 (display repetition: repeated, novel) repeated measures ANOVA on the time from target fixation to response (**Figure 5c**) revealed a significant main effect of epoch, $F(9,252) = 2.70$, $p = .031$, $\eta^2 = .09$. There was no main effect of display repetition and no interaction, Fs < 1. Thus, in Experiment 3, we find no evidence of display repetitions influencing response processes.

Discussion

The results of Experiment 3 provide further evidence that the contextual cueing effect is produced primarily by improved attentional guidance. As in the previous experiments, reaction time benefits in repeated displays were accompanied by similar reductions in the number of fixations required to find the target, while no difference between repeated and novel displays was observed in the time from fixating the target to emitting a response. Our use of gaze contingent targets provides strong evidence that response related processes were not improved by repeated display configurations, as it cannot be the case that participants identified the target in the periphery and initiated response selection prior to fixating the target. Thus, all response related processing must have been captured in the time between target onset (the time of target fixation) and the button press. These results argue strongly against the proposal that attentional guidance is only facilitated in slow search tasks (Kunar, Flusberg, & Wolfe, 2008). With this experiment we cannot definitively rule out the possibility that the gaze contingent targets induced a strategy of attentional guidance improvement, whereas with targets visible in the periphery the benefit may accrue to response related processes. However, the similarity of the results of Experiment 3 to those of the previous experiments that did not use gaze contingent targets argues against such an account.

Correlations

To further quantify the strength of the relationship between the contextual cueing benefit observed in reaction times, and both attentional guidance and response related processes, we computed correlations between individuals' contextual cueing effects observed in the reaction times, and in each of the two eye tracking measures –number of fixations, and time from target fixation to response. Data for these correlations were collapsed across experiments for the separate spatial cueing conditions (uncued, validly cued, and invalidly cued), with the pop-out experiment combined with the validly cued conditions of Experiments 1, 2a, and 2b. Contextual cueing magnitudes were calculated as the difference between repeated and novel conditions, averaged across all epochs for the relevant condition and measure.

These analyses (**Figure 6**) showed extremely strong correlations between individuals' contextual cueing magnitude in the reaction time data, and in the number of fixations required to find the target, for each of the three cueing conditions; uncued, $r = .93$, $p < .001$; validly cued/pop-out, $r = .95$, $p < .001$, and invalidly cued, $r = .88$, $p < .001$. In contrast, there were no significant correlations between contextual cueing magnitude in the reaction time data and the time from fixating the target to emitting a response; uncued, $r = -.14$, $p = .370$; validly cued/pop-out, $r = -.09$, $p = .469$, and invalidly cued, $r = -.09$, $p = .572$. Importantly, the correlations for the validly cued condition are qualitatively the same if Experiment 3 is not included in that condition; RT – number of fixations: $r = .96$, $p < .001$, and RT – time between fixation and response: $r = -.11$, $p = .472$, and so are not driven by the pop-out trials. The correlations between contextual cueing magnitude in the reaction times and number of fixations were all significantly larger than those between reaction times and the time from target fixation to response (assessed using a test for differences in correlated correlation coefficients; Meng, Rosenthal, & Rubin, 1992); uncued, $Z = 7.63$, $p < .001$; validly cued/pop-out, $Z = 10.78$, $p < .001$, and invalidly cued, $Z = 5.93$, $p < .001$.

