Distractor rejection in parallel search tasks takes time but does not benefit from context repetition

Gavin Jun Peng Ng, Simona Buetti, Sanda Dolcos, Florin Dolcos, & Alejandro Lleras

University of Illinois at Urbana-Champaign

Word Count: 7139 (excluding References)

Number of Figures: 3

Number of Tables: 2

Running title: Lure processing and context repetition

Abstract

When the spatial configuration of a search display is presented repeatedly, response times to finding the target within that configuration are shorter compared to completely novel configurations, even though observers do not have explicit recognition of the repetition. This phenomenon is known as Contextual Cueing and selective attention is thought to be necessary for the effect. Previous research has suggested that repetition of the context of unattended items does not appear to improve performance; only repetition of attended items does. It has been proposed that this occurs because unattended items are pre-attentively filtered and thus do not contribute to performance. Here we demonstrate that so-called “unattended” items do contribute to performance, just not to contextual cueing. We approach this question from the perspective of the parallel processing of the scene that unfolds at the start of each item and that has been recently modeled by the Target Contrast Signal Theory. We show that the processing time per item during parallel evaluation of the scene is not affected by context repetition, suggesting that the locations of the items rejected in this stage are not integrated into the memory representation underlying contextual cueing. Other alternatives are also discussed.

Keywords: contextual cueing, visual search, target contrast signal theory

Introduction

Contextual cueing is a phenomenon in visual search tasks whereby response times are faster in displays where contextual information is repeated compared to novel displays. As the name suggests, there are two elements in Contextual Cueing. *Context* refers to information that co-occurs with the target, such as the spatial layout (Chun & Jiang, 1998), identity (Chun & Jiang, 1999; Goujon, Didierjean, & Marmèche, 2007), or motion trajectory (Chun & Jiang, 1999) of the stimuli in the search display. *Cueing* refers to the guidance of attention. Contextual cueing thus refers to situations in which attention is guided by the contextual information that has been learned. Most often, the context is the spatial layout of the stimuli in the search display (Chun, 2000; Lleras & Von Mühlenen, 2004; Rosenbaum & Jiang, 2013). In these experiments, half of the search displays are repeated (i.e. the spatial layout of the search stimuli remains identical throughout the experiment), while the other half are novel (i.e. the spatial layout is randomly generated each time). Response times to repeated displays are faster than to novel displays, suggesting that observers learn the context, which allows their attention to rapidly move to the target location.

Selective attention is thought to be necessary for the development of Contextual Cueing (Jiang & Chun, 2001). In a series of experiments, observers were tasked to search for a target with a pre-defined color among distractors of the same color ("candidates") or a very different color ("lures")[1]. When the spatial layout of the candidates was repeated and that of the lures was

[1] Jiang and Chun (2001) referred to these as the "attended" and "unattended" color respectively. However, we prefer to use the term "candidate" to describe a distractor that is very similar to the target such that selective attention is required to distinguish it from the target, and the term "lure" to describe a distractor that is sufficiently different from the target such that the visual system can distinguish it from the target in peripheral vision without the need for focused selective attention (Buetti et al., 2016; Lleras, Buetti, & Mordkoff, 2013; Neider & Zelinsky, 2008)

random, Contextual Cueing was observed even though only half the context was repeated. Furthermore, in the reverse scenario when the spatial layout of the lures repeated but the one for the candidates was random, evidence for Contextual Cueing was mixed. Specifically, in Experiment 2, no Contextual Cueing by the repeated lure configuration was found, whereas in Experiment 3, only a small effect was observed. Finally, in Experiment 4 when the search was made harder, repeating the lure context once again failed to produce a Contextual Cueing effect. Jiang and Chun (2001) thus argued that Contextual Cueing is dependent on selective attention. Importantly, the authors argued that the attentional process in these tasks with both candidates and lures is markedly different from the standard Contextual Cueing paradigm where all distractors are candidates. They argued that in candidate-only displays, candidates are attended to by selective attention before being rejected (e.g. Duncan & Humphreys, 1989; Jiang & Chun, 2003; Treisman & Sato, 1990). On the other hand, in mixed displays, lures are first filtered by a pre-attentive process (Palmer, 1995) since they differ from the target on at least one salient feature. Selective attention then evaluates and rejects the candidates. This process of selection and rejection, which does not take place for lures, is proposed to be the locus of the Contextual Cueing effect. That said, other studies found that lure-context repetition can facilitate search (Geyer, Zehetleitner, & Müller, 2010; Harris & Remington, 2017; Kunar, Flusberg, Horowitz, & Wolfe, 2007). Thus, evidence regarding the contribution of lure context to Contextual Cueing is mixed.

The goal of the present study is to re-examine the role of lure-context repetition on Contextual Cueing, given the somewhat mixed results in Jiang and Chun’s (2001) study, as well as the conflicting evidence from other studies (e.g. Geyer, Zehetleitner, & Müller, 2010; Harris & Remington, 2017; Kunar, Flusberg, Horowitz, & Wolfe, 2007). One difficulty in Jiang and

Chun's (2001) design arises from the fact that when lure-context was repeated, displays also contained candidates whose locations were not repeating. Thus, the entire display context was varying on each presentation. Further, because candidates attract focused attention (and lures do not), varying the locations of the candidates might overwhelm whatever guiding influence the lure-context might be exerting on attention. This might explain why in their Experiment 3, Jiang and Chun found the greatest Contextual Cueing effect in a condition where *both* candidate- and lure-context repeated. Another difficulty of the design was the use of lures that were letter-like (Ls). While it is tempting to say that Ls of a non-target color are unattended, there is strong evidence that letters are compulsory stimuli (Teichner & Krebs, 1974), and as such, focused attention to these stimuli might be stronger than to lures that are not letters. Furthermore, no direct evidence was provided that the lures were unattended and in fact, by their shape, the letter lures were very similar to the candidates and the target. Finally, Jiang and Chun's result that when search was hard (Experiment 4), lure-repetition did not impact Contextual Cueing is not conclusive. Indeed, such a null result is also consistent with the idea that the region of the display that is processed in parallel is reduced to a small region around fixation (Hulleman & Olivers, 2017; Lavie, 2010; Lleras, Wang, Ng, Ballew, & Buetti, revise and resubmit) . In sum, it is difficult to make strong conclusions at this point regarding the possible contribution of repeating lure-contexts to reaction time.