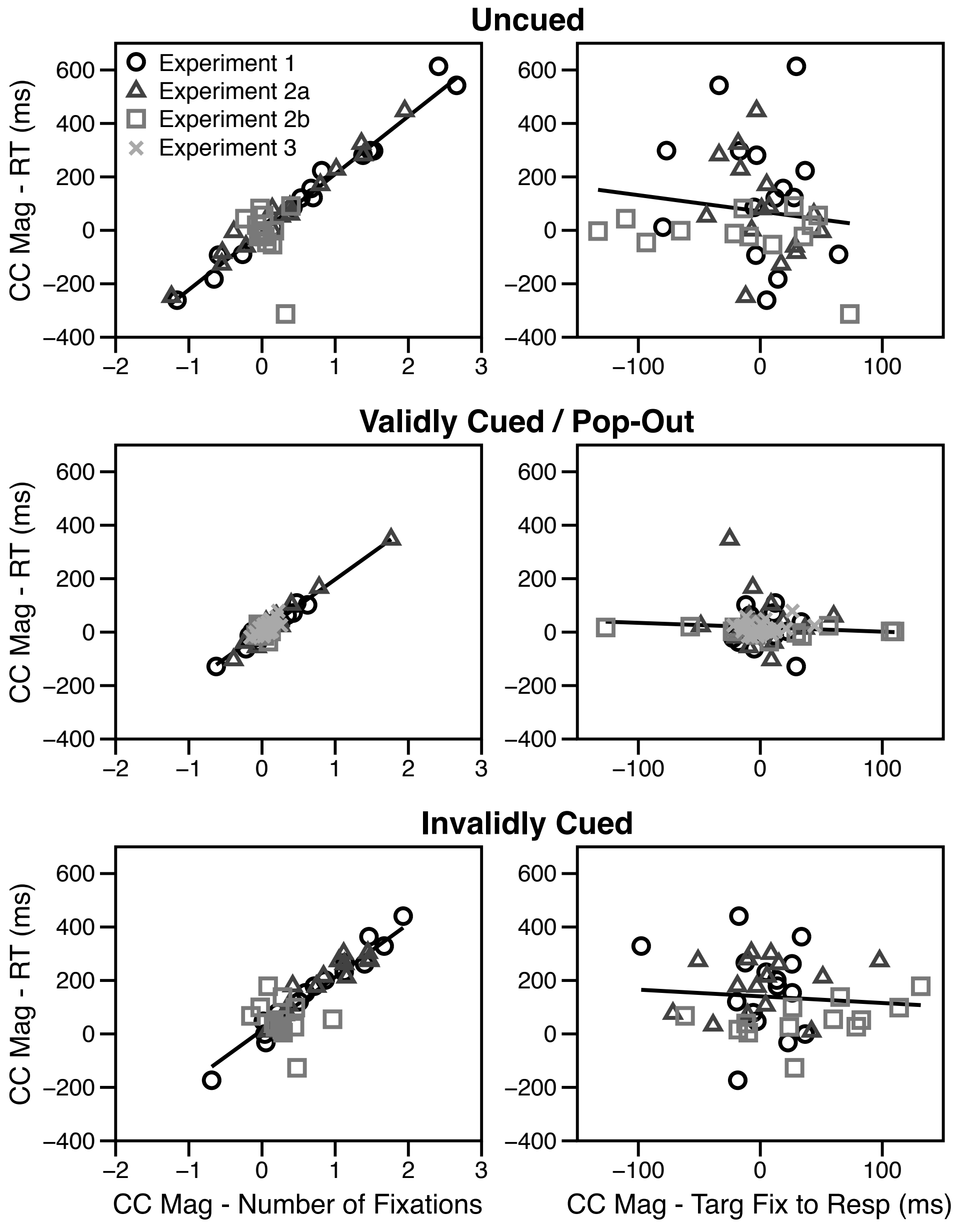

Figure 6: Correlations between contextual cueing magnitude (*CC Mag*, calculated as Novel – Repeated) produced in the reaction time data, and the two eye-tracking measures – number of fixations prior to response (left column) and time from target fixation to response (right column), for the three cueing conditions. Data have been collapsed across epochs and combined across all experiments. Experiment 3 data was included in the validly cued condition.

These results are interesting as they suggest that even in the validly cued condition, that showed little evidence of a contextual cueing benefit on average, any benefit of repeated context that may have been present at the individual level was still driven entirely by differences in the number of fixations compared to novel displays. Interestingly, computing the contextual cueing effect for validly cued trials with the combined data from Experiments 1, 2a, and 2b does show a significant reaction time difference between validly cued repeated

(M = 651ms) and novel (M = 673ms) displays, one-tailed $t(45) = 2.02$, $p = .025$, $d = .30$, consistent with the results of Schankin and Schubö (2010). Once again, this was reflected in a significant reduction in the number of fixations required to find the target on validly cued repeated displays (M = 2.18 fixations) relative to novel displays (M = 2.28 fixations), one-tailed $t(45) = 1.92$, $p = .031$, $d = .28$. No significant difference in the time from fixating the target and responding was observed between repeated (M = 457ms) and novel displays (M = 459ms) in the combined data, $t < 1$. Under the current conditions, contextual cueing on validly cued trials seems to be extremely variable, and so was not clearly observed in the individual experiments. Nonetheless, the results of this analysis suggest that although contextual cueing may be observed when targets are validly cued (Schankin & Schubö, 2010), it is still driven by improvements in attentional guidance, just as in the uncued and invalidly cued conditions.

General Discussion

This study aimed to examine the mechanisms underlying the contextual cueing effect, particularly in those conditions that have been argued to provide evidence against an attentional guidance account of contextual cueing – spatially cued and pop-out contextual cueing. We used eye tracking to divide the search into separate measures that map to distinct cognitive processes. The number of fixations required to complete the search was taken as an index of attentional guidance, and the time from when the target was fixated to when the response was made was taken as an index of response related processes (response selection, programming, and execution). We used these measures to determine whether the contextual cueing effect is produced by improvements to attentional guidance, response related processes, or both. Experiments 1, 2a, and 2b, employed spatial cues; validly cueing the target location on 50% of trials in the cued period. Experiment 3 employed pop-out search with gaze contingent targets that allowed us to have a precise measure of the amount of time available for response related processing. The results of all experiments were in agreement. In all cases, improved attentional guidance – expressed as a reduction in the number of fixations required to find the target – seems to be the driving force behind the contextual cueing effect. No differences between repeated and novel displays in the time from fixating the target to responding were found, suggesting no involvement of response related processes in contextual cueing.

Previous eye tracking studies of standard contextual cueing paradigms have consistently provided evidence for improved attentional guidance (Manelis & Reder, 2012; Manginelli &

Pollmann, 2009; Peterson & Kramer, 2001; Tseng & Li, 2004; Zhou et al., 2012), and little or no evidence of improvements to response related processes (Tseng & Li, 2004; Zhou et al., 2012). The argument that attentional guidance improvements are not the major source of contextual cueing came, in part, from the presence of contextual cueing benefits in search paradigms that supposedly produce optimal attentional guidance – namely spatially cued and pop-out search. The results of our experiments show that even under these conditions, contextual cueing effects are driven primarily by improvements to attentional guidance with repeated display exposure.

Our conclusion that contextual cueing effects are driven by improved attentional guidance even under conditions of near-optimal attentional guidance requires us to show not just a correspondence between reaction time benefits and number of fixations, but also to show that under conditions of near-optimal attentional guidance we can observe contextual cueing in the first place. In support of this, in Experiment 3 we show unambiguous contextual cueing under conditions of feature-based pop-out. Furthermore, although we observed no significant contextual cueing on validly cued trials of Experiments 1, 2a, and 2b, it is important to note that these experiments all produced small contextual cueing effects in the expected direction, and, when combined across experiments, the contextual cueing on validly cued trials was indeed significant in both the reaction times and the number of fixations required to find the target. Thus, we show evidence of contextual cueing in both of the near-optimally guided conditions employed in this study.

Our use of the term 'optimal' in the context of our valid onset cues and pop-out feature search conditions was chosen to be consistent with previous authors who have characterized these manipulations as guiding attention optimally (Kunar et al., 2007), or guiding attention in a "fast and mostly involuntary manner […], so that the visual context becomes redundant" (Schankin & Schubö, 2010, p. 718). Obviously, no formal or mathematical proof of optimality is possible, nor do we mean to imply that search cannot be improved -- in fact we were able to show improved guidance on repeated displays with both spatial cues and pop-out feature singletons. This is the reason for our repeated use of the more tentative terms 'near-optimal' and 'supposedly optimal'. Nonetheless, valid spatial cues and feature singletons greatly reduce the search required and, thus, the potential benefit of improved attentional guidance. This further emphasizes the significance of our findings of contextual cueing in these conditions and the close association between contextual cueing and number of eye fixations.

Our findings that eye movements account for virtually all of the variance in search times with repeated displays provides strong support for the view that the learning that occurs with context repetition biases attention to promote a more efficient search. By showing this with displays that are already highly efficient we extend previous work on the role of guidance in contextual cueing to show that guidance, not response related factors, underlies the learning even in efficient search conditions. Our focus on near-optimal search conditions allowed us to examine whether attentional guidance improvements are associated with contextual cueing, even under the conditions in which this is least likely to be the case. The results from these highly efficient search conditions provide important new evidence for the mechanisms that produce the contextual cueing effect more generally. Our failure to observe evidence of response facilitation with display repetition in any condition of any experiment was somewhat surprising, as there was no *a priori* reason to think that contextual cueing would not facilitate response related processes. However, the data show unambiguously that it did not. Thus, the weight of evidence provided by the combination of past research and our experiments here, suggests that, in general, contextual cueing may be produced solely by improved attentional guidance.