A second motivation for the current study comes from a recent mathematical model of parallel processing that proposes that the rejection of lures is not a pre-attentive process of filtering but instead an active process of rejection that takes time (Buetti, Cronin, Madison, Wang, & Lleras, 2016), that has often been overlooked. The Target Contrast Signal Theory provides a framework to quantify the processing cost of lures that are rejected in parallel through peripheral vision (Lleras et al., revise and resubmit). According to the theory, processing of a scene begins with a parallel accumulation of evidence at *all* locations across the display. The evidence accumulated is a contrast signal between information at each location and the target template. Evidence accumulates toward a ‘non-target’ threshold. Locations that reach this threshold are rejected in parallel; items at these locations will not be selected by focused attention for further processing. The evidence accumulation process is stochastic, and results in a logarithmic increase in response times as a function of both set size and lure-target similarity (Buetti et al., 2016; Townsend & Ashby, 1983). Note that this is in contrast to a pre-attentive filtering process as suggested by Jiang and Chun (2003; Palmer 1995), where no meaningful increase in response times as a function of set size is supposed to take place. Focused attention (and/or an eye movement) is directed towards locations that have not reached the non-target threshold after a certain period of time has passed without any accumulators reaching threshold. In sum, according to the Target Contrast Signal Theory, lures do not get filtered out en masse by a pre-attentive process. Rather, all items in a search display, including lures, are processed in parallel, producing significant and systematic processing costs. Recent work also confirmed that the theory is successful at predicting search performance in novel heterogeneous displays based on search performance in simple homogeneous displays, across participants (for displays with geometric shapes see Lleras, Wang, Madison, & Buetti, 2019; for displays with real-world objects see Wang, Buetti, & Lleras, 2017).

Given this new understanding regarding how lures are rejected, it is important to re-examine the effect of lure-context repetition. Here, we used lure stimuli that are clearly different from candidates (as in Buetti et al., 2016) and we also manipulated lure set size to better characterize the lure rejection process in the presence of context repetitions. As we will review in the General Discussion, using stimuli that are clearly not the target (i.e., lures) is crucial for getting clear evidence from the experiments regarding the contribution of these stimuli to performance (not just Contextual Cueing). Having “unattended” stimuli that are too similar to the target (as has been done before) may lead to a mix of results, depending on the position of those items in the display: items close to fixation might be discarded in parallel, whereas items in the periphery might not, or not consistently so. Thus, stimuli selection is critically important to avoid mixed results. Further, a fine manipulation of lure set size will also help us have a good and stable estimate of the lure-rejection process (Buetti et al., 2016) and evaluate whether or not this process is impacted by context repetition.

Experiment 1 is an initial investigation to select the candidate set size in the subsequent experiments. Experiment 2 will examine context-repetition effects in displays where both candidates and lures are repeated, and Experiment 3 will evaluate context-repetition effects in lure-only displays. Finally, in the Contextual Cueing literature, there is ample evidence that participants are at chance at identifying the displays that were repeated throughout the experiment (e.g., Chun & Jiang, 1998; Colagiuri & Livesey, 2016; Jiang & Chun, 2003; but see Vadillo, Konstantinidis, & Shanks, 2015). That said, we were interested in investigating whether participants would have a better memory for the lures displays, reasoning that perhaps this relatively easy to process displays might be more memorable (Khosla, Raju, Torralba, & Oliva, 2015). All code and data can be found on the Open Science Framework (Ng, Buetti, Dolcos, Dolcos, & Lleras, 2019)

Experiment 1

The goal of this experiment was to identify what is a small number of candidates that, when repeated, can produce reliable Contextual Cueing effects. This is necessary because when lures are introduced into the displays in subsequent experiments, the set size of lures will be varied across a wide range. The manipulation of lure set size across a wide range of values is necessary to observe the logarithmic increase in RT as a function of set size that indexes parallel evidence accumulation processes (see Buetti et al., 2016 for details).

Methods

Participants

All participants were recruited from the subject pool from the University of Illinois at Urbana-Champaign. Participants were given course credit for taking part in the experiment. All participants were tested with the Ishihara color plates and determined to be non-colorblind. All participants also had normal or corrected-to-normal vision. We aimed to collect 20 participants per experiment.

Stimuli and apparatus

The target was a red letter T that was rotated 90 degrees either clockwise or anti-clockwise. Participants had to respond to the target orientation. The distractor stimuli were red letter ‘L’s rotated either 0, 90, 190, or 270 degrees clockwise. The experiment was programmed and ran in MATLAB using the Psychophysics Toolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007).

In each search display, the stimuli were distributed across a 6-by-6 grid. There was a total of 12 displays (six for set size 4 and six for set size 8) that were repeated throughout the entire experiment. Each of these 12 repeated displays had unique target locations. A separate set of 12 target locations, which did not overlap with those in the repeated displays, was randomly selected and served as the target locations for the novel displays. This was done to equate target location probability between the repeated and novel displays. The novel displays were never repeated and were checked against the repeated displays to ensure that there were also no repeats. All stimuli were presented against a 1024 x 768 pixel black background on a 22-inch (400mm x 300mm) cathode ray tube monitor with a refresh rate of 85Hz. Participants viewed the display, unrestrained, from a distance of approximately 60cm.

Design and procedure

There were two within-subject independent variables: display type (repeat or novel) and set size (4 or 8). Participants viewed 25 blocks of 24 trials each, for a total of 600 trials. In each block, half the trials were repeated while the other half was novel. Within the repeated and novel trials, half were set size 4 while the other were set size 8. There was thus a total of 4 cells, with 150 trials in each cell. Trial order was randomized.

The participants' task was to find the oriented T target and report whether it was pointing to the right or to the left. Before the start of the experiment, participants were also presented with the following instructions aimed at improving the likelihood of obtaining Contextual Cueing (Lleras & Von Mühlenen, 2004):

The best strategy for this task, and the one that we want you to use in this study, is to be as receptive as possible and let the unique item "pop" into your mind as you look at the screen. The idea is to let the display and your intuition determine your response. Sometimes people find it difficult or strange to tune into their "gut feelings", but we would like you to try your best. Try to respond as quickly and accurately as you can while using this strategy. Remember, it is very critical for this experiment that you let the unique item just 'pop' into your mind.

The experiment began with a block of six practice trials to familiarize the participant with the experiment and to emphasize the passive instructions. Recording of experimental data began after the practice block after the participant acknowledged that they were ready to begin by pressing the left or right arrow key. Each trial began with a fixation cross that lasted 1000ms, after which the search display appeared. Participants had 5 seconds to respond to the identity of the target by pressing the left arrow when the T was rotated to the left (90 degrees anti-clockwise), or the right arrow when the T was rotated to the right (90 degrees clockwise). A loud beep was provided when an incorrect response was made or when no response was made after 5 seconds; no feedback was provided for correct responses. Each trial terminated when a response was made, or when 5 seconds passed without any response. A blank screen was then presented for 1 second before the next trial.

At the end of the experiment, participants were presented with two questions. They were first asked: “Some of the trials had the same arrangement of objects in the display. Did you notice?”. Participants responded either “yes” or “no”. After which, the next question was presented: “"What proportion of trials do you think had repeated spatial arrangements?”. Participants responded by entering a number between 0 and 100 on the keyboard.

Results

Analyses for all experiments were conducted in R (R Development Core Team, 2008). For each participant, response times (RTs) beyond 2.5 standard deviations of each condition were excluded from analyses. Trials on which participants made an error were also excluded. This led to the removal of 4.4% of trials. No participants had to be replaced in this experiment.

To increase the signal-to-noise ratio, the data was split into 5 epochs of 5 blocks each, with a total of 30 trials per cell per epoch (Jiang & Chun, 2003). A 2 (display type) by 2 (set size) by 5 (epoch) repeated measures ANOVA on RT was performed. RTs were slower for novel (M = 837 ms, SD = 192 ms) compared to repeated (M = 799 ms, SD = 164 ms) displays, $F(1, 19) = 5.09$, $p = .0361$, $\omega_p^2 = .163$. RTs were faster with four (M = 723ms, SD = 121ms) compared to eight (M = 914ms, SD = 178ms) candidates, $F(1, 19) = 126.81$, p <.001, $\omega_p^2 = .857$. Lastly, RTs decreased as a function of epoch (from epoch 1 to 5: M = 863, 832, 819, 794, 782 ms, SD = 197, 188, 180, 164, 158 ms), $F(4, 76) = 10.48$, p_c = .0033, $\varepsilon = 0.472$, $\omega_p^2 = .319$. Three interactions were marginally significant, suggesting that the design could be under-powered to detect these more subtle effects: display by epoch, $F(4, 76) = 2.38$, $p = .0593$, $\eta_p^2 = .111$; set size by display, $F(1, 19) = 3.47$, $p = .0779$, $\eta_p^2 = .154$. The interaction of set size by epoch was not significant, $F(4, 76) = 1.66$, $p = .169$, $\eta_p^{22} = .0803$, and the three-way interaction between set size, display, and epoch, $F(4, 76) = 0.55$, $p = .702$, $\eta_p^2 = .0281$.

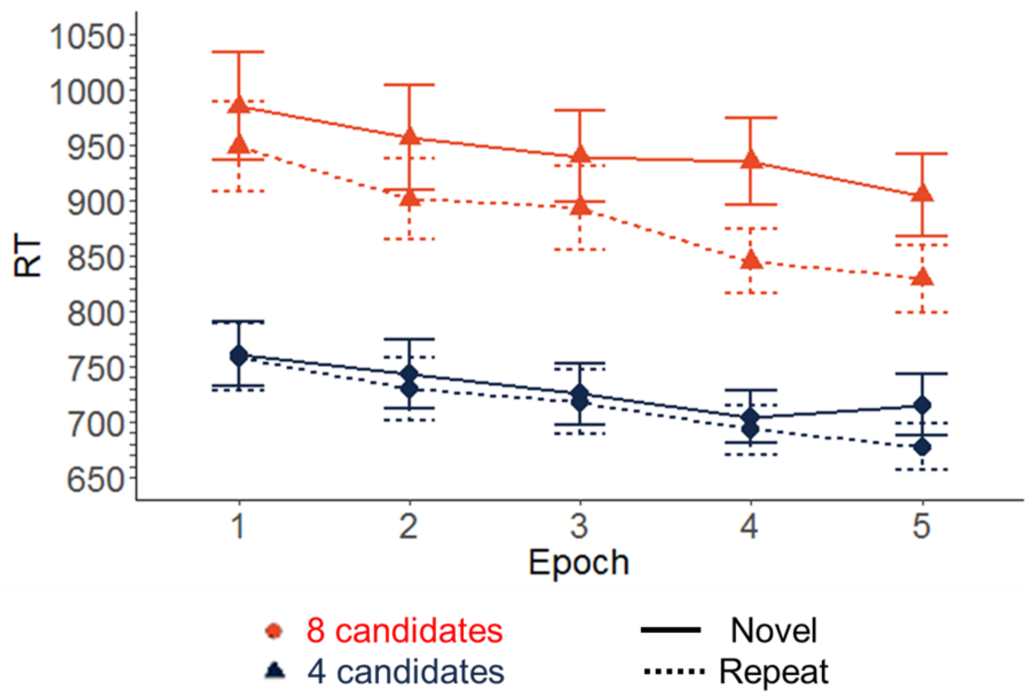

Figure 1. Response times for novel displays (solid lines) were significantly longer than for repeated displays (dashed lines), for both set size 8 (orange circles) and set size 4 (black triangles) in Experiment 1. The average magnitude of the Contextual Cueing effect was larger in set size 8 (61ms) compared to set size 4 (15ms).

At the end of the experiment, 25% of the participants responded “yes” when they were asked whether they noticed that some displays were repeated throughout the search experiment. On average, these participants estimated that 32.4% of the displays were repeated.

In sum, Experiment 1 showed that while the Contextual Cueing effect was observed both with 4 and 8 candidates. The effect was greater with 8 candidates compared to 4 candidates (61 vs. 15ms on average) and thus we decided to use 8 candidates in Experiment 2, to maximize the chances of obtaining a Contextual Cueing effect driven by the repetition of the candidates, so that we could examine whether the magnitude of this effect is impacted by the repetition of the lures.