A guidance-only account of contextual cueing is attractive for its parsimony. If attentional guidance can be improved by display repetition, even when guidance is already highly efficient, there is no need to invoke an additional improvement to other processes such as response selection. However, one might ask, if all contextual cueing is driven by improvements to attentional guidance, why do we not see more consistent reductions in search slopes with changes in set size? We cannot answer this with certainty at this point. It is worth noting, however, the evidence showing the contextual cueing effect to be driven primarily by items local to the target region (Brady & Chun, 2007; Olson & Chun, 2002). If the effect of repetition is primarily associated with the configuration of distractors in the local vicinity of the target, then guidance improvements would not be expected across the entirety of the search episode. This would significantly weaken the link between set size and guidance.

Finally, it may be suggested that the extra fixations required to find the target on some novel trials could be related to lower confidence on those trials, where lower 'confidence' is analogous to a higher threshold for emitting a response. For example, when the eyes land close to the target, putting the target in parafoveal view, the resulting evidence could be sufficient to exceed the response criterion in repeated displays owing to increased confidence.

On novel displays, however, an extra fixation may often be required to foveate the target to achieve the needed confidence that the stimulus is indeed a target. In this way, response facilitation may masquerade as improved attentional guidance. However, the gaze-contingent method used in Experiment 3 shows that this is not the case. As the targets were masked until directly fixated, participants could not identify the target in the periphery, and confidence regarding the identity of peripheral stimuli could not have influenced the number of fixations required to perform the task. Thus, improved attentional guidance remains the strongest explanation for the observed reduction in number of fixations on repeated displays.

To summarize, in this paper we find strong support for attentional guidance as the driving force behind contextual cueing, even under conditions of near-optimal attentional guidance such as those produced by valid spatial cues or feature-based pop-out. No support was found for the suggestion that response related processes are facilitated by repeated distractor contexts. Furthermore, reductions in the number of fixations required to find the target account for the overwhelming majority of variance in the reaction time benefits produced by repeated displays, which argues against the involvement of additional processes in producing the contextual cueing effect. If other processes are involved in contextual cueing, it is now for future studies to demonstrate convincing evidence of such, beyond simply showing contextual cueing in situations in which attentional guidance seems unlikely.

Acknowledgements: The authors would like to thank Hannah Biddell for assistance with data collection. This work was funded by Discovery Grant DP1301001 from the Australian Research Council to Roger Remington.

References

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Brockmole, J. R. & Henderson, J. M. (2006). Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology, 59*(7), 1177-1187.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*(1), 28–71.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology, 1*, 42-45.

Geringswald, F., Baumgartner, F., & Pollmann, S. (2012). Simulated loss of foveal vision eliminates visual search advantage in repeated displays. *Frontiers in Human Neuroscience, 6*(134), 1-9.

Geringswald, F., Herbik, A., Hoffmann, M. B., & Pollmann, S. (2013). Contextual cueing impairment in patients with age-related macular degeneration. *Journal of Vision, 13*(3):28, 1-18.

Geringswald, F., & Pollmann, S. (2015). Central and peripheral vision loss differentially affects contextual cueing in visual search. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(5), 1485-1496.

Geyer, T., Baumgartner, F., Müller, H. J., & Pollmann, S. (2012). Medial temporal lobe-dependent repetition suppression and enhancement due to implicit vs. explicit processing of individual repeated search displays. *Frontiers in Human Neuroscience, 6*(272), 1-13.

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.

Giesbrecht, B., Sy, J. L., & Guerin, S. A. (2013). Both memory and attention systems contribute to visual search for targets cued by implicitly learned context. *Vision Research, 85*, 80-89.

Goldfarb, E. V., Chun, M. M., & Phelps, E. A. (2016). Memory-guided attention: Independent contributions of the hippocampus and striatum. *Neuron, 89*, 317-324.