Experiment 2

The goal of Experiment 2 was to examine the effect of lure-context repetition on Contextual Cueing. In Experiment 2 we used 8 candidates on every display, while varying the number of lures. In contrast to Jiang and Chun (2001), the lure stimuli we used were simple geometric shapes (orange diamonds) that were neither letter-like nor similar in shape to the candidates. Buetti et al. (2016) showed that these lures are rejected in parallel when searching for a T target amongst L candidates (in novel displays), producing logarithmic processing costs. Importantly, in old displays, both candidate and lure stimuli were repeated. If lures are processed in parallel in repeated displays, as proposed by the Target Contrast Signal Theory, response times should increase logarithmically as a function of lure set size (Buetti et al., 2016). In contrast, if lures are filtered out by a pre-attentive process, then there should be no effect of lure set size on response times (Duncan & Humphreys, 1989; Palmer, 1995; A. Treisman & Sato, 1990).

From the perspective of the Target Contrast Signal Theory, it is unclear what the fate of lures that are rejected in parallel is with regard to Contextual Cueing. On the one hand, since lures undergo an active process of evidence accumulation, their locations might be implicitly learned and form part of the spatial context that determines Contextual Cueing. If this is the case, searching through repeated contexts should be faster (more efficient) than searching through novel displays, and thus search should be more efficient in repeated displays (leading to smaller search slopes). That is to say, Contextual Cueing would be determined by the entire list of locations where evidence initially accumulated. On the other hand, since lures are discarded prior to attentional scrutiny of individual items, it is also possible that they will not contribute to Contextual Cueing. That is to say, what determines the "context" in Contextual Cueing might be the list of candidate locations only (where accumulators did not reach the non-target threshold). This latter possibility is consistent with Jiang and Chun's (2001) proposal that Contextual Cueing is determined by the set of locations that are "selectively attended". Finally, from the perspective of the Target Contrast Signal Theory, it is important to evaluate whether or not the parallel lure rejection process would improve when lure context is repeated.

Method

Twenty participants were recruited from the same subject pool as Experiment 1. These participants did not take part in any of the other experiments in this paper. All methods are identical to Experiment 1, except for the following changes: in addition to the candidate Ls, there were symmetric orange diamonds (lures). Each display always contained 8 candidates. There were 4 different lure set sizes: 0, 5, 10, 20. As such, there were 3 repeated displays per set size and a total of 12 repeated displays throughout the entire experiment. In the repeated displays, both lures and candidates, as well as the target, was always in the same location.

Results

For each participant, response times beyond 2.5 standard deviations of each condition were excluded from analyses. Trials on which participants made an error were also excluded. This led to the removal of 3.4% of trials. No participants had to be replaced in this experiment.

A 2 (display type) by 4 (lure set size) by 5 (epoch) fully within ANOVA was performed. RTs for novel (*M* = 990 ms, *SD* = 174 ms) displays were slower than repeated (*M* = 953 ms, *SD* = 178 ms) displays, $F(1, 19) = 4.64$, $p = .0443$, $\omega_p^2 = .148$. RTs increased with lure set size (set size 0, 5, 10, 20, respectively: *M* = 916, 962, 978, 1030 ms, *SD* = 152, 183, 167, 185 ms), $F(3, 57) = 9.68$, $p < .001$, $\omega_p^2 = .299$, RTs decreased as a function of epoch (from epoch 1 to 5: *M* = 1042, 981, 947, 953, 935 ms, *SD* = 189, 164, 165, 172, 174 ms), $F(4, 76) = 12.17$, $p_c < .001$, $\varepsilon = .585$, $\omega_p^2 = .356$. The main effects were qualified by a significant display type by epoch interaction, $F(4, 76) = 6.80$, $p < .001$, $\omega_p^2 = .223$. Follow-up paired t-tests revealed that RTs for repeated displays were faster than for novel displays only in the last two epochs (Table 1), after adjusting the p-value to .01 (= .05/5, Bonferroni correction). The interactions between set size and display and between set size and epoch were not significant, $F(3, 57) = 0.24$, $p = .868$, $\eta_p^2 = .0125$; and, $F(12, 228) = 1.12$, $p = .341$, $\eta_p^2 = .0557$, respectively. Importantly, the three-way interaction between display type, set size and epoch was not significant, $F(12, 228) = 1.34$, $p = .198$, $\eta_p^2 = .0659$.

Table 1. Follow-up t-tests for the significant display type by epoch interaction in Experiment 2. Asterisks indicate statistical significance at p < .01 (after Bonferroni correction).

Epoch	$t(79)$	*P*	*M* (SD)	Cohen's d_z
1	-0.16	.877	- 3 (191)	-0.018

2	1.42	.161		23 (142)	.159
3	2.01	.0482		36 (162)	.225
4	4.65	<.001	*	80 (154)	.52
5	2.88	.00507	*	51 (160)	.322

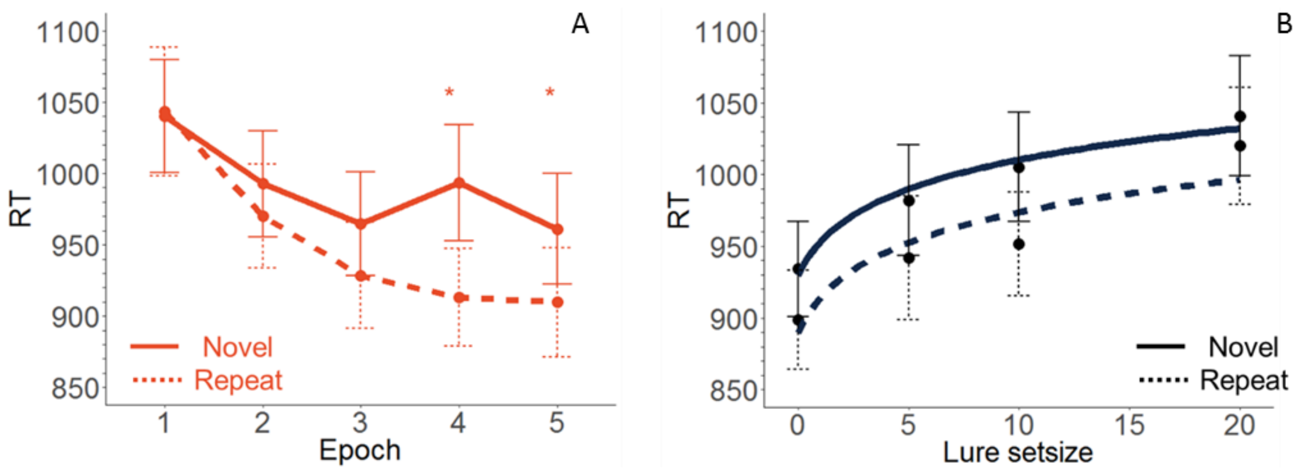

Figure 2. (A) Response times for novel displays (solid lines) were significantly longer than for repeated displays (dashed lines) in Epochs 4 and 5. *(B)* There was no significant difference in logarithmic slopes between novel (solid lines) and repeated (dashed lines) displays, indicating that there was no difference in search efficiency.

To follow up on the absence of a significant effect in the three-way interaction, we used a Bayes factor approach. To better characterize search efficiency, we fitted each subject's data with a logarithmic function. To determine whether the logarithmic RT by set size slopes were meaningfully different between the novel and repeated displays, the Bayes factor was calculated for a model with display type as a predictor. Bayes factors are preferred over null hypothesis testing when the goal is to provide evidence for null effects (Rouder, Speckman, Sun, Morey, & Iverson, 2009). The analysis revealed moderate evidence (Kass & Raftery, 1995) for the hypothesis that there was no meaningful difference in slopes between the novel and repeated displays, $BF_{01} = 3.22$. Search efficiency did not improve with repeated displays.

At the end of the experimental session, 30% of the participants responded “yes” when they were asked whether they noticed that some displays were repeated throughout the search experiment. On average, these participants estimated that 32.6% of the displays were repeated (the response from one participant was missing due to a keyboard malfunction).

In sum, Experiment 2 showed that the magnitude of the Contextual Cueing effect was not affected by lure set size, even though lures contributed to processing times and were clearly distinct from the target. The results also showed that lures were not filtered out pre-attentively: lures were processed in parallel (as indexed by the logarithmic functions) and, the more there were, the longer it took to reject all of them. Yet, the lure rejection process did not benefit from repeating the locations were lures were presented. If indeed lures do not contribute to Contextual Cueing despite being actively processed, then we should not observe any Contextual Cueing with lure-only displays. This prediction was tested in Experiment 3.

Experiment 3

Here, we tested the extent to which the repetition of the spatial configuration occupied by lures helps the lure rejection process by evaluating the processing cost of lures. In addition, this experiment also included a more precise memory test to assess whether participants had an explicit recollection of the repeated displays. Multiple previous studies have shown that memory for candidate-repeated contexts is implicit (e.g. Chun & Jiang, 1998; Goujon, Didierjean, & Thorpe, 2015; but see Vadillo et al., 2015), and we wanted to investigate whether participants had any memory traces of the repeated lure-only displays.

Method

Twenty-two participants were recruited from the same subject pool as previous experiments. One subject did not complete the experiment due to a computer error. One additional participant was replaced because their average RT that was more than 2.5 standard deviations higher than the group mean. The included participants did not take part in any of the other experiments in this paper. All methods are identical to Experiment 1, except for the following changes. All distractors were the orange diamond lures used in Experiment 2. There were 5 different lure set sizes (0, 3, 9, 19, 31). There were 13 instead of 12 repeated displays: an additional one was included for the target-only condition (lure set-size = 0). There were 3 repeated displays for each of the non-zero lure set sizes, and one for the target-only display (set size 0). Finally, there was also a recognition test at the end of the experiment.

The recognition test started by asking participants whether they noticed anything unusual with the experiment. If they answered ‘Yes’, they were prompted to describe it by using the keyboard. Regardless of whether they answered ‘yes’ or ‘no’, they were informed on the next screen that some of the displays were repeated, and were asked whether they had noticed this or not. After which, they were asked what percentage of the trials they thought were repeated. Participants were then informed that they would be presented with a recognition test.

The recognition test consisted of 104 trials in total. Each of the 13 repeated displays was presented 4 times, twice with the target rotated 90 degrees clockwise and twice anti-clockwise. Thirteen novel displays, which were never presented in the search task, were created. These novel displays were also presented 4 times, two with the target rotated 90 degrees clockwise and twice anti-clockwise. This was to equate for learning within the recognition task. The target location for these novel displays were the same as the target locations for the novel displays in the search task to equate for target probability. The recognition test was blocked such that each repeated and novel display was presented once before it was presented again.

On each trial, the display was presented until a response was made. Participants pressed the ‘z’ key to indicate whether they had seen the display before during the search task, and the ‘/’ key to indicate that the current display was a novel one. Upon response, a confidence rating screen was presented. Participants had to indicate their level of confidence in their response, ranging from 1 (‘completely guessing’) to 5 (‘completely confident’). Upon response, a blank screen was presented for a randomly selected duration between 600 and 800 ms before the next trial began.

Results

One participant was excluded from analyses as they had a mean RT that was more than 2.5 standard deviations away from the overall group mean. An additional participant was run to replace this subject. Response times beyond 2.5 standard deviations of the mean of each participant were excluded from analyses. Trials on which participants made an error were also excluded. This led to the removal of 4.5% of trials.

A 2 (display type) by 5 (lure set size) by 5 (epoch) within-subjects ANOVA was performed. RTs increased with lure set size (set size 0, 3, 9, 19, 31, respectively: M = 507, 528, 548, 561, 570 ms, SD = 110, 80, 77, 84, 83 ms), $F(4, 76) = 31.88$, $p < .001$, $\omega_p^2 = .604$. RTs for novel (M = 545 ms, SD = 91 ms) displays were not significantly different from repeated (M = 541ms, SD = 90ms) displays, $F(1, 19) = 0.74$, $p = .401$, $\omega_p^2 = -.0125$. RTs did not differ significantly between epochs (from epoch 1 to 5: M = 560, 538, 535, 546, 536 ms, SD = 95, 79, 76, 103, 94 ms), F(4, 76) = 2.17, $p_c = .128$, $\varepsilon = .5$, $\omega_p^2 = .0546$.