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences, 19*(9), 524-533.

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 & Memory, 14*, 548-553.

Hannula, D. E., & Ranganath, C. (2009). The eyes have it: Hippocampal activity predicts expression of memory in eye movements. *Neuron, 63*, 592-599.

JASP Team (2016). JASP (Version 0.7.5.5)[Computer software].

Johnson, J. S., Woodman, G. F., Braun, E., & Luck, S. J. (2007). Implicit memory influences the allocation of attention in visual cortex. *Psychonomic Bulletin & Review, 14*(5), 834-839.

Kasper, R. W., Grafton, S. T., Eckstein, M. P., & Giesbrecht, B. (2015). Multimodal neuroimaging evidence linking memory and attention systems during visual search cued by context. *Annals of the New York Academy of Sciences*, *1339*(1), 176-189.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.1–16.

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*(4), 816–828.

Kunar, M. A., Flusberg, S. J., & Wolfe, J. M. (2008). Time to guide: Evidence for delayed attentional guidance in contextual cueing. *Visual Cognition*, *16*(6), 804–825.

Kunar, M. A., & Wolfe, J. M. (2011). Target absent trials in configural contextual cueing. *Attention, Perception, & Psychophysics, 73*, 2077–2091.

Luck, S. J. (2011). Electrophysiological correlates of the focusing of attention within complex visual scenes: N2pc and related ERP components. In: The Oxford handbook of event-related potential components (Luck, S. J., Kappenman, E. S., eds.), pp329–360. Oxford, UK: Oxford University Press.

Manelis, A., & Reder, L. M. (2012). Procedural learning and associative memory mechanisms contribute to contextual cueing: Evidence from fMRI and eye-tracking. *Learning & Memory, 19*, 527-534.

Manginelli, A. A., Baumgartner, F., & Pollmann, S. (2013). Dorsal and ventral working memory-related brain areas support distinct processes in contextual cueing. *NeuroImage, 67*, 363-374.

Manginelli, A. A., & Pollmann, S. (2009). Misleading contextual cues: How do they affect visual search? *Psychological Research, 73*, 212-221.

Meng, X-L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111*(1), 172-175.

Mordkoff, J. T., & Gianaros, P. J. (2000). Detecting the onset of the lateralized readiness potential: A comparison of available methods and procedures. *Psychophysiology*, *37*(3), 347–360.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology, 4*(2), 61-64.

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.

Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics*, *63*(7), 1239–1249.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Schankin, A., Hagemann, D., & Schubö, A. (2011). Is contextual cueing more than the guidance of visual-spatial attention? *Biological Psychology, 87*, 58-65.

Schankin, A., & Schubö, A. (2009). Cognitive processes facilitated by contextual cueing: Evidence from event-related brain potentials. *Psychophysiology*, *46*(3), 668–679.

Schankin, A., & Schubö, A. (2010). Contextual cueing effects despite spatially cued target locations. *Psychophysiology*.

Summerfield, J. J., Lepsien, J., Gitelman, D. R., Mesulam, M. M., & Nobre, A. C. (2006). Orienting attention based on long-term memory experience. *Neuron, 49*, 905-916.

Tseng, Y-C., & Li, C-S. R. (2004). Oculomotor correlates of context-guided learning in visual search. *Perception & Psychophysics, 66*(8), 1363-1378.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495–501.

Yantis, S. & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 601-621.

Zang, X., Jia, L., Müller, H. J., & Shi, Z. (2015). Invariant spatial context is learning but not retrieved in gaze-contingent tunnel-view search. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(3), 807-819.

Zellin, M., Conci, M., von Mühlenen, A., & Müller, H. J. (2011). Two (or three) is one too many: testing the flexibility of contextual cueing with multiple target locations. *Attention, Perception, & Psychophysics, 73*, 2065–2076.

Zellin, M., von Mühlenen, A., Müller, H. J., & Conci, M. (2013). Statistical learning in the past modulates contextual cueing in the future. *Journal of Vision, 13*(3):19, 1–14.

Zhao, G., Liu, Q., Jiao, J., Zhou, P., Li, H., & Sun, H-J. (2012). Dual-state modulation of the contextual cueing effect: Evidence from eye movement recordings. *Journal of Vision, 12*(6):11, 1-13.