There was no significant interaction between set size and display, $F(4, 76) = 0.72$, $p_c = .505$, $\varepsilon = .554$, $\eta_p^2 = .0365$; set size and epoch, $F(16, 304) = 0.555$, $p = .71$, $\varepsilon = .279$, $\eta_p^2 = .0394$; and epoch and display, $F(4, 76) = 1.15$, $p = .342$, $\eta_p^2 = .0571$. The three-way interaction between display type, set size and epoch was also not significant, $F(16, 304) = 0.553$, $p_c = .917$, $\varepsilon = .262$, $\eta_p^2 = .0283$.

The lack of a significant difference in response times between repeated and novel displays suggests that lure processing did not contribute to Contextual Cueing. The Bayes factor for a model with display type as a factor indicated that there was *strong evidence* (Kass & Raftery, 1995) that the response times for repeated and novel displays were not meaningfully different, $BF_{01} = 10.961$.

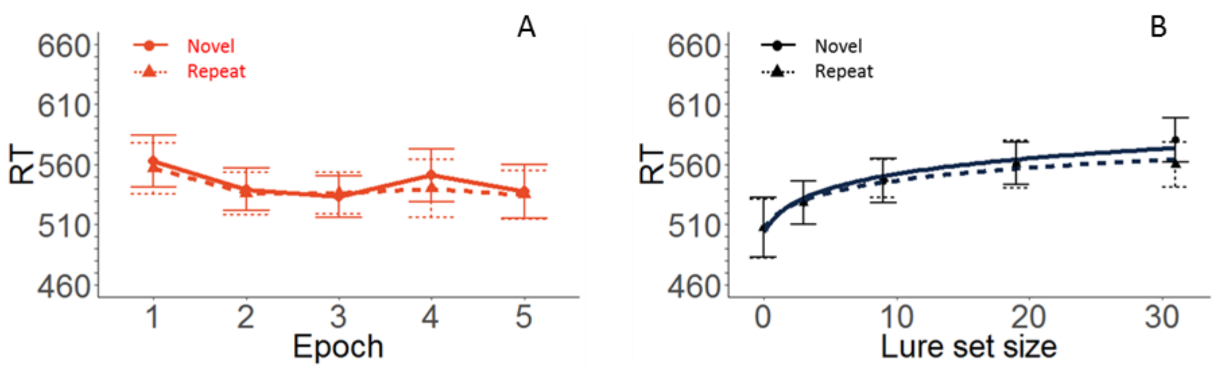

Figure 3. Response times (in ms) as a function of lure set size for Experiment 3. Repeated displays are represented by the solid line, while novel displays are represented by the dashed line. *(A)* Response times did not significantly differ with epoch or display type. *(B)* Although response times increased logarithmically as a function of set size, suggesting that lures were processed, there was no statistically significant Contextual Cueing effect. Error bars indicate the standard error of mean.

At the end of the experimental session, only one participant responded “yes” when asked whether they noticed anything strange about the experiment but did not elaborate. This participant estimated that 2% of the displays were repeated.

We next turn to the results of the recognition test. First, one-sample t-tests were conducted for each of the lure set sizes (0, 3, 9 ,19, 31) to determine whether *d’* was significantly different from zero, considering recognition performance across the four memory blocks (Table 2). Furthermore, the same analysis was conducted on the first block only. It is indeed possible *d’* decreased throughout the testing as participants became confused as to whether they recognized the displays from the search task or the previous presentations during the recognition test (Table 3). The results indicated that none of the comparisons were statistically significant after adjusting the p=value to $\alpha = .005$ (Bonferroni correction), suggesting that participants did not report any conscious awareness of the repeated displays.

Table 2. d’ for the recognition test in Experiment 3, collapsed across the four blocks (top) and for the first block only (bottom). All *d’*s were not significantly different from zero after adjusting the p-value significance level by the number of comparisons using Bonferroni Correction, suggesting that there was no conscious awareness of the repeated displays.

Overall recognition call performance (four presentations)				
Lure set size	*t*(18)	P	*d’*	Cohen’s d_z
0	2.21	.0407	.457	.507
3	-0.73	.478	-0.0709	-.166
9	-0.27	.794	-0.02	-.061
19	-0.16	.874	-0.0168	-.0369
31	1.86	.08	0.127	.427

Recognition performance during the first presentation				
Lure set size	*t*(18)	P	*d'*	Effect size
0	3.02	.0073	.484	.694
3	-1.42	.172	-0.132	-.326
9	-1.18	.253	-.102	-.271
19	-0.38	.71	-0.0528	-.0872
31	0.039	.969	0.00558	.00895

Finally, we also examined the possibility that some displays were explicitly recognized. Out of the total two hundred and twenty eight lure-context displays that were repeated throughout the experiment across all nineteen participants (twelve lure-context displays times 19 participants), forty displays (17.5%) had a perfect accuracy score in the recognition test: fifteen participants indicated that they had seen at least one display during the search task on all four presentations during the recognition test. The mean confidence rating for these displays was 3.59 out of 5 (the average for all other repeated displays was 2.94). This seems to suggest that there is at least some percentage of displays that were explicitly recognized by participants. The Bayes Factors were calculated for a model with accuracy score as a factor. There was strong evidence for the hypothesis that the magnitude of the contextual cueing effect was not affected by accuracy scores on the recognition test, BF01 = 14.352.

Between-experiment comparison on the awareness of repeated displays

The percentage of participants who reported awareness of repeated displays in Experiments 1,2, and 3 were 25%, 30%, and 5% respectively. A chi-square test of independence revealed that noticing rates were not significantly different across the three experiments, $\chi 2(3, N= 59) = 4.09, p = .13$.

General discussion

The goal of the present study was to evaluate whether repeating lure contexts over time would produce a similar Contextual Cueing effect to that observed when candidate contexts repeat. The lure stimuli used in the present study were clearly distinguishable from the target. Experiments 2 and 3 showed converging evidence that lure-context repetition does not in fact contribute to Contextual Cueing. These results are consistent with the findings in Jiang and Chun (2001). That said, in spite of the fact that lure-context repetition does not produce a Contextual Cueing effect, what is novel in our study is that we have evidence that lures *were* indeed processed and produced significant costs on reaction time (about 50-100 ms, comparing the zero lure condition to the largest lure set size condition; see Figure 2 and 3). Thus, this finding goes counter theories that assume lures are filtered out pre-attentively, or “unattended”, and thus rejecting them carries no processing cost. In terms of the mechanistic locus of Contextual Cueing, the results imply that Contextual Cueing emerges late, after lures have been discarded from a scene, in what is often referred to as the second stage of visual search (e.g. Treisman & Gelade, 1980; Wolfe, 2006). From the perspective of parallel peripheral processing, *all locations* in the display are processed initially, yet it is not this set of locations that produces the effect. Only the subset of locations that are not rejected during parallel processing will form the memory basis that leads to the facilitation in Contextual Cueing. Just as Jiang and Chun (2001) suggested, contextual information is learned only for distractors that undergo processing by selective attention.

Coupled with the fact that in typical Contextual Cueing experiments, memory for repeated displays is only implicit, the current results suggest that Contextual Cueing might be a form of procedural knowledge: given a set of locations to attend to (the list of non-rejected locations), the visual system is faced with a series of decisions regarding the order in which these locations are to be inspected by the eyes/attention. Thus, the benefit of repetition might be to lessen the demands on this procedural decision process and to improve performance as the same set of contexts repeat over and over throughout the experiment. If so, this mechanism might also help explain why there is, at best, a minimal guidance effect of context repetition in Contextual Cueing (Kunar et al., 2007). There have only been a few studies where the set size of the candidate set is manipulated (Chun & Jiang, 1998; Kunar et al., 2007; Kunar, Flusberg, & Wolfe, 2008). If display repetition guided attention towards the target, then one would expect an interaction between display repetition and set size (a smaller set size effect on repeated displays than on novel displays), which has not been consistently observed (or only to a small extent). Thus, perhaps Contextual Cueing is less an attentional effect and more a procedural memory effect: the advantage that comes from repeating the same actions/decisions over time, just like repetition improves playing the same musical piece on the piano.

It should be noted, though, that several studies have found evidence of Contextual Cueing in lure-only displays. However, it is unclear whether these studies truly reflect the influence of lure-context repetition. In Geyer, Zehetleitner, and Müller (2010), the search display was preceded by placeholders which previewed the spatial locations before the distractors appeared. The distractors were always green bars rotated 45 degrees to the right, while the target could be either a red bar rotated 45 degrees to the right or a green bar rotated 45 degrees to the left. The target was randomly defined on each trial. Thus, the search process in this scenario would be different from a typical Contextual Cueing experiment. There would not be a process by which lures are first rejected; the context that is learned in this case would be the spatial layout of the placeholders, rather than the layout of the lures. Furthermore, the search task could be considered as an oddball search, instead of search with a fixed target. We have previously shown that the lure-rejection process in oddball search is very different than in fixed-target search tasks (Buetti et al., 2016).

Other studies reported very small Contextual Cueing effects in lure-only displays (12 - 33ms; Harris & Remington, 2017; Kunar et al., 2007). In Experiment 2b in Kunar et al. (2007), lures (either 8 or 12) were always accompanied by placeholders. The target was a letter T and was red in color, as was the placeholder it appeared in. Lures were green letter Ls placed in green placeholders. Although they reported a statistically significant Contextual Cueing effect with 12 lures, the main effect of display repetition was not statistically significant, and the Contextual Cueing effect for set size 8 was only marginally significant ($p = .09$) after collapsing across the last 3 epochs. Thus, the difference across studies could arise either from a false positive result in Kunar et al. (2007) or from a lack of power to detect a small effect in the present study. However, we believe the latter possibility is less likely than the former. First, we had almost twice as many participants than in Kunar et al. (20 vs. 12). Second, Bayes factors analyses in our experiments revealed strong evidence for the lack of any effect of lures on Contextual Cueing. Therefore, we believe it is more likely that the Contextual Cueing effect detected in Kunar et al. (2007) was a false positive. A final possibility could be that some of the lures may have acted as candidates instead. The categorization of a stimulus as a lure or candidate is determined by whether the visual system can differentiate it from a target in the periphery. It is possible that some of the lures functioned as candidates, especially when they were in the far periphery, and thus contributed a small Contextual Cueing effect.

Considerations for future studies

There has been considerable debate on whether the Contextual Cueing effect arises from an implicit memory trace of the repeated displays (Colagiuri & Livesey, 2016; Goujon et al., 2015; Schlagbauer, Muller, Zehetleitner, & Geyer, 2012; Vadillo et al., 2015). The results at the recognition task in Experiment 3 indicated no overall memory for the repeated-lure contexts, but at the same time, display-level analysis suggested perfect memory for a small subset of lure contexts. This finding suggests that future studies should consider more precise tests of the memory for repeated displays and a more careful analysis of whether or not explicit memory for a subset of displays can drive the overall Contextual Cueing effect.

An important implication of our findings is that the manner in which stimuli are rejected determines whether or not they contribute to the spatial context driving Contextual Cueing. When items can be rejected via peripheral vision through parallel processing, the locations of these items do not form part of that context. Those items that cannot be rejected in parallel in the periphery probably do contribute to that context. In other words, it is not the stimulus per se that matters, but the interaction between a stimulus and its location on the visual field: a stimulus that is somewhat similar to the target might *act* as a lure in the near periphery (where resolution is somewhat high), but it might act as a candidate farther in the periphery. As a result, experiments with such “lures” might be difficult to interpret because these peripheral “lures” wouldn’t be rejected in parallel, and thus their locations would be part of the list of to-be-scrutinized locations. If so, these locations would create the sort of spatial context that does facilitate repeated search. This scenario might be aggravated by other peripheral processing constraints like crowding. This hypothesis should be tested in future experiments.

Future experiments should also explore the possibility that lure locations are in fact stored in memory but are too slow to emerge or too weak to have an impact on efficient search. Suppose it takes 500 ms to recognize a repeated context (irrespective of whether it is composed of lures or candidates). When the repeated display only contains lures, by the time the context has been implicitly recognized by the visual system, the target has already been found. When the repeated display contains candidates, it takes longer to find the target; the recollection of the context thus has time to impact the deployment of attention and therefore facilitate search, producing a Contextual Cueing effect. That said, we doubt this account as it is inconsistent with the data of Jiang and Chun (2001), where the slower search task eliminated (rather than increased) the lure-context repetition effect.

Summary

In conclusion, the results show that lure processing does not benefit from context repetition, even though lure processing incurs a significant time cost. This suggests that lure locations are not stored in the memory trace that drives Contextual Cueing, consistent with the proposal by Jiang and Chun (2001). The results are also consistent with the Target Contrast Signal Theory, which proposes that locations containing lures are rejected early on, during parallel processing, and are not considered as targets for the attention and eye movement system. Only those locations that are likely to be fixated contribute to Contextual Cueing.

References

Brainard, D. H. (1997) The Psychophysics Toolbox, *Spatial Vision*, *10*, 433-436.

Buetti, S., Cronin, D. A., Madison, A. M., Wang, Z., & Lleras, A. (2016). Towards a better understanding of parallel visual processing in human vision: Evidence for exhaustive analysis of visual information. *Journal of Experimental Psychology: General*, *145*(6), 672–707. http://doi.org/10.1037/xge0000163

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170–178. http://doi.org/10.1016/S1364-6613(00)01476-5

Chun, M. M., & Jiang, Y. V. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*(1), 28–71. http://doi.org/10.1006/cogp.1998.0681

Chun, M. M., & Jiang, Y. V. (1999). Top-Down Attentional Guidance Based on Implicit Learning of Visual Covariation. *Psychological Science* , *10*(4), 360–365. http://doi.org/10.1111/1467-9280.00168

Colagiuri, B., & Livesey, E. J. (2016). Contextual cuing as a form of nonconscious learning: Theoretical and empirical analysis in large and very large samples. *Psychonomic Bulletin and Review*, *23*(6), 1996–2009. http://doi.org/10.3758/s13423-016-1063-0

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

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*(2010), 20. http://doi.org/10.1167/10.5.20

Goujon, A., Didierjean, A., & Marmèche, E. (2007). Contextual cueing based on specific and categorical properties of the environment. *Visual Cognition*, *15*(3), 257–275. http://doi.org/10.1080/13506280600677744

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences*, *19*(9), 524–533. http://doi.org/10.1016/j.tics.2015.07.009

Harris, A. M., & Remington, R. W. (2017). Contextual cueing improves attentional guidance, even when guidance is supposedly optimal. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(5), 926–940. http://doi.org/10.1037/xhp0000394

Hulleman, J., & Olivers, C. N. L. (2017). The impending demise of the item in visual search. *Behavioral and Brain Sciences*, *40*, 1–69. http://doi.org/10.1017/S0140525X15002794

Jiang, Y. V., & Chun, M. M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology Section A : Human Experimental Psychology*, (February 2014), 37–41.

Jiang, Y. V., & Chun, M. M. (2003). Contextual cueing: Reciprocal influences between attention and implicit learning. In *Attention and Implicit Learning* (pp. 277–296).

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Khosla, A., Raju, A. S., Torralba, A., & Oliva, A. (2015). Understanding and predicting image memorability at a large scale. *International Conference on Computer Vision*. http://doi.org/10.1109/ICCV.2015.275

Kleiner M, Brainard D, Pelli D, 2007, “What’s new in Psychtoolbox-3?” *Perception 36 ECVP Abstract Supplement*.

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. http://doi.org/10.1037/0096-1523.33.4.816

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. http://doi.org/10.1080/13506280701751224

Lavie, N. (2010). Attention, Distraction, and Cognitive Control Under Load. *Current Directions in Psychological Science*, *19*(3), 143–148. http://doi.org/10.1177/0963721410370295

Lleras, A., Buetti, S., & Mordkoff, J. T. (2013). *When do the effects of distractors provide a measure of distractibility? Psychology of Learning and Motivation* (Vol. 59). Elsevier. http://doi.org/10.1016/B978-0-12-407187-2.00007-1

Lleras, A., & Von Mühlenen, A. (2004). Spatial context and top-down strategies in visual search. *Spatial Vision*, *17*(4–5), 465–482. http://doi.org/10.1163/1568568041920113

Lleras, A., Wang, Z., Madison, A., & Buetti, S. (2019). Predicting search performance in heterogeneous scenes: Quantifying the impact of homogeneity effects in efficient search. *Collabra*, *5*, 1–15.

Lleras, A., Wang, Z., Ng, G. J. P., Ballew, K., & Buetti, S. (revise and resubmit). A contrast signal theory of parallel processing in goal-directed search. *Attention, Perception & Psychophysics*.

Neider, M. B., & Zelinsky, G. J. (2008). Exploring set size effects in scenes: Identifying the objects of search. *Visual Cognition*, *16*(1), 1–10. http://doi.org/10.1080/13506280701381691

Ng, G. J. P., Buetti, S., Dolcos, S., Dolcos, F., & Lleras, A. (2019, March 18). Distractor rejection in parallel search tasks takes time but does not benefit from context repetition. Retrieved from osf.io/zwxbh

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect.

Current Directions in Psychological Science, *4*(4), 118–123.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rosenbaum, G. M., & Jiang, Y. V. (2013). Interaction between scene-based and array-based contextual cueing. *Attention, Perception, & Psychophysics*, *75*(5), 888–99. http://doi.org/10.3758/s13414-013-0446-9

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review*, *16*(2), 225–237. http://doi.org/10.3758/PBR.16.2.225

Schlagbauer, B., Muller, H. J., Zehetleitner, M., & Geyer, T. (2012). Awareness in contextual cueing of visual search as measured with concurrent access- and phenomenal-consciousness tasks. *Journal of Vision*, *12*(11), 25–25. http://doi.org/10.1167/12.11.25

Teichner, W. H., & Krebs, M. J. (1974). Visual search for simple targets. *Psychological Bulletin*, *81*(1), 15–28. http://doi.org/10.1037/h0035449

Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press. http://doi.org/10.2307/1422636

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136. http://doi.org/10.1016/0010-0285(80)90005-5

Treisman, A., & Sato, S. (1990). Conjunction Search Revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(3), 459–478.

Vadillo, M. a., Konstantinidis, E., & Shanks, D. R. (2015). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*.

http://doi.org/10.3758/s13423-015-0892-6

Wang, Z., Buetti, S., & Lleras, A. (2017). Predicting search performance in heterogeneous visual search scenes with real-world objects. *Collabra*, *3*(1), 6. http://doi.org/http://doi.org/10.1525/collabra.53

Wolfe, J. M. (2006). Guided search 4.0. *Integrated Models of Cognitive Systems*, (3), 99–120. http://doi.org/10.1007/978-94-011-5698-1_